CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	The exercise and diet-microbiome paradigm: influences of physical activity and dietary nutrition on the human gut microbiome
Authors	Barton, Wiley
Publication date	2018
Original Citation	Barton, W. 2018. The exercise and diet-microbiome paradigm: influences of physical activity and dietary nutrition on the human gut microbiome. PhD Thesis, University College Cork.
Type of publication	Doctoral thesis
Rights	© 2018, Wiley Barton. - http://creativecommons.org/licenses/by-nc-nd/3.0/
Download date	2024-05-02 10:16:39
Item downloaded from	https://hdl.handle.net/10468/7736

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

THE EXERCISE AND DIET-MICROBIOME PARADIGM: INFLUENCES OF PHYSICAL ACTIVITY AND DIETARY NUTRITION ON THE HUMAN GUT MICROBIOME

Presented to the National University of Ireland, Cork, in fulfilment of the requirements for the degree of Doctorate of Philosophy

A thesis in 1 volume

By

Wiley Barton, B.Sc.

Department of Food Biosciences, Teagasc Food Research Centre, Cork, Ireland
Head of Department: Dr Paul Cotter

Department of Medicine, National University of Ireland, Cork, Ireland
Head of Department: Professor Fergus Shanahan

Submitted December 2018

Committee of Supervision:

Dr. Orla O'Sullivan
PhD

Dr. Paul D. Cotter
PhD

Professor Fergus Shanahan
MD, DSc, FRCPI, FRCP (UK), FACP, FRCP

TABLE OF CONTENTS

Declaration of Original Work----------i

Acknowledgements----------iii

List of Abbreviations---------- ix

List of Visual Resources---------- xiii

Abstract ---------- 1

1.1 | Introduction 1
1.2 | Methods 1
1.3 | Results 2
1.4 | Conclusions 3

Chapter 1---------- 4

INTRODUCTION AND BACKGROUND

1.1 | Abstract 4
1.2 | Author's relevant publications 5
1.3 | Dissemination of author's research 7
1.3.1 | Conference participation 7
1.3.2 | Campus participation 8
1.4 | Introduction 9
1.5 | Purpose of research 10
1.6 | Thesis structure 10
1.7 | The human microbiome as it relates to metabolic function and health 12
1.7.1 | Membership 12
1.7.2 | Womb to Tomb 14
1.8 | The Human Body – a Microbiome Perspective 16
1.8.1 | The Skin 18
1.8.2 | The Mouth 19
1.8.3 | The Nose and Respiratory System 22
1.8.4 | The Vagina and Urinary Tract 24
1.8.5 | The Gut 27
1.9 | Systemic Implication of the Gut Microbiome in Health and Disease 31
1.9.1 | Short-chain fatty acids 31
1.9.2 | Bile acids 36
1.10 | Conclusions and Outlook 39
1.11 | Hypothesis 39
1.12 | Objective 40
1.13 | References 41

Chapter 2 ---------- 62

MATERIALS AND METHODS

2.1 | Ethical approval **62**
2.2 | Study compliance and withdrawal **62**
2.3 | Participant data collection **63**
2.3.1 | Diet & nutrition 63
2.3.2 | Physical activity 63
2.3.3 | Body composition 64
2.3.4 | Inflammatory cytokine measurement 64
2.4 | Exercise intervention **65**
2.5 | Extraction and sequencing of faecal microbiome DNA **67**
2.5.1 | Processing of faecal samples 67
2.5.2 | Metagenomic library preparation and sequencing 67
2.6 | Bioinformatic processing of microbial sequencing data **69**
2.7 | Metabolomic phenotyping **71**
2.7.1 | Sample preparation 71
2.7.2 | LC-MS Metabolic profiling analysis 73
2.7.3 | ^{1}H-NMR Metabolic profiling analysis 73
2.7.4 | GC-MS SCFA targeted analysis 74
2.7.5 | LC-MS data treatment 74
2.7.6 | ^{1}H-NMR data treatment 75
2.7.7 | GC-MS data treatment 75
2.7.8 | Metabolite ID 75
2.8 | Quantification and statistical analysis **76**
2.8.1 | Power analysis and sample size 76
2.8.2 | Clinical data 77
2.8.3 | Microbiome analysis 78
2.8.4 | Metabonomic Statistical and Bioinformatics Analysis 79
2.9 | References **81**

Chapter 3 ---------- 84

THE MICROBIOME OF PROFESSIONAL ATHLETES DIFFERS FROM THAT OF CONTROLS NOT ONLY IN COMPOSITION BUT PARTICULARLY AT THE FUNCTIONAL METABOLIC LEVEL

3.1 | Abstract **84**
3.2 | Introduction **86**
3.3 | Study objective **87**
3.4 | Methods **87**
3.4.1 | Study population 87
3.4.2 | Acquisition of clinical, exercise and dietary data 88
3.4.3 | Preparation of Metagenomic libraries 88
3.4.4 | Metagenomic statistical and bioinformatic analysis 88
3.4.5 | Metabolic profiling 89
3.5 | Results **91**

3.5.1 | Functional structure of the enteric microbiome correlates with athletic state 91
3.5.2 | Distinct differences between host and microbial metabolites in athletes and controls 97
3.5.3 | Correlating metabonomic and metagenomic results 100
3.6 | Discussion 102
3.7 | References 106
3.8 | Supplementary Content 111

Chapter 4---------- 121

EXTENDING THE DIET-EXERCISE-MICROBIOME PARADIGM

4.1 | Abstract 121
4.2 | Introduction 123
4.3 | Materials and Methods 124
4.3.1 | Study population 124
4.3.2 | Acquisition of clinical exercise and dietary data 124
4.3.3 | Preparation of metagenomic libraries and Sequencing 124
4.3.4 | Metagenomic statistical and bioinformatics analysis 125
4.3.5 | Metabolic profiling 126
4.3.6 | Objective assessment of dietary intake 126
4.3.7 | Statistical Analysis 126
4.3.8 | Metabolite identification 128
4.3.9 | Metabolite quantification 128
4.4 | Results 129
4.4.1 | Model generation through clustering of individuals according to objective dietary healthiness scores and exercise levels. 129
4.4.2 | Metabolic characterisation of athletes and less active controls stratified by predicted dietary healthiness. 131
4.4.3 | Quantitative analysis of short chain fatty acids and organic acids. 132
4.4.4 | Bacterial diversity 135
4.4.5 | Diet – Gut microbiome – Gut metabolome – Host metabolome interactions 141
4.5 | Discussion & Biological Interpretation 145
4.5.1 | Metabolic phenotype associated with diet and exercise 145
4.5.2 | Evidence of altered gut microbial – host co-metabolism 146
4.5.3 | SCFA and Branched chain amino acid Metabolism 149
4.6 | Conclusion 151
4.7 | References 153
4.8 | Supplementary Content 160

Chapter 5 ---------- 171

COMPUTATIONAL ELUCIDATION OF MICROBIAL DYNAMICS IN THE GUT MICROBIOME OF ELITE ATHLETES

5.1 | Abstract **171**
5.2 | Introduction **173**
5.3 | Study objective **174**
5.4 | Methods **174**
5.4.1 | Study population 174
5.4.2 | Acquisition of clinical, exercise and dietary data 174
5.4.3 | Preparation of Metagenomic libraries 175
5.4.4 | Metagenomic statistical and bioinformatic analysis 175
5.4.5 | Microbiome modelling 176
5.4.6 | Metabolic profiling 176
5.5 | Results **177**
5.5.1 | Community dynamics captured by AGORA models 177
5.5.2 | Influence of dietary nutrition construction on reactions 181
5.5.3 | Reaction variation between athletes and controls 181
5.6 | Discussion **186**
5.7 | References **188**
5.8 | Supplementary Content **192**

Chapter 6 ---------- 196

A PROSPECTIVE METAGENOMIC AND METABOLOMIC ANALYSIS OF THE IMPACT OF EXERCISE AND/OR WHEY PROTEIN SUPPLEMENTATION ON THE GUT MICROBIOME OF SEDENTARY ADULTS

6.1 | Abstract **196**
6.2 | Introduction **198**
6.3 | Materials and Methods **199**
6.3.1 | Experimental models and subject details 199
6.3.2 | Study recruitment and safe participation 199
6.3.3 | Intervention group allocation 200
6.3.4 | Combined exercise intervention 200
6.3.5 | Measurement visits 201
6.3.6 | Inflammatory cytokine measurement 202
6.3.7 | Dietary data collection 202
6.3.8 | DNA extraction and metagenomic sequencing of faecal microbiome and whey protein supplement 203
6.3.9 | Bioinformatic processing of microbial metagenomic sequencing 203
6.3.10 | Metabolomic sample preparation 204
6.3.11 | Quantification and statistical analysis 204
6.3.12 | Data and software availability 205
6.3.13 | Acknowledgements 205
6.4 | Results **207**
6.4.1 | Study overview 207

6.4.2 | 8 weeks of aerobic and resistance training improves body composition and cardio-respiratory fitness profiles in sedentary subjects 212
6.4.3 | Metagenomic assessment of microbiota after exercise and/or dietary adjustment 214
6.4.4 | Characterization of whey protein supplement microbial content 221
6.5 | Discussion 221
6.6 | References 226
6.7 | Supplementary Content 231

Chapter 7---------- 247

INFLUENCE OF EXERCISE ON THE INTESTINAL MICROBIOME OF IBD PATIENTS

7.1 | Abstract 247
7.2 | Introduction 249
7.3 | Study objective 249
7.4 | Methods 249
7.4.1 | Experimental model and ethical approval 249
7.4.2 | Study recruitment and safe participation 250
7.4.3 | Intervention group allocation 250
7.4.4 | Combined exercise intervention 251
7.4.5 | Compliance and withdrawal from the study 251
7.4.6 | Measurement visits 251
7.4.7 | Inflammatory cytokine measurement 252
7.4.8 | Disease activity and quality of life assessment 252
7.4.9 | DNA extraction and metagenomic sequencing of faecal microbiome 253
7.4.10 | Bioinformatic processing of microbial metagenomic sequencing 253
7.4.11 | Metabolomic sample preparation 253
7.4.12 | Quantification and statistical analysis 253
7.4.13 | Data and software availability 254
7.5 | Results 255
7.5.1 | Study overview 255
7.5.2 | Improvements in body composition and cardio-respiratory fitness following exercise training period 256
7.5.3 | Metagenomic assessment of gut microbiota after exercise and control periods 261
7.5.4 | Pathway assessment 264
7.5.5 | Metabolomic profiling 264
7.6 | Discussion 267
7.7 | Conclusion 269
7.8 | References 270
7.9 | Supplementary Content 272

Chapter 8---------- 273

LONGITUDINAL EFFECTS OF PHYSICAL ACTIVITY ON THE GUT MICROBIOME

8.1 | Abstract 273
8.2 | Introduction 275
8.3 | Study objective 275

8.4 | Methods **276**
8.4.1 | Experimental model and ethical approval 276
8.4.2 | Study recruitment and safe participation 276
8.4.3 | Exercise intervention 276
8.4.4 | Measurement visits 277
8.4.5 | VO_{2max} and body composition measurement 277
8.4.6 | Dietary data collection 278
8.4.7 | DNA extraction and metagenomic sequencing of faecal 278
8.4.8 | Bioinformatic processing of microbial metagenomic sequencing 278
8.4.9 | Metabolomic sample preparation 279
8.4.10 | Quantification and statistical analysis 279
8.4.11 | Data and software availability 279
8.5 | Results **280**
8.5.1 | Study overview 280
8.5.2 | Improvements in body composition and cardio-respiratory fitness 280
8.5.3 | Recording of dietary habits 283
8.5.4 | Metagenomic assessment of gut microbiota 283
8.5.5 | Pathway assessment 288
8.5.6 | Metabolomic profiling 288
8.6 | Discussion **292**
8.7 | References **294**
8.8 | Supplementary Content **296**

Chapter 9 ---------- 297

SUMMARY AND CONCLUSION

9.1 | Summary of original research **297**
9.1.1 | The athlete microbiome 297
9.1.2 | Computational interrogations of the athlete microbiome 298
9.1.3 | Influence of short-term exercise and whey protein supplementation on the gut microbiome 299
9.1.4 | The vulnerable IBD gut microbiome and exercise 300
9.1.5 | Influence of prolonged exercise on the intestinal microbiome 301
9.2 | Implications for related research **301**
9.3 | Relevant future efforts **303**
9.4 | References **304**

Appendix A ---------- 308

CHAPTER 3 ADDITIONAL MATERIAL

Supplementary Table 3.1 308
Supplementary Table 3.2 311
Supplementary Table 3.8 323

Appendix B ---------- 369

CHAPTER 6 ADDITIONAL MATERIAL

Supplementary Table 6.6 – Component C 369
Supplementary Table 6.6 – Component D 375

Appendix C ---------- 378

CHAPTER 8 ADDITIONAL MATERIAL

Supplementary Table 8.2 378

Appendix D ---------- 417

PUBLICATIONS

A Prospective Metagenomic and Metabolomic Analysis of the Impact of Exercise and/or Whey Protein Supplementation on the Gut Microbiome of Sedentary Adults 417
The Microbiome of Professional Athletes Differs from that of More Sedentary Subjects in Composition and Particularly at the Functional Metabolic Level 434
The Metabolic Role of the Microbiota 443

Appendix E---------- 446

ORIGINAL PROGRAMMING

Declaration of Original Work

I, the undersigned, hereby declare that except where otherwise acknowledged, all work presented in this thesis is original and entirely my own. This thesis has not been submitted in whole or in part for a higher degree to any university.

_____

Wiley Barton, BSc

December 2018

"If you believe that you can damage, then believe that you can fix."

— Reb Nachman of Breslov, *Likutey Moharan II*

Acknowledgements

Reaching the reality of compiling this body of work, and even more so, to have participated in the completion of the scientific endeavours from which it is constructed, is the result of enumerable instances of compassionate guidance, desperate intervention, and sober candour. Comprehensively describing my gratitude to every individual that has contributed to the enabling of my pursuit in science is regrettably far beyond the scope of this thesis. I will however continue to honour the village that I have been reared by through the perpetuation of my being and dedication to discovery. I thank you all with the greatest sincerity.

To those with a direct and immediate influence on the PhD which is embodied here:

Academic Supervision

My supervisors, Dr. Orla O'Sullivan, Dr. Paul Cotter, and Professor Fergus Shanahan, I feel that I have been immensely fortunate to be a pupil under the guidance of such a distinguished committee. The diverse expertise made available to me through these exceptional investigators has granted essential perspective on conducting impactful research and engaging with academia.

I offer Dr. O'Sullivan deep gratitude for accepting me as her student, and diligently fostering my growth as a scientist. Specifically, I have been encouraged with efforts holding potential for impact, challenged to provide firm rationale for my claims and pursuits, and guided to understanding when I've hit a dead end. Further, Dr. O'Sullivan has continuously provided the necessary structure to keep me on track and in line with the relentless march

of deadlines. Incredibly, this has been accomplished while still granting me the opportunity for self-development with the exploration of new ideas and approaches. It would be difficult to measure how far into so many rabbit holes I would be without Dr. O'Sullivan's assertive pragmatism. The numerous recalibrations of my focus towards finishing things when they are done have been crucial in reaching the conclusion of my PhD. I like to think I have developed my capacity for this practicality through Dr. O'Sullivan's guidance, and hope to continue to do so as I move forward to future roles she has helped prepare me for.

Similarly, I am extremely thankful to Dr. Cotter for providing an exceptionally stable and intellectually nourishing research environment; one that has promoted my creativity with problem-solving and self-directed professional growth, while constructively confronting it with reality. I am further, deeply impressed by Dr. Cotter's demeanour of perpetual goodwill and Buddha-like understanding and graceful guidance. Authentic kindness of the calibre borne by Dr. Cotter is rare, and there is no corner of our society immune to its benefit. This quality of benevolence is something I will always pursue in myself, even if it means running like a madman.

I am also incredibly grateful to Professor Shanahan for granting me the extraordinary opportunity to be under his supervision. Exposure to the processes of thought and operation refined from complete entrenchment in elite clinical academia has given me invaluable perspective of what it takes to push research out from the shadows of adequacy and into excellence. I have also never met someone with a greater capacity for unequivocal frankness than Professor Shanahan. It is a trait I genuinely admire, even when directed towards myself, and will aim to incorporate it in my own conduct with similar vigour. Through the great patience and supreme wisdom of this committee, there is no question in my mind that I have been moulded into a better scientist, and indeed a better person.

Teagasc and UCC

Dr. Owen Cronin has been a spectacular constant of my PhD's course, and is owed my great thanks for his generosity in my inclusion with so many fantastic studies. Dr. Cronin's insights with the clinical implications of the various studies we have been involved with together have greatly enhanced the impact and relevance of our research.

The Vision 1 laboratory as a whole has been a continuous source of stability, encouragement, and enjoyment. For the years of sharing incredible research, brilliant insights, rich conversations, and tolerance of my comedic endeavours, I thank you all. And while I have a heart filled with fondness for the group as a whole, a particularly dear subset of this group is with the crew of lads better known as the one and only, the legendary, the LadPhlan/WuTangPhlan. Dr. Aaron Walsh, Dr. Calum Walsh, and Dr. Conor Doyle, despite it all, I have an astounding amount of respect and gratitude for these gentlemen. With the crucial exchange of varied insights on *R*, statistics, microbiology, and life, you lads kept my sanity safe.

Amy Murphy, honorary WuTangPhlan member and first ever Teagasc pal, has been a constant source of kindness and generosity. Always encouraging, and one to quickly throw down banter on cats, shows, and everything between, you have my thanks. Furthermore, I'm grateful for Amy's patience with thoroughly proofreading this thesis, shes got one he'll of a I.

I am also extremely grateful to the many volunteers from the various studies of this thesis for providing the data that made wonderful findings and my work possible.

Friends and Past Mentors

Professor Richard M. Cripps at the Biology Department of the University of New Mexico, taking a chance on me as a research assistant, and opening my eyes to the wonder of research was an incredible turning point in my life, and I very well would not be where I am without this opportunity. Thank you Rich, for this, and an everlasting love of the ultimate model and magnificent arthropod, *D. melanogaster*.

Dr. Paul Szauter, founder and Chief Scientific Officer of EquiSeq, an unnatural and spectacular force of science and art. The guy that always has the right answer. The impact your guidance on science and academia has had on me is immense, however it pales in comparison to the profound influence of your philosophy. You were a source of much needed direction in the face of serious challenges, and I will always be grateful for the person you are and the help you provided.

Edward Chacon of New Mexico, and his family will always have my loving appreciation for providing security and understanding during a very difficult period of my life. I will never forget Ed's belief in me, and I wish he could see where I am, and where I'm heading. Marti Mills of Santa Fe, those many years ago, was the first person to suggest a future for me in academia, and I am grateful for her sewing that seed. Stephanie Walstrom of Albuquerque, I will always have the sincerest gratitude for you providing stability when I desperately needed it, and for granting me the opportunity to pursue my education again. Alfonso Sandoval of Santa Fe, my dear friend. Through many transitions in my life you have remained as a companion to marvel at our species' absurdity, explore reality, and to overcome the past with. Thank you for sticking with me over the years and allowing me to do the same for you.

Family

My sister Katherine Barton, it has been such a pleasure to grow up with you, and to have such a wonderful person present for the entirety of my life. I wouldn't have survived it all without you. Darin Brown, my brother, thank you for being in my life; it feels like you were always there. Knowing you were out there working on your PhD at the same time as me has offered a strange comfort, if for no other reason than knowing you will understand exactly why I have been so terrible about communicating!

My parents, Susan Hall, David Barton and Lynne Laier, thank you for never giving up on me, and for always committing to doing your best towards carrying my life forward. I am further, extremely grateful for the capacities of creativity, rationality, and understanding that you have fostered in me. Veronica's parents, Clorinda Torrez, Glen Peterson, and Angie Bratton, you have all been a tremendous support over the years with your kindness and generosity. Most importantly though, your contributions towards the generation of the magnificent being that is Veronica, has indirectly influenced my life in profound ways that are well beyond this acknowledgment, and I am eternally grateful for this.

All the cats that have gifted me with their being, particularly Bubele, Osiris, and my darling Amitābha, you likely will never read this, but there has been no other element of my existence that has been as pure and familiar, you have been the foundation of life as I have experienced it.

Veronica Peterson, my beloved partner, no words, nor actions exist to accurately define my gratitude for you. You are an explosive catalyst to my science, and sentience, and humanity. Our unity has been instrumental in completing this work, as well as moving forward to face each day. I feel fortunate that our work is not yet done, and delight in the world our hands will help shape.

Accomplices

Lastly, to my brothers and sisters united in opposition of the perverse and obsolete structures of governance that divide and subdue our species, and to all who are embedded in the struggle to topple and destroy those who advocate and tyrannically enforce ignorance, consumption, and violence, I am with you.

Our future of equality, opportunity, and progress *will* come to be.

List of Abbreviations

Abbreviation:	Meaning:
°C	Degrees Celsius
1H-NMR	Proton nuclear magnetic resonance
1RM	repetition maximum value
AAB	Amino acid biosynthesis
BA	Bile acids
BMI	Body mass index
BMTagger	NCBI Best Match Tagger
BP	Blood pressure
BPM	Beats per minute
BSH	Bile salt hydrolase
BV	Bacterial vaginosis
CA	correspondence analysis
CB	Carbohydrate biosynthesis
CBA	Conjugated bile acids
CD	Crohn's Disease
CfB	Cofactor biosynthesis
CIN	Cervical intra-epithelial neoplasia
CK	Creatine kinase
CMPF	3-carboxy-4-methyl-5-propyl-2-furanpropionic acid
CREC	Clinical Research Ethics Committee

Abbreviation:	Meaning:
CRP	C-reactive protein
CVD	Cardio-vascular disease
DCA	deoxycholic acid
DEXA	Duel Energy X-ray Absorptiometry
DNA	Deoxyribonucleic acid
EGFR	Epidermal growth factor receptor
EM	Energy metabolism
EPIC	European Prospective Investigation into Cancer
FDR	False Discovery Rate
FFQ	Food frequency questionnaire
GC-MS	Gas chromatography mass-spectrometry
GI	Gastrointestinal
HDAC	Histone deacetylase
HILIC UPLC-MS	Hydrophilic interaction ultra performance liquid chromatography mass-spectroscopy
HIV	human immunodeficiency virus
HMP	Human Microbiome Project
HUMAnN2	Human Microbiome Project Unified Metabolic Analysis Network
IBD	Inflammatory Bowel Disease
IBS	Irritable Bowel Syndrome
IFN-gamma	Interferon gamma
IPAQ	International Physical Activity Questionnaire

Abbreviation:	Meaning:
IQR	Interquartile range
IRFU	Irish Rugby Football Union
k-NN	k-nearest neighbour
LB	Lipid biosynthesis
LCA	lithocholic acid
LPS	Lipopolysaccharides
MAPK	Mitogen activated protein kinase
MSD	Meso-Scale Discovery
NAFLD	Non-alcoholic fatty liver disease
NMDS	Non-metric multidimensional scaling
NO	Nitric oxide
OPLS-DA	Orthogonal Partial Least Squares-Discriminant Analysis
PAG	Phenylacetylglutamine
PCA	Principal Component Analysis
PCoA	Principal Coordinate Analysis
PERMANOVA	Permutational analysis of variance
QC	Quality Control
RP UPLC-MS	Reverse-phase ultra performance liquid chromatography mass-spectroscopy
RPE	Rating of perceived exertion
rRNA	Ribosomal ribonucleic acid
SASP	Senescence-associated secretory phenotype

Abbreviation:	Meaning:
SCFA	Short chain fatty acid
SD	Standard deviation
TMA	Trimethylamine
TMAO	Trimethylamine-N-oxide
TNF-alpha	Tumour necrosis factor alpha
TSP	3-(Trimethylsilyl)propionic-2,2,3,3-d4 acid
UC	Ulcerative Colitis
VB	Vitamin biosynthesis
VO2max	Maximum aerobic capacity

List of Visual Resources

Figure 1.1 | Demonstration of key microbiota and metabolites of the human microbiome, delineated according to human physiology. 17

Figure 1.2 | Host-microbe metabolic interaction. 33

Figure 2.1 | Comparison of sequences from all studies. 70

Figure 3.1 | Comparison of phylogenetic constructions from metagenomic and 16S rRNA gene sequencing sourced from all participants. 93

Figure 3.2 | Group-wise comparison of microbial metagenomic and metabolomic profiles. 94

Figure 3.3 | Group variation of microbial metabolic function and associations between pathways and clinical and dietary variables. 96

Figure 3.4 | Athletes display a profile of Short Chain Fatty Acids that alters from that of the controls. ..99

Figure 3.5 | Distinctive association profiles of metabolic pathways to metabolites in athletes and controls. 101

Figure 4.1 | Clustering of individuals according to exercise status and predicted adherence to healthy dietary habits, calculated from urinary 1H-NMR spectra of individuals, to form comparative models. 130

Figure 4.2 | ^{1}H-NMR, SCFA and bacterial diversity results. 134

Figure 4.3 | Taxonomic β-diversity of participant classes. 136

Figure 4.4 | Diversity of metabolic pathways between classes. 138

Figure 4.5 | Linear regression of faecal water ^{1}H-NMR spectra against bacterial diversity (Shannon index), corrected for confounding factors (age, lean mass and fat mass). ...140

Figure 4.6 | Faecal – Urinary metabolic interactions. 142

Figure 4.7 | Microbial metabolic superpathway–metabolite interactions. 143

Figure 5.1 | Comparison of complete taxonomic profiles with selected AGORA models. 178

Figure 5.2 | Comparison between metabolic pathways and FBA reactions. 179

Figure 5.3 | Influence of nutritional inputs on reactions. 182

Figure 5.4 | Reaction variation between athletes and controls. 184

Figure 6.1 | Study design. 208

Figure 6.2 | Alterations in cardiorespiratory fitness and body composition following exercise interventions, protein interventions, and combined interventions. 213

Figure 6.3 | Intervention effects on taxonomic and functional pathway diversity of the intestinal microbiome. 216

Figure 6.4 | Pairwise analysis of detected virus taxonomy prior to and following intervention. 218

Figure 7.1 | Study outline. 258

Figure 7.2 | α-diversity of taxonomic profiling. 262

Figure 7.3 | β-diversity of taxonomic profiles. 263

Figure 7.4 | α-diversity of metabolic pathways. 265

Figure 7.5 | β-diversity of metabolic pathways. 266

Figure 8.1 | Study overview. 282

Figure 8.2 | Subject body composition and anthropometric data. 284

Figure 8.3 | Nutritional constituents of participants over study period. 285

Figure 8.4 | α-diversity of bacterial species and pathways. 286

Figure 8.5 | β-diversity of bacterial taxonomy. 287

Figure 8.6 | β-diversity of metabolic pathways. 289

Figure 8.7 | Metabolite trajectories for both participants. 290

Figure 8.8 | Linear regression of ^{1}H-NMR urinary spectra to BMI. 291

Abstract

1.1 | Introduction

Revolutionary insights of the human microbiome, the conglomeration of microorganisms that stably occupy an expansive array of anatomical locations of the human body, have certified the phenomenon as an integral component of human biology. As progress has been made in elucidating the various factors of influence that the microbiome imposes on human health, it has become apparent that modulation of the community structure and functional activity of the microbiome is an essential step in utilizing these microbes as an element of intervention in healthcare. Advancement in understanding of how the microbiome is manipulated is accordingly of great interest and importance.

1.2 | Methods

Physical activity is a recently emergent prospective modulator of the human intestinal microbiome. This development in microbiome-host interaction presents great potential, but has presently been investigated in limited depth. To expand the understanding of this biological dynamic we engaged the subject with focus on metabolic function of the microbiome, utilising advanced computation methods and molecular technologies. Our approach has sought to address multiple aspects of exercise induced microbiome alteration, and accordingly was conducted with randomized control trials and prospective observational study designs.

1.3 | Results

This work describes further evidence that physical exercise has a role in directing the community structure and metabolic activity of the gut microbiome. Our inspection of professional athletes, whose prolonged engagement in rigorous exercise grants perspective of the extreme end of the spectrum of physical activity, revealed a divergent microbiome from that of more sedentary controls. Metabolic pathways from the athletes were configured to offer enhanced energy recovery from the intestinal environment. Quantified metabolomic phenotyping of this system similarly revealed that the athlete microbiome had a favourable profile of SCFA enrichment.

Examination of structured short-term exercise on exercise naïve individuals revealed subtle alterations of the microbiome, both in terms of phylogenetic composition and metabolic output. In our investigation it was also observed for the first time that supplementation of whey protein resulted in an alteration of the virome. Further sequencing of the supplement itself revealed a highly similar viral composition to the participants, suggesting that whey protein, a widely used supplement, is directly transmitting virus particles.

To explore the implementation of exercise as a supportive intervention for health conditions related to the GI system, patients with inflammatory bowel disease (IBD) were recruited for a short-term exercise intervention study. While minimal changes in the composition of the IBD patients' gut microbiome were identified, the patients were shown to undergo the treatment without insult to, or perturbation of their disease state. This important finding illustrates that sufferers of IBD can engage in physical exercise, reaping the well-known health benefits from the activity, without obvious influence on the gut microbiome.

Elucidation of the longitudinal effects of exercise on the gut microbiome was achieved through the examination of two individuals engaging in exercise over the course of six months. Through this *N of 1* style of study, resources were concentrated on the two participants, enabling the acquisition of much greater detail on the proposed question. Here again, it was observed that while the participants were rewarded with improvements in health parameters, changes of the gut microbiome were subtle.

Complementing the impressive results produced from initial investigations of professional athletes, effort was put forth to apply advanced computational approaches to the previously generated data in order to advance understanding of the athlete microbiome. A novel algorithm designed to predict adherence to healthy dietary habits based on metabolomic profiling was used to better define dietary influence on the metabolomic dynamics of the gut environment. Separately, a computational modelling method centred on Flux Balance Analysis generated models of microbial metabolic exchange within the athletes and corresponding controls.

1.4 | Conclusions

The findings supporting this thesis have immediate relevance in the area of athleticism, with implications for professional athletes and individuals casually engaged in exercise. There is further impetus provided for continued investigation of the athlete microbiome, the influence of prolonged periods of exercise on populations with low fitness levels, and the impact of dietary supplement derived microbe transmission.

Chapter 1

INTRODUCTION AND BACKGROUND

1.1 | Abstract

The human enteric microbiome has been identified as having key roles in health and numerous diseases. Comprised of trillions of microbial cells and viral particles, the microbiome represents a bioactive ecosystem of exceedingly high complexity. Investigation of the importance of enteric microbes has progressed through culture dependant (e.g. media-based methods) and molecular (e.g. genetic sequencing and metabolomic analysis) techniques. The latter have become increasingly popular and evolved from being used for taxonomic identification of microbiota to elucidation of functional capacity (sequencing) and metabolic activity (metabolomics).

The following chapter provides details of this thesis; the core rationale behind it, outcomes of its course, and an outline of its structure. Following this description, an overview of the metabolic capabilities of the gut microbiome will be presented in the context of health and disease.

1.2 | Author's relevant publications

Publications derived from this thesis:

Cronin, O*; **Barton, W***; Skuse, P; Penney, NC; Garcia-Perez, I; Murphy, EF; Woods, T; Nugent, H; Fanning, A; Melgar, S; Falvey, EC; Holmes, E; Cotter, PD; O'Sullivan, O; Molloy, MG; Shanahan F. A Prospective Metagenomic and Metabolomic Analysis of the Impact of Exercise and/or Whey Protein Supplementation on the Gut Microbiome of Sedentary Adults. *mSystems.* 2018 June 26; 3(3):e00044-18.

Cronin, O; O'Sullivan, O; **Barton, W**; Cotter, PD; Molloy, MG; Shanahan, F. Gut microbiota: implications for sports and exercise medicine. *British Journal of Sports Medicine*, 2017 May; 51(9):700-701.

Barton, W; Penney, NC; Cronin, O; Garcia-Perez, I; Molloy, MG; Holmes, E; Shanahan, F; Cotter, PD; O'Sullivan, O. The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level. *Gut*. 2017 March 30. doi: 10.1136/gutjnl-2016-313627.

Barton, W; Shanahan, F; Cotter, PD; O'Sullivan, O. The Metabolic Role of the Microbiota. *Clinical Liver Disease*. 2015 May 7; 5 (4), 91-93

Articles from this thesis in preparation for publication:

Metabolic Phenotyping in Relation to the Gut Microflora.

Barton, W; Shanahan, F; Cotter, PD; O'Sullivan, O.

Extending the Diet-Exercise-Microbiome Paradigm.

Barton, W*; Penney, NC*; Garcia-Perez, I; Holmes, E; Shanahan, F; Cotter, PD; O'Sullivan, O.

Computational modelling of the elite athlete gut microbiome.

Barton, W*; Baldini, F*; Penney, NC; Garcia-Perez, I; Holmes, E; Shanahan, F; Cotter, PD; O'Sullivan, O; Thiele, I.

The impact of exercise on the gut microbiome of sedentary adults with Inflammatory Bowel Disease: A prospective metagenomic and metabolomic analysis.

Cronin, O*; **Barton, W***; Whiston, R; Penney, NC; Garcia-Perez, I; Murphy, EF; Woods, T; Nugent, H; Fanning, A; Melgar, S; Falvey, EC; Holmes, E; Cotter, PD; O'Sullivan, O; Molloy, MG; Shanahan F.

Influence of prolonged exercise on the gut microbiome: An 'n of 1'case study.

Cronin, O*; **Barton, W***; Whiston, R; Penney, NC; Garcia-Perez, I; Murphy, EF; Woods, T; Nugent, H; Fanning, A; Melgar, S; Falvey, EC; Holmes, E; Cotter, PD; O'Sullivan, O; Molloy, MG; Shanahan F.

(* Shared 1st authorship placement)

1.3 | Dissemination of author's research

1.3.1 | Conference participation

Invited talks

The Diet and Exercise-Microbiome Paradigm. IPA World Congress and Probiota Conference. (February 2018) Barcelona, Spain.

Is there a future for targeted manipulation of the microbiome in athletes? National Sports Nutrition Conference. (November 2017) Oslo, Norway.

DNA sequencing and bioinformatics workshop. Alimentary Pharmabiotic Centre-Imperial College London Student Research Conference. (October 2017) Cork, Republic of Ireland.

The Diet and Exercise-Microbiome Paradigm: Distinct Functional Profiles of the Athlete Microbiome Revealed by Metagenomic and Metabonomic Analysis. Alimentary Pharmabiotic Centre – Imperial College London Student Research Conference. (September 2016) London, United Kingdom.

Poster presentations

Wiley Barton, Nicholas C. Penney, Owen Cronin, Isabel Garcia Perez, Michael G. Molloy, Elaine Holmes, Fergus Shanahan, Paul D. Cotter, Orla O'Sullivan. *The Diet and Exercise-Microbiome Paradigm: Distinct Functional Profiles of the Athlete Microbiome Revealed by Metagenomic and Metabonomic Analysis.*

(May 2017) Cell Symposium: Exercise Metabolism. Gothenburg, Sweden.

(September 2016) Exploring Human Host-Microbiome Interactions in Health and Disease Conference. London, United Kingdom.

Orla OSullivan, **Wiley Barton**, Owen Cronin, Peter Skuse, Michael G. Molloy, Paul D. Cotter, Fergus Shanahan. *Metabolomic and Metagenomic analysis of Exercise and dietary protein.* (July 2015) ISMB conference. Dublin, Republic of Ireland.

1.3.2 | Campus participation

Wiley Barton, Nicholas C. Penney, Owen Cronin, Isabel Garcia Perez, Michael G. Molloy, Elaine Holmes, Fergus Shanahan, Paul D. Cotter, Orla O'Sullivan.

The Exercise and Diet-Microbiome Paradigm: AthMET 'Omics.

(December 2016) Oral presentation at New Horizons Conference. Cork, Republic of Ireland.

Wiley Barton, Owen Cronin, Michael G. Molloy, Fergus Shanahan, Paul D. Cotter, Orla O'Sullivan.

Metabolomic and Metagenomic analysis of Exercise and dietary protein.

(July 2015) Poster presentation at ICL and APC Symposium. Cork, Republic of Ireland.

1.4 | Introduction

Microbial communities are ubiquitously integrated with the physical world we occupy, and indeed, are intricately linked to our corporeal forms [1-10]. Comprised of numerous phylogenetic lineages, such as protozoa [11-16], fungi [17-20], viruses [21-24], and prokaryota [25-29], these *microbiome* communities exist in complex ecological states, with widely varying capacities of function [29-34]. While the influences of these microbial systems are intently studied in diverse research areas, such as food [7, 35], agriculture [36], and ecology [37], all with profound implications, the microbiome as it pertains to human health has gained immense momentum as an area of research. Underlying the human microbiome is recognition that the microbial communities integrated throughout the human body act as important factors in metabolic function, essentially operating as previously unknown organs [38-43]. Unlike the mammalian components of the human body, the microbiome can be modified to achieve specific health aims. Specifically, the use of pharmaceuticals [44-46], pre- and probiotics [47-49], diet [50-55], and exercise [56-61] can be utilised to alter the composition and function of the microbiome. The latter of which, exercise, has particular importance in this regard due to the numerous benefits exerted by the behaviour on its own, such as with brain function [62], immunity [63], and metabolism [64, 65]. To this end, the current chapter and the thesis it introduces outlines advances made towards interaction of exercise and the human gut microbiome.

1.5 | Purpose of research

The primary aim of this thesis is to elaborate on the understanding of the human gut microbiome in its relation to physical activity and diet. In particular, the effort of investigating the *diet and exercise—microbiome paradigm* has been driven by the utilization of computationally centred practices and methodologies. As with many areas of the microbiome, advances in high-throughput technologies (e.g. shotgun metagenomic sequencing) have enabled researchers to probe inquiries with unprecedented depth, while simultaneously creating a bottleneck with the volume of data generated. As such, navigation of complex high-volume data has become a limiting factor in the perpetuation of research (i.e. drawing evidence for hypotheses, disseminating results, and laying groundwork for the generation of further inquiries).

1.6 | Thesis structure

The final portions of this chapter will establish the base knowledge underpinning this thesis, describing the current relevant information related to the interaction of the human gut microbiome and host physiology. Specific attention is given to the microbiome of the gastrointestinal system and its relation to physical activity (i.e. exercise).

Prior to description of experimental work in chapters 3-8, materials and methods utilised throughout the course of this PhD are outlined in chapter 2. Beginning with chapter 3, the first such investigation within this thesis of professional athletes is presented. Here, via combined metabolomic and metagenomic analysis, the functional capacity of the elite athlete gut microbiome was presented for the first time. Chapter 4 continues the exploration of the elite athlete microbiome with the application of a novel mathematical tool for dietary analysis based on urine metabolomic profile with the professional athlete cohort and accompanying controls. Similarly, chapter 5 applies the flux balance analysis (FBA)

metabolomic modelling approach to the original high-throughput sequencing data generated from the elite athletes.

Beginning with chapter 6, the focus of this thesis shifts from the investigation of professional athletes to exercise naïve individuals. Chapter 6, the first such entry, describes an intervention study in unfit, but otherwise healthy adults. Here, participants were subject to treatment of a structured exercise regime, whey protein supplementation, or both. The study revealed modest influence of exercise on the gut microbiome, however for the first time, reported that viral particles in whey protein transfer to humans via consumption. Chapter 7 continues in this vein with an exercise intervention in adult patients with inflammatory bowel disease. In this chapter, it is revealed that short –term exercise confers no insult to disease state or composition of gut microbiome in IBD patients, while still granting the well documented positive influence of physical activity on health. Concluding the description of experimental efforts, chapter 8 details an *n of 1* type study in which two individuals were prospectively examined over the course of 6 months. With this approach resources were focused on the two participants, granting more elaborate assessment of longitudinal influences of moderate intensity self-structured exercise. Finally, chapter 9 concludes this thesis, with discussion of the cumulative implications of previous chapters, and consideration of future directions.

1.7 | The human microbiome as it relates to metabolic function and health

It has been established that communities of microorganisms, *microbiomes*, reside on or within nearly every physical substrate on our planet (and associated artificial satellites) [1-10]. The importance of the microbiome as an integral component of the human biological system has become clear, not only to microbiologists, but to clinicians and the general public. Encompassing multiple divisions of the tree of life such as protozoa [11-16], fungi [17-20], viruses [21-24], and prokaryota [25-29], microbiomes consist of compositions of microorganisms in various ecological structures and functional capacities [29-34]. Comprehension of the microbiome's vast influence on human health has become formidably wide. Accordingly, the ever expansive sum of such information presents a continuously growing hurdle in the field for those involved in the field to overcome. The remainder of this chapter aims to offer essential background to the human microbiome, providing overview of microbiomes delineated by human anatomy within the framework of microbe-host metabolic interaction.

1.7.1 | Membership

Within the context of human health, protozoa have historically been regarded as opportunistic parasites, acting as agents of disease [14, 16, 66-69]. Notable examples of protists implicated in morbidity include *Giardia*, *Entamoeba histolytica*, and *Cryptosporidium* [16, 66-68]. A feature common to infections with these organisms is invasion of the gastrointestinal (GI) mucosa, resulting in diarrhoea. Inconsistencies in the development and presentation of symptoms in response to infection, along with disagreement on phylogenetic classification of the organisms, have led to an incomplete acceptance of the etiological role of the protozoa in disease. In contrast, there is evidence for a protective role of a commensal protist, *Tritrichomonas musculis* (*T. musculis*) in murine models [14, 15]. Intestinal coloniza-

tion of non-human animals with *T. musculis* produced a host-protective immunological response, which attenuated the deleterious effects of a challenge with *Salmonella* [15].

The fungal microbiome, or *mycobiome*, like that of the protozoan contribution to the microbiome, is incompletely understood. As shortcomings of the selective nature of culture-dependent methods for taxonomic profiling were relocated to the past for prokaryotic organisms with the revolution of 16S sequencing, the fungal counterpart of the 16S rRNA gene, the internal transcribed spacers (ITS) rRNA gene loci have received less attention in terms of supporting sequence analysis software, and databases [19, 70]. Although the catalogue of taxonomic members of the human mycobiome is still expanding, the role of *Candida* species in health has been relatively well characterised. While present at various anatomical sites as a stable constituent of the human microbiome, aberrations in *Candida* abundance and morphology increase pathogenicity of the organism, which, regardless of immunocompetency can result in infection with widely varying virulence [18-20, 32, 71]. Within the GI tract, pathogenicity of *Candida* has been shown, in part, to be regulated by *Lactobacillus* species. Although the precise mechanism remains unknown, it is evident that some lactobacilli suppress *Candida* through host epithelial immunomodulation and competition for epithelial adherence [48, 72-74]. Advantageously reducing competition for proliferation at the epithelium, lactobacilli interfere with hyphal formation of *Candida* with the production of SCFAs and exopolysaccharides (EPS) [72, 74].

Viral infections are a significant cause of human disease and morbidity [22, 23, 75]. Despite the challenges viruses pose to health, it has become evident that bacteriophage, virus particles that target bacteria, have a beneficial role in the microbiome. Indeed, these viruses that attack bacteria may be harnessed for therapeutic use. Bacteriophage are both highly efficient and specific. Each virus will typically infect only a small number of bacterial

strains. Elimination of bacterial cells occurs via lysis, resulting from propagation of viral particles within the infected cell. Publically disclosed in 1915, the discovery of bacteriophage preceded that of antibiotic drugs by over a decade [76-78]. Despite this, therapeutic use of bacteriophage has historically been overshadowed by antibiotics. A revival of interest in bacteriophage therapy has been driven by the emergence of antibiotic resistance and limitations in novel antibiotic discovery [30, 79-82]. In addition to the therapeutic potential of bacteriophage, the viruses integrally influence the larger microbial community through transmission of genetic material between hosts (e.g. antibiotic-resistance genes), and lowering host density via lysis [83].

Each known component of the microbiome presents specific challenges in its investigation and role in human health. Bacteria, however, are understood as the predominate drivers of the microbiome, and thus, are the focal point of microbiome research [25-27, 29, 84].

1.7.2 | Womb to Tomb

Present from birth to death, an individual's microbiome maintains a constant presence as a chimeric organ [39, 40, 43, 85]. Seeding of this microbial system occurs at the beginning of life with birth, via transmission of a mother's microbiome to her infant during the birthing process [86-90]. Influenced by direct environmental transmission, a delivered infant will inherit either the mother's vaginal and faecal microbiota as it passes through the birthing canal, or the skin microbiota during caesarean delivery [86-88]. Either route of delivery imposes prolonged multi-faceted effects on the infant [91, 92]. Vaginal birth confers a microbiome of the mother's urogenital system which has undergone specific alterations throughout the pregnancy that are conducive to the development of robust and functional immune and GI systems of the infant [89]. Alternatively, numerous deleterious health

effects for infants delivered by caesarean section have been identified. Immediate influences upon the infant include increased risk of exposure to antibiotic resistant bacteria from the mother's skin [87]. Long-term insults to health arising from caesarean delivery include greater risk of developing obesity, sensitivity to food and inhalant allergens, and asthma [91-95]. In light of increasing awareness of potential negative health effects associated with caesarean delivery, an experimental procedure of vaginal seeding has been developed to simulate the microbial exposures present in vaginal birth via administration of vaginal swabs to newly-delivered infants [96]. Implementing vaginal seeding is an issue of contention, and many clinical practitioners are wary of the intervention prior to elaborate investigation of its effects [97, 98].

Throughout infancy, an individual's core microbiome is continuously influenced by the mother and environment. Whether nourished by the mother's natural breast milk or formula, the infant microbiome continues to be moulded through supplied nutrition. In this regard, a positive health bias towards biological 'tradition' persists, as both the process of breast feeding and breast milk itself, convey health benefits superior to formula [89, 99, 100]. Progressing through infancy the microbiome goes through highly variable changes, beginning to stabilize at approximately 2 years of age. Flux of the microbiome during this period is attributed to numerous factors, including dietary variations (e.g. milk vs. solid food), immunological development, introduction to novel microbes, and antibiotic exposure [87, 89, 90, 100-102].

Through the transition from infancy to childhood, and onto adulthood, the microbiome of an individual stabilizes, while still being influenced by drug exposure [29, 44, 45, 103, 104], physical activity [57-61, 105-110], the environment [3], and diet [21, 49, 53]

(discussed more elaborately in proceeding sections) [111, 112]. The microbiome changes again with old age [50, 113, 114], and again finally with death [115, 116].

1.8 | The Human Body – a Microbiome Perspective

Microbial communities take form within any accessible area of a host's body. The defined niches with stable communities in humans and other mammals are currently generalized to the respiratory system [20, 117-120], nasal [25, 121, 122] and oral cavities [17, 25, 26, 123], skin [22, 25, 26, 75, 88, 124-128], vagina and urinary tract [25, 87, 88, 96, 129-131], and gastro-intestinal system [21, 25-27, 29, 40, 43, 84, 85, 87]. For each of these unique communities, varied technical challenges and health related prospects are present.

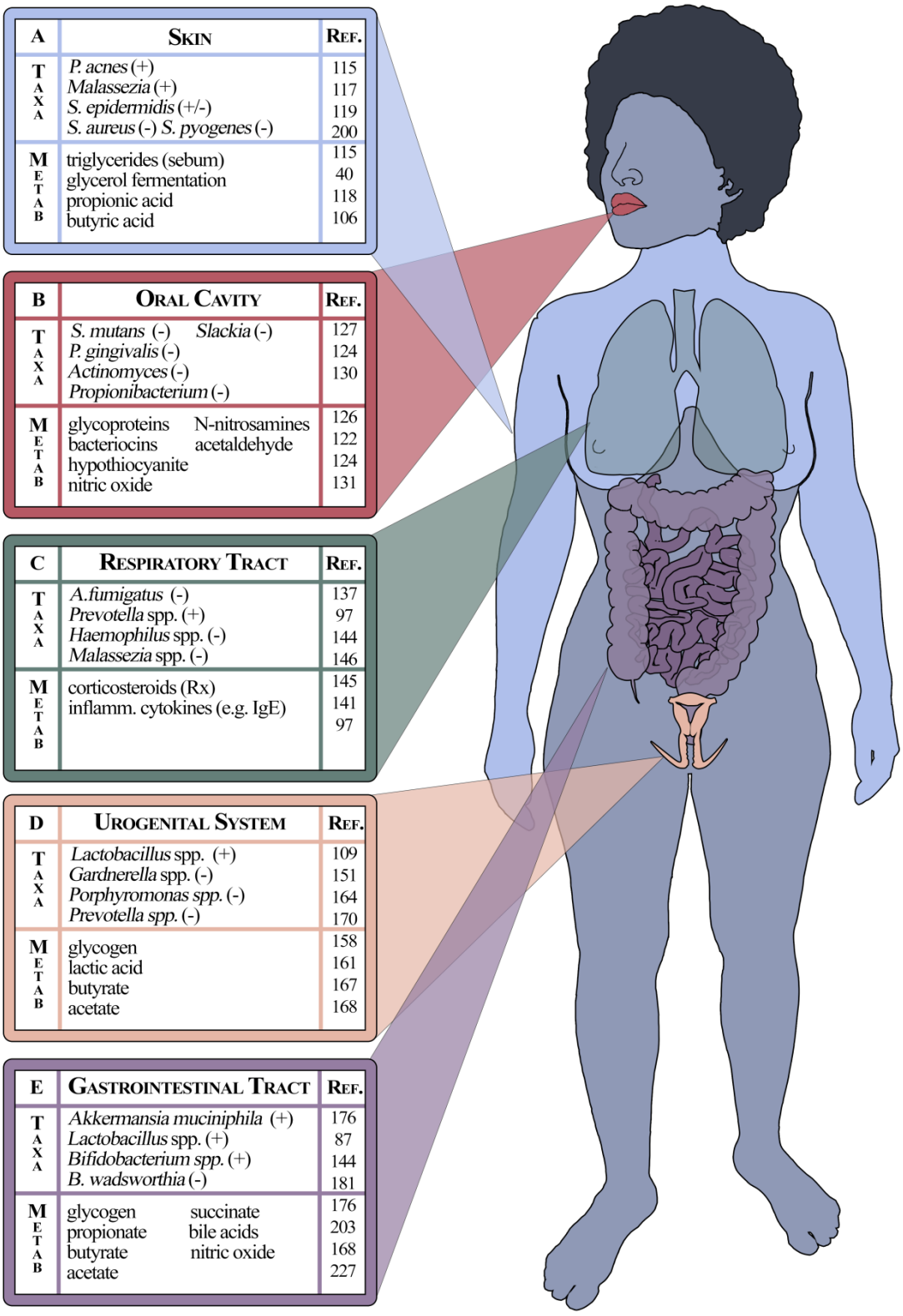


Figure 1.1 | Demonstration of key microbiota and metabolites of the human microbiome, delineated according to human physiology. (**A**) The skin, (**B**) oral cavity, (**C**) respiratory tract, (**D**) urogenital system, (**E**) and gastrointestinal tract are each highlighted with examples of microbiota (Taxa) and relevant metabolic activity (Metab). Beneficial associations to host health are denoted as (+), while negative associations are (-).

1.8.1 | The Skin

Comprising a relatively vast space (~1.8 m^2 for an adult human) with an equally expansive range of subsystems defined by folds, crevices, pH, secretion profiles, and environmental exposures, the skin supports highly varied microbial communities functioning in diverse ecological constraints (see Figure 1.1 A) [125, 132, 133].

Ecological partitioning of the skin microbiome is further defined by elementary biological traits of the host. Microbial composition at specific anatomical locations coordinates with gender [132, 134, 135]. Topical sampling of hand palms demonstrates greater diversity of bacterial taxa in women than men, with specific taxa differentially abundant between the two sexes [134, 135]. Similar results have been presented for other body sites, such as the thigh and torso [132, 134]. Expectedly, cohabitation of sexually active partners results in a shared skin microbiome that accurately matches couples 86% of the time [134]. Ancestral host genetics have also been demonstrated to influence the composition of the skin microbiome. Male participants of diverse ethnic backgrounds, all dwelling in a single geographic location, were shown to have microbial differences specific to ethnicity [136]. Despite such associations with the skin microbiome, ancestral genetics have been shown to exert negligible influence on the gut microbiome, where instead other factors, such as environment, play a more profound role in the form and function of the microbial community [137]. Microbiome structure also corresponds with age, with diversity increasing towards its peak in adulthood, before sharply declining with elderhood [132, 138]. Guidance of skin microbiome composition is not limited to host biology, but also subject to extrinsic factors.

Continuous environmental interaction leaves the skin communities as our most exposed microbiome. Environmental factors shown to be influential include: hygiene routines,

topical medication and cosmetic use, and residential environment (e.g. rural vs. urban) [125, 126, 132, 135, 139]. Despite the vulnerability to external perturbations, an individual's skin microbiome maintains a consistent core structure [140]. Primary constituents of this stable inner-community are residential commensal taxa which perform homeostatic functions and act as a barrier against transient and potentially pathogenic species, subsequently maintaining a role in a variety of cutaneous conditions [128, 138, 140, 141]. Among these residential members are strains of *Propionibacterium acnes* (*P. acnes*), *Malassezia*, and *Staphylococcus epidermidis* (*S. epidermidis*) [138, 140, 142, 143]. Lipophilic, *P. acnes* and species of *Malassezia* proliferate in sebaceous gland rich body sites, such as the face and back [125, 138, 142]. The rich pool of triglycerides found in sebum are hydrolysed by microbiota to produce fatty acids which assist in bacterial adherence and maintaining an acidic pH [138, 144]. Such low pH environmental conditions reinforce the presence of lipophilic commensals, while simultaneously inhibiting colonization by potentially pathogenic strains of *Staphylococcus aureus* and *Staphylococcus pyogenes* (*S. aureus* and *S. pyogenes*, respectively) [138, 145]. *P. acnes* additionally contributes to suppression of methicillin-resistant *S. aureus* through glycerol fermentation to short-chain fatty acids (SCFA), particularly so with propionic acid which also inhibits growth of *Escherichia coli* and *Candida albicans* [74, 141, 145]. Despite its contribution to homeostasis of the skin, *P. acnes*, as its name suggests, plays a clear role in the condition of acne.

1.8.2 | The Mouth

The oral cavity provides a rich landscape wherein microbes establish a well-defined ecosystem (see Figure 1.1 B). Structure morphology and tissue type of the human mouth offers a variety of microbial habitats, further delineated by conditions of oxygenation, pH,

and nutrient availability [146, 147]. Control of the oral microbiome is mediated in concert by factors derived both by host and microbiota [146, 148-150].

Immunological training by microbiota seeded early in life defines the distinction between the commensal core and transient pathogens, wherein selected commensals create biological barriers with biofilm formation, alterations of pH and oxygen levels, and production of growth interfering molecules [148, 150, 151]. Bacteriocins, i.e., small peptide molecules including lantibiotics and microcins, are one such vehicle of microbial-derived molecular regulation of community composition within the mouth (and other microbial systems) [150]. Underlying mechanisms coordinating this antagonistic inter-microbe regulation of community structure remain unresolved, however its complexity is highlighted by findings of at least 1,169 putative lantibiotic gene clusters within the oral metagenomes defined by the Human Microbiome Project (HMP) [152].

Complementing microbial influences, the host exerts an array of modulating effects upon the oral cavity. Prominently featured throughout this environment, saliva, is comprehensively understood for moistening the mouth and aiding in the mastication, swallowing, and digestion of food. Saliva also provides an essential nutrient source for microbiota, containing complex molecules such as glycoproteins (e.g. mucins) [146, 148, 153, 154]. Similarly, saliva-derived proline-rich glycoproteins contribute to pellicle formation on mouth surfaces, immobilizing microbiota through their adherence to the structures [146, 148]. Bioactive compounds found within saliva also include potent factors that inhibit growth or otherwise modify the microbial complex's activity within the mouth. For example, bacterial growth is curbed by lysozyme mediated cell lysis and interference of glucose metabolism with lactoperoxidase catalysed conversion of hydrogen peroxide and thiocyanate to hypothiocyanite [146, 148].

Sustaining a balanced oral microbiome purportedly results in numerous local and systemic health benefits. Nitric oxide (NO) is an important cellular signalling molecule, crucially involved with various physiological functions: metabolism, nerve function, and cardiovascular function. Key oral microbiome constituents have demonstrated the ability to reduce dietary nitrates to nitrite [42, 148, 154]. Converted nitrite is deposited into saliva, which is ingested after oral cavity circulation, leading to NO conversion and the subsequent transmission to tissues across the body [42, 154]. Countering the potential health benefits of bacterial nitrite supplementation, the compound may stimulate cancer development through formation of carcinogenic N-nitrosamines [42]. Posing similar risk of carcinogenesis, acetaldehyde is converted from ethanol by oral bacteria [154].

Dysfunction of the oral microbiome contributes directly to dental diseases; the most widely recognized such condition being tooth decay or dental caries. Caries formation begins with bacterial fermentation of carbohydrates to organic acids, resulting in localised pH reduction and subsequent tooth demineralization [146, 148, 151, 154]. Once the site has been acidified, the affected environment becomes increasingly hospitable to bacteria that are tolerant of low pH conditions, thus stimulating proliferation of destructive communities, and worsening of the condition [146, 148, 154]. While *Streptococcus mutans* is implicated as an integral component of tooth decay, it is evident that no single perpetrating organism is the causative agent, and instead polymicrobial activity drives the condition, with diverse actors from genera such as *Actinomyces, Slackia, Propionibacterium,* and *Lactobacillus* [151].

Periodontal disease is similarly notorious and well-understood for its influence by microbiota. Prolonged biofilm formation at the interface of gingival tissue and the tooth surface leads to accumulation of pathogenic bacteria that exacerbate inflammation through secretion of cytotoxic compounds such as lipopolysaccharides (LPS) [148, 154]. Resultant

bleeding from the inflammation provides a source of iron from heme, a molecule utilized by pathogenic microbes (e.g. *Porphyromonas gingivalis*) [148, 154]. Without disruption, the dysbiotic periodontal microbial community thrives, and with enhanced immunological antagonisation of the gingival tissue, contributes to induction of a dysregulated inflammatory response, permanently damaging connective tissue and bone [148, 154].

Despite long-standing associations between oral health and microorganisms, and increasing understanding of such associations with advances of genetic sequencing and metabolomic technologies, a great deal more remains to be investigated within the oral microbiome.

1.8.3 | The Nose and Respiratory System

The human lung has long been considered a sterile biological system unless challenged with disease. Awareness of the microbiome has led to revision of this perception, and general acceptance of the respiratory microbiome (see Figure 1.1 C).

Under conditions of general health, the lung environment reflects many characteristics of the mouth and nose interiors: moderate thermal stability, high oxygen availability, mucosa-lined internal surfaces, and a continuous influx of environmental microbes. Despite being similarly well-suited for thriving microbial activity as the respiratory system openings, modern investigation of respiratory-related microbes deeper in the system projects a healthy lung microbiome with low phylogenetic diversity [155-157]. Simplicity of the lung microbiome is most notable in comparison to the oral cavity, which acts as a major channel for microbiota translocation with microaspiration of aerosolized material from the upper respiratory tract and direct migration along the oropharynx mucosa [157, 158].

While many human microbial communities exhibit high levels of diversity when healthy, presenting associations between disease and reduced diversity, the respiratory

microbiome deviates from this trend and is indicated to be more susceptible to malignancy when complexity of its composition increases [25, 148, 157, 159, 160]. This is observed as far up in the respiratory system as the nasal cavity, with elevated diversity of the inner nostril associated with number of allergies [134]. Conversely, postsurgical outcome of sinus surgery is better with more diverse sinonasal microbiomes, suggesting an unpredictable relationship between upper respiratory tract microbial diversity and health [161].

Association between heightened lung microbiome complexity and compromised pulmonary health has been demonstrated with asthma, an inflammatory disease [20, 119, 121, 160, 162]. Aligned with much of microbiome-health interaction, evidence supports early life microbial exposures as critically influential on respiratory health. Strong epidemiological associations assert an increased risk of inflammatory respiratory disease with caesarean birth, and reduced risk from diverse antigen presentation (i.e. rural and farm exposures) [93, 94, 163, 164]. Mechanistic detail of the microbial role in asthma remains undescribed, however associations have been made between a deviation from the typical predominance of Bacteroidetes members (e.g. species of *Prevotella*) to those of Proteobacteria (e.g. *Haemophilus* species) [119, 165, 166]. Intriguingly, Proteobacteria are a predominant component of the skin microbiome, suggesting detrimental transposition of skin associated microbiota into the lungs [134, 138]. Similarly, analysis of fungal contributions to the pulmonary microbiome implicates the presence of *Malassezia* species in asthma [167]. This particular fungal species is more prominently known as a factor in atopic and seborrhoeic dermatitis, further supporting a role of deleterious skin microbiota cross-over in asthma [143]. It remains to be determined whether translocation of undesirable microbiota exacerbates the disease, or rather, the condition establishes a dysfunctional environment that is more conducive to infiltration of foreign microbes.

Although elaborate understanding of the respiratory microbiome's general functionality is lacking, evidence has been presented for alterations of the asthmatic lung microbiome composition within the context of corticosteroid treatment [162, 166]. Patients with asthma either resistant or sensitive to corticosteroid treatment both demonstrated compositional derangements in comparison to controls, with reduced Bacteroidetes abundance, and increased levels of Proteobacteria and Actinobacteria species [162]. Additionally, host-derived peripheral blood monocytes from the lungs of corticosteroid resistant patients had inhibited corticosteroid response when co-cultured with an isolate of *Haemophilus parainfluenzae*, a potential pathogen associated with asthma [162].

Although initial findings of the respiratory microbiome are promising, future developments face significant challenges. While other body sites are highly accessible for sampling with topical swabbing or secretion collection, acquiring samples from the lung is considerably more complicated. Not only is advanced training necessary to safely reach the tissue of the lower respiratory tract, but contamination risk throughout the respiratory system becomes exceedingly high when attempting to sample from specific pulmonary sites [155, 156, 158, 168].

1.8.4 | The Vagina and Urinary Tract

The female urogenital microbiome maintains a paramount health role, seeding the microbiome of infants passing through the birth canal. Contribution of the mother's urogenital microbiome during vaginal birth (or skin with caesarean delivery) on the establishment of an individual's microbiome marks a critical stage of biological development, with potentially life-long influences on asthma and allergy susceptibility, obesity, and neurocognitive development [90, 91, 169-171].

Highlighted by its epidemiological influence, substantial effort has been put towards characterisation of vaginal microbial components and associated metabolic function (see Figure 1.1D). Similarly to the lungs, the healthy vaginal microbiome is characterized as maintaining low microbial diversity, however *Lactobacillus* species dominate its composition [25, 131, 172]. Disruptions to the healthy vaginal microbiome's stable low-complexity are linked to severity of cervical intra-epithelial neoplasia (CIN) and bacterial vaginosis (BV), of which the latter presents increased risk of sexually transmitted infection acquisition, pelvic inflammatory disease, and preterm birth [129, 173-178].

Lactobacillus dominance of the vaginal microbiome appears to be unique in humans when compared to other animals (>70% and ~1%, respectively) [179]. In response to this anomaly, several theories have been proposed for the *Lactobacillus*-centric human vaginal microbiome: conserved common function of vaginal microorganisms that in humans happens to be fulfilled by *Lactobacillus* species, and such microbes adapting to starch rich diets that are typical of humans [179]. The diet hypothesis further suggests that high glycogen concentrations found within the human vaginal tract reflect dietary carbohydrate catabolism which is facilitated by abundant salivary amylase levels.

Irrespective of evolutionary basis, thriving lactobacilli growth is fostered in the vaginal environment by glycoprotein and mucin rich genital fluid and high levels of glycogen and α-amylase, of which the latter increases energy availability of glycogen through its by-products [179-181]. With *Lactobacillus* proliferation, oestrogen mediated pH of the vagina is further acidified by microbial-derived lactic acid, which is metabolized from glycogen through anaerobic glycolysis [182-187]. Low pH (~3.5) and high lactic acid concentrations contribute in conjunction with cervicovaginal fluid, a highly effective antimicrobial and antiviral medium, to stabilization of the vaginal environment [185, 187]. With BV, when

vaginal pH rises (>4.5) and microbial composition shifts away from lactobacilli to *Gardnerella* species, lactic acid levels drop and a more prominent SCFA profile develops [185]. While SCFAs are generally observed to bestow health benefits, particularly in the gut, within the vaginal tract an undesirable pro-inflammatory response appears to be induced by acetate and butyrate [55, 74, 128, 141, 185, 188].

The vaginal microbiome's dichotomous state between *Lactobacillus* and *Gardnerella* dominance appears to considerably influence the efficacy of microbicide HIV prevention therapy [129]. Tenofovir microbicide gel was 59.2% effective in HIV infection prevention for *Lactobacillus* dominant vaginal communities, while the *Gardnerella* based BV microbiome had a prevention rate of 18% [129]. Controlled doses of tenofovir administered to patients with either *Gardnerella* or *Lactobacillus* oriented microbiomes showed significantly lower concentrations of the drug in *Gardnerella* dominated vaginal communities, and indeed, drug concentration negatively correlated with *Gardnerella* abundance [129]. *In vitro* analysis demonstrated that *Gardnerella* and other BV associated microbes efficiently metabolised the drug through a cleavage of an oxy-methylphosphonic acid side-chain of the compound [129].

The male urogenital tract microbiome has received less attention; however budding investigation of the subject suggests health-relevant microbial activity within this system. Circumcision significantly modifies microbial composition of the coronal sulci of the penis, decreasing the total microbial load, including anaerobic taxa putatively associated with BV [189, 190]. Reduced HIV infection rates have independently been associated with circumcision, however the underlying factors of this protective effect are unknown [191]. Characterization of the penile microbiome suggests that microbial modulations from circumcision reduce epithelial immune activity, subsequently reducing HIV presentation to

the immune system [189-191]. Although the human urogenital microbiome has already begun to yield promising insights to a variety of important aspects of health, there continues to be numerous questions that remain unanswered.

1.8.5 | The Gut

Of the microbial communities delineated by human physiology, the consortium of prokaryotes, viruses, Archaea, and eukaryotes found within the gastrointestinal (GI) system have been investigated with the greatest intensity, subsequently yielding the most elaborate insights of microbial contributions to health and disease (see Figure 1.1 E) [12, 21, 27, 29].

Present knowledge of microbial introduction to the GI tract dictates uni-directional passage of microbiota within ingested material. Microbes are transported along the GI system, following a gradient of community complexity that peaks in the colon [192-194]. Once established, the gut microbiome is subject to influence from a growing number of known factors, both from the host and other microbial cohabitants. Perhaps most profoundly affecting this community is host diet, supplying microbes and the arsenal of nutrients that contribute to the microbiome's function and proliferation [53, 55, 102, 195, 196]. Plant-based complex carbohydrates, which intestinal microbiota process with enzymes that are absent from the human host, keenly exemplify dietary influence on the microbiome [55, 196, 197]. Through metabolism of these polysaccharides, microbial fermentation yields SCFAs, compounds with a broad range of purportedly profound effects on the host [55, 196, 197].

Host-derived metabolites, similarly to dietary constituents, are utilized by the microbiome [196, 198-201]. Examples highlighting this host-microbe interaction include: bile acids (BA), which upon metabolization by bacteria, engage in complex host-microbe signalling cascades, and intestinal mucins, compounds utilized by mucin specialists (e.g.

Akkermansia muciniphila), providing protective properties to the host [38, 196, 198-201]. In addition to drugs explicitly affecting microorganisms, i.e. antibiotics, other medications demonstrate profound microbial interactions, affecting microbes and pharmacokinetics of the drugs [200, 202-205]. An *in vitro* screen of more than 1000 pharmaceutical compounds against core representative strains of gut bacteria, demonstrated that growth of at least one strain was inhibited by 24% of compounds intended to target human cells [205]. The type II diabetes drug metformin has separately been implicated in altering both composition and function of human intestinal microbiota, moulding a microbiome with an enrichment of SCFA metabolism genes and faecal concentrations of propionate and butyrate [204]. Specific microbial metabolic interaction with metformin however, remains to be elucidated.

Indeed, drugs of intoxication (e.g. alcohol and cannabis) are indicated as having microbiome interactions, although such insights are sparse and often limited to non-human animal models [46, 104, 206-209]. An exception to the limited reporting on recreational drug use and the microbiome is the profiling of the gut microbiome of chronic cannabis users [208]. It was demonstrated that in comparison to controls, chronic cannabis users had a 13-fold reduction in the *Prevotella:Bacteroides* ratio. Lower *Prevotella* abundance was further associated with poor cognition test performance and reduced mitochondrial ATP production [208].

Host behaviour, specifically physical exercise, is also recognized as an integral factor in modulation of the intestinal microbiome's composition and function [57-61, 105-110]. Illustrating the influence of the extremes of exercise, professional athletes have been shown to harbour a gut microbiome, that in addition to having high compositional diversity of microbial taxa, contains a gene profile with robust potential for environmental energy capture [57, 58]. Specifically with the cohort of professional rugby players, in comparison to

age matched controls with similar BMI to represent the range of body composition in the athletes, a substantial increase in metabolic pathways with potential for health benefit, and a unique profile of faecal metabolites were observed. The pathways, which, given levels of expression matching their abundance offer opportunity for health modulation, ranged from organic cofactor and antibiotic biosynthesis to degradation and biosynthesis of carbohydrates. The latter of which could potentially result in an increased capacity for energy utilisation by the microbiome [57]. Metabolomic profiling of the athlete gut microbiome revealed elevated levels of SCFAs, metabolites with wide health implications (detailed further below) and association to a lean body composition [210]. The faecal metabolome of these athletes also presented elevated levels of trimethylamine-N-oxide (TMAO), a compound that has been associated with cardiovascular disease and atherosclerosis, although these negative associations have been disputed due the occurrence of high levels of TMAO in populations with low occurrence of cardiovascular disease [211]. Examining the microbiome of high-performance bicyclists, it was shown that the genus of *Prevotella* was significantly associated with reported time of exercising, while transcriptional levels of *Methanobrevibacter smithii* products were upregulated, particularly so for methanogenesis genes [60]. Investigation of amateur half-marathon runners demonstrated that through the course of high-intensity running, significant changes occurred in certain taxa (e.g. *Coriobacteriaceae*) and metabolites within the gut environment [61]. Intriguingly, the introduction of exercise as a novel stimulus appears to elicit more subtle changes in the gut microbiome. After undergoing a short period (8 weeks) of moderate-intensity exercise, healthy but inactive adults were shown to have only minor changes in the composition of their gut microbiome [59]. A separate analysis of a combination of lean and obese individuals undergoing a period of structured exercise conversely asserted that concentra-

tions of faecal SCFAs increased in lean participants following exercise, while an obesity dependant shift in microbial diversity was present after exercise, and dissipated after a washout period [56]. Taken together, there still remains work to be done in understanding the mechanisms underlying the interaction of exercise and the gut microbiome.

Gut microbiome analysis is predominantly carried out on the terminal end of the GI tract with the large intestine. Focus on the large intestine results from the relative ease and reliability by which gut specimens can be non-invasively acquired as stool samples. These samples importantly provide accurate representation of the intestinal microbiome as excreted samples retain microbial cells and metabolites from the lumen and mucosa, however, without exact recapitulation of the intestine's various sub-sites [192, 193, 212]. From these samples, understanding of the gut microbiome has advanced through utilization of culture-dependant methodologies, to microbial taxonomic characterization with 16S rRNA gene sequencing, and on to the functional potential of the microbiome, progressing in stride with advancements in genetic sequencing and metabolomic technologies [25, 28, 54, 57, 213].

These collective analyses have revealed numerous roles for the gut microbiome in a broad range of disease states. Early sequencing based studies of the intestinal microbiome presented an influence of the microbiome in obesity, first with mice and later with humans [28, 214]. This early work paved the way for numerous other proposed health insults resulting from undesirable intestinal microbiota composition or function, ranging from diabetes and inflammatory bowel disease to psychiatric disorder and cancer [24, 215-222]. While initial contemporary gut microbiome research presented microbiota taxonomically, which drastically limits potential understanding of microbial influence on disease, shotgun metagenomic sequencing and metabolomic analysis strategies place the microbiome within

a framework of genetic pathways and metabolic products, enabling more elaborate investigations of mechanistic interactions between host and microbe [38, 46, 57, 213, 223].

1.9 | Systemic Implication of the Gut Microbiome in Health and Disease

The GI tract acts as the primary site for the uptake and metabolic processing of nutrients, and subsequent systemic allocations of such products. As the human body's central portal, the gut plays substantial role in health regulation. Gut microbiome investigation supports a similarly crucial role of intestinal microbes in health maintenance and modulation of various disease states, compounding with native biology of the host intestinal system. Microbial contributions to this health dynamic are mediated by numerous metabolic modalities operating independently or in concert. The most prominent metabolic circuit is at the interface of the microbiome and ingested nutrients, wherein microbiota utilize host diet to proliferate and produce metabolites that integrate with host cell metabolism (see Figure 1.2) [29, 43, 53, 195, 196, 224, 225].

1.9.1 | Short-chain fatty acids

Of microbially produced metabolites influencing the host, SCFAs have demonstrated immense contribution to host biology, both locally within the intestinal system and distally with hepatic, neurological, and immunological function [55, 188, 225-229]. As previously noted, microbial SCFA generation occurs through polysaccharide utilization. Although the primary substrate for production of these compounds, it has also been demonstrated that gut microbiota have the capacity to convert butyrate from the metabolism of protein [225, 230-232]. Primary SCFA biosynthesis from polysaccharides is facilitated by various gut microbiota, employing a diverse repertoire of metabolic and enzymatic capacities [55, 225, 233]. Acetate, the SCFA generally found in the highest concentrations of intestinal SCFA profiling, can be derived from pyruvate metabolism,

following the Wood-Ljungdahl pathway or acetyl-coA synthesis [55, 225, 233]. Propionate is produced via two known processes involving sugar fermentation pathways, the succinate and propanediol pathways [55, 225]. Carbohydrate-derived propionate predominantly occurs through the succinate pathway, utilizes hexose and pentose sugars (e.g. glucose and ribose, respectively), and is dependent on vitamin B_{12} for complete conversion of succinate to propionate [225]. Propionate biosynthesis mediated by the propanediol pathway uses deoxy sugars fucose and rhamnose as the main reactants, however lactate conversion to lactaldehyde by *Lactobacillus buchneri* can feed generation of 1,2-propanediol [225]. Butyrate, the third most common SCFA, is most frequently formed through glycolysis of carbohydrates in which two acetyl-CoA molecules are combined to acetoacetyl-CoA, which is then reduced to butyryl-CoA [55, 225]. Final formation of butyrate within this metabolic scheme is carried out through two known pathways; butyryl-CoA: acetate CoA-transferase or phosphotransbutyrylase and butyrate kinase [55, 225].

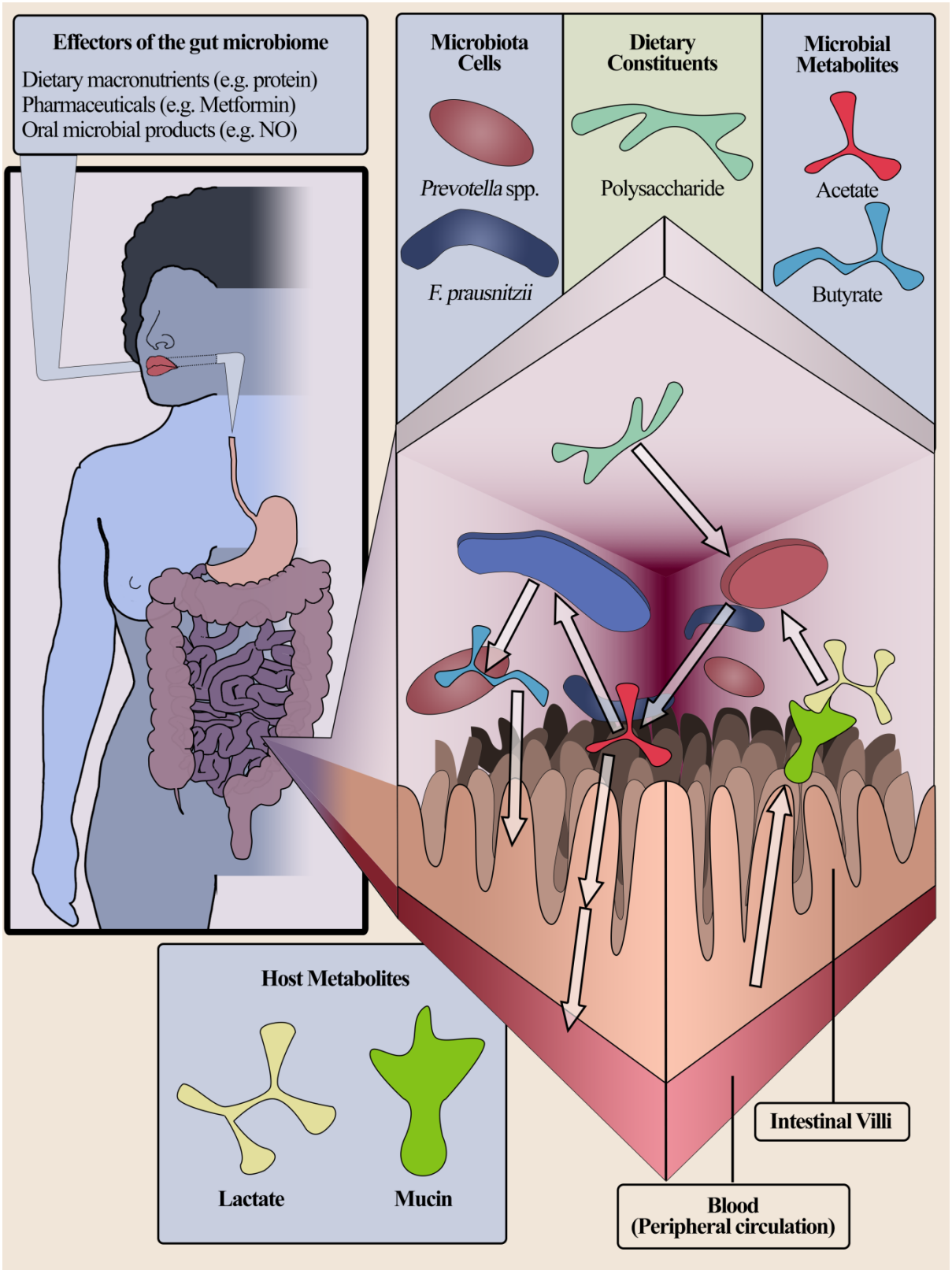


Figure 1.2 | Host-microbe metabolic interaction.

Figure 1.2 continued | A simplified demonstration of the metabolic interactions between host and microbiome. The cross-section of the small intestine illustrates the metabolic exchange between the intestine and two taxonomic representatives (*Prevotella* spp. and *F. prausnitzii*). Polysaccharides act as an example of dietary substrate utilized by the microbiota for the production of SCFA (butyrate and acetate). Similarly, host-derived substrate in the form of lactate presented with excretion of mucin from the intestine can be used by the microbiota. Within the example, acetate can either be absorbed by the intestine and subsequently the blood stream where systematic influences take place, or converted to butyrate, exerting localized effect on intestinal epithelial cells.

Upon excretion from microbial cells, SCFAs entering the intestinal environment are utilized by colonocytes as an energy source or pass into broader circulation via the portal vein [55, 225]. Acting locally on colonocytes, butyrate is incorporated into luminal cells through diffusion or direct transport mediated by the Na(+)-coupled transporter SLC5A8 [55, 234]. Butyrate within colonocytes contributes to energy production through conversion of the compound to acetyl-CoA, or alternatively will inhibit histone deacetylase (HDAC) activity [55, 234, 235]. HDAC inhibition occurs within colorectal cancer cells, wherein glucose is preferentially used as an energy source, leading to butyrate accumulation and the subsequent action upon HDAC which results in a cascade of effect on cell proliferation, differentiation, and apoptosis [55, 234, 235].

Upon entering wider circulation through the portal vein, propionate is primarily metabolized in the liver while acetate is more broadly circulated, for example, crossing the blood-brain barrier where it may influence satiety through action on the hypothalamus [227]. Demonstrated in mice, gut-derived acetate and propionate have separately been suggested to influence asthma [55, 236, 237]. While regulatory T cell activity is enhanced by acetate inhibition of histone deacetylase 9 (HDAC9), resulting in suppression of environmental allergen hypersensitivity, propionate affects lung dendritic cells dampening promotion of T helper type 2 cell driven inflammation while leaving the cells' phagocytic ability intact [55, 117, 236-238].

Of particular relevance to this thesis, SCFAs have been shown to be detected by metabolomic profiling in higher concentrations within the gut of athletes when compared to controls. Our work with elite athletes demonstrated that an enrichment of SCFAs is present in the intestinal environment of the athletes when compared to controls [57]. Separately it was shown that lean individuals undergoing a short-term exercise intervention had

increased levels of SCFAs in response to the treatment [56]. This intriguingly suggests that physical fitness may foster an environment within the gut that, likely in conjunction with dietary changes associated with increased exercise, is more conducive to the production of SCFAs

1.9.2 | Bile acids

With similarly profound implications as SCFAs, bile acids (BA) are shown to be at the centre of a metabolic interplay between host and microbe [53, 198, 199, 202, 204, 239-241]. Following post-meal metabolic cues, bile released from the canalicular membrane of hepatocytes enters the intestinal system. Primary BAs, cholic acid and chenodeoxycholic acid, are converted from cholesterol and conjugated to taurine or glycine, and within context of host physiology, are utilized as detergents to allow intestinal absorption of dietary lipids and fat-soluble vitamins [239, 241, 242].

Microbial bile salt hydrolase (BSH) facilitates conjugated bile acid (CBA) hydrolysis, reverting the compounds back to BAs which permits small intestine reabsorption or additional metabolic processing [241, 242]. Unconjugated and glycine-conjugated BA absorption by passive diffusion and active transport creates a circulating pool of BAs, establishing continuous bioavailability of the compounds [240-242]. As detergents, BAs have the capacity to disrupt the lipid membrane of bacterial cells, subsequently exerting considerable influence on the structure of the microbiome. Microbes accordingly employ myriad strategies to circumvent the antimicrobial action of BAs, such as outer membrane lipid and protein modifications [241, 242]. In conjunction with BA resistance, microbiota exploit their ability to alter BA hydrophobicity through modification of the compounds, permitting either evasion of destruction or an inhospitable environment for competing

organisms [241, 242]. Microbial-BA activity impacting host health has consequentially begun to be elucidated.

Microbial BSH driven hydrolysis of CBAs to unconjugated primary BAs enables subsequent conversion to secondary BAs deoxycholic acid and lithocholic acid (DCA and LCA, respectively) [241, 242]. DCA in particular, accumulates in the enterohepatic BA pool. Relatively high concentrations of DCA result from intestinal diffusion and hepatic reuptake conduced by the compound's hydrophobicity, and the human liver's inability to rehydroxylate DCA [241].

Obesity is a health condition with demonstrated influence by BAs, for example dietary insult to weight-gain and glycaemic control was ameliorated with intervention of BA binding resins [243]. Anatomical influence on BA Roux-en-Y gastric bypass surgery has intriguingly been shown to also have an effect on BAs, with serum concentrations raised in individuals that have undergone the procedure when compared to obese and severely obese controls, suggesting that anatomical manipulation of the procedure modifies the dynamics of the BA pool [244, 245]. The fat and protein enriched 'Western' diet that contributes to obesity development modifies not only gut microbiome composition but also microbial BA pool contributions [53, 196, 214, 216, 240].

Among the numerous detrimental effects of obesity, evidence supports microbial derived DCA as a potent tumour promoter, contributing to development of hepatocellular carcinoma and the colorectal cancer precursor colorectal adenomas [53, 240, 246-248]. Although mechanistic detail in the human gut is limited, DCA-driven hepatocellular carcinoma in mice is suggested to result from the compound's provocation of the senescence-associated secretory phenotype (SASP) in hepatic stellate cells [247]. SASP is characterized by broad alterations in gene expression and secretory profile, which affect

neighbouring cells through numerous factors: release of cytokines (e.g. IL-1α and -β), insulin-like growth factor binding proteins, nitric oxide and reactive oxygen species, and potentially the glycoprotein, fibronectin [247, 249]. Influence of DCA on colorectal tumorigenesis is proposed to mediate derangement of epidermal growth factor receptor-mitogen activated protein kinase (EGFR-MAPK) regulation, specifically with DCA preventing degradation of EGFR through calcium signalling of MAPK [246]. Within this valuable framework established around the interaction of gut microbes and BAs, a wide range of metabolic facets remain to be elucidated. There continue to be poorly understood aspects of the roles BAs have on microbial function and ecology, and indeed, the influences of microbial-derived BAs on host health. Furthermore, the dynamics of BA-microbiome-host interaction are only a small component of the numerous bioactive compounds within the gut environment, highlighting the vast areas of the microbiome that still require investigation.

1.10 | Conclusions and Outlook

Examination of microbiome-host interaction has defined an integral role of microbiota in health and disease. The challenge now lies in comprehensive elucidation of the underlying mechanisms. Metabolic phenotyping and identification of the microbial metabolites interacting with the host will be pivotal to this challenge. With robust characterization of metabolic activity within the various microbiomes of the human body, progress can be made in the development of defined microbial cultures (probiotics) and substrates conducive to selective growth or function of microbes (prebiotics) for health enhancement. In short, there is need and opportunity for the imaginative deployment of metabolic phenotyping within the field of microbiome research.

1.11 | Hypothesis

The research of this thesis is directed towards expanding the understanding of physical exercise as an agent in the modification of the gut microbiome, particularly in how the microbiome pertains to health. This goal encompasses the profiling of intestinal microbiota, both in terms of taxonomy and metabolic function, within the context of various degrees of exercise intensity, and human populations.

The specific hypotheses of this thesis are as follows:

1. The athlete microbiome exhibits unique characteristics, including functional adaptations to the dietary and energetic extremes associated to the demands of high-performance sport.
2. Identifiable changes occur in the composition and function of the gut microbiome in response to exercise and/or protein supplementation as novel stimuli.
3. Computational based investigations can be utilized to elucidate additional details of the microbiome dynamics in athletic populations.
4. Exercise can be used by inflammatory bowel disease patients without detrimental impact on their intestinal microbiome.
5. Prolonged engagement in physical activity results in pronounced alterations of the gut microbiome.

1.12 | Objective

The objectives of this thesis are as follows:

1. Describe the taxonomic and metabolic potential profiles for the gut microbiome of high-performance athletes.
2. Identify alterations in the gut microbiome of exercise naïve individuals in specific response to physical exercise, whey protein supplementation, or both.
3. Assess the viability of exercise as an adjunct treatment in physically inactive inflammatory bowel disease patients.
4. Evaluate the influence of moderate intensity exercise on the microbiome longitudinally over the course of 6 months.
5. Determine the effectiveness of elucidating additional details of gut microbiome dynamics with computational modelling strategies.

1.13 | References

[1] W. Xie, F. Wang, L. Guo, Z. Chen, S. M. Sievert, J. Meng, G. Huang, Y. Li, Q. Yan, S. Wu, X. Wang, S. Chen, G. He, X. Xiao, A. Xu, Comparative metagenomics of microbial communities inhabiting deep-sea hydrothermal vent chimneys with contrasting chemistries, ISME J 5 (3) (2011) 414-26.

[2] E. Afshinnekoo, C. Meydan, S. Chowdhury, D. Jaroudi, C. Boyer, N. Bernstein, J. M. Maritz, D. Reeves, J. Gandara, S. Chhangawala, S. Ahsanuddin, A. Simmons, T. Nessel, B. Sundaresh, E. Pereira, E. Jorgensen, S. O. Kolokotronis, N. Kirchberger, I. Garcia, D. Gandara, S. Dhanraj, T. Nawrin, Y. Saletore, N. Alexander, P. Vijay, E. M. Henaff, P. Zumbo, M. Walsh, G. D. O'Mullan, S. Tighe, J. T. Dudley, A. Dunaif, S. Ennis, E. O'Halloran, T. R. Magalhaes, B. Boone, A. L. Jones, T. R. Muth, K. S. Paolantonio, E. Alter, E. E. Schadt, J. Garbarino, R. J. Prill, J. M. Carlton, S. Levy, C. E. Mason, Geospatial Resolution of Human and Bacterial Diversity with City-Scale Metagenomics, Cell Syst 1 (1) (2015) 72-87.

[3] J. F. Ruiz-Calderon, H. Cavallin, S. J. Song, A. Novoselac, L. R. Pericchi, J. N. Hernandez, R. Rios, O. H. Branch, H. Pereira, L. C. Paulino, M. J. Blaser, R. Knight, M. G. Dominguez-Bello, Walls talk: Microbial biogeography of homes spanning urbanization, Sci Adv 2 (2) (2016) e1501061.

[4] L. M. Coughlan, P. D. Cotter, C. Hill, A. Alvarez-Ordonez, New Weapons to Fight Old Enemies: Novel Strategies for the (Bio)control of Bacterial Biofilms in the Food Industry, Front Microbiol 7 (2016) 1641.

[5] B. C. Bourrie, B. P. Willing, P. D. Cotter, The Microbiota and Health Promoting Characteristics of the Fermented Beverage Kefir, Front Microbiol 7 (2016) 647.

[6] C. J. Doyle, D. Gleeson, P. W. O'Toole, P. D. Cotter, High-throughput metataxonomic characterization of the raw milk microbiota identifies changes reflecting lactation stage and storage conditions, Int J Food Microbiol 255 (2017) 1-6.

[7] A. M. Walsh, F. Crispie, K. Daari, O. O'Sullivan, J. C. Martin, C. T. Arthur, M. J. Claesson, K. P. Scott, P. D. Cotter, Strain-level metagenomic analysis of the fermented dairy beverage nunu highlights potential food safety risks, Appl Environ Microbiol (2017)

[8] A. J. McHugh, C. Feehily, C. Hill, P. D. Cotter, Detection and Enumeration of Spore-Forming Bacteria in Powdered Dairy Products, Front Microbiol 8 (2017) 109.

[9] K. Venkateswaran, P. Vaishampayan, J. Cisneros, D. L. Pierson, S. O. Rogers, J. Perry, International Space Station environmental microbiome - microbial inventories of ISS filter debris, Appl Microbiol Biotechnol 98 (14) (2014) 6453-66.

[10] N. A. Be, A. Avila-Herrera, J. E. Allen, N. Singh, A. Checinska Sielaff, C. Jaing, K. Venkateswaran, Whole metagenome profiles of particulates collected from the International Space Station, Microbiome 5 (1) (2017) 81.

[11] P. D. Scanlan, Blastocystis: past pitfalls and future perspectives, Trends Parasitol 28 (8) (2012) 327-34.

[12] P. D. Scanlan, R. Knight, S. J. Song, G. Ackermann, P. D. Cotter, Prevalence and genetic diversity of Blastocystis in family units living in the United States, Infect Genet Evol 45 (2016) 95-97.

[13] P. D. Scanlan, C. R. Stensvold, M. Rajilic-Stojanovic, H. G. Heilig, W. M. De Vos, P. W. O'Toole, P. D. Cotter, The microbial eukaryote Blastocystis is a prevalent and diverse member of the healthy human gut microbiota, FEMS Microbiol Ecol 90 (1) (2014) 326-30.

[14] S. L. Burgess, C. A. Gilchrist, T. C. Lynn, W. A. Petri, Jr., Parasitic protozoa and interactions with the host intestinal microbiota, Infect Immun (2017)

[15] A. Chudnovskiy, A. Mortha, V. Kana, A. Kennard, J. D. Ramirez, A. Rahman, R. Remark, I. Mogno, R. Ng, S. Gnjatic, E. D. Amir, A. Solovyov, B. Greenbaum, J. Clemente, J. Faith, Y. Belkaid, M. E. Grigg, M. Merad, Host-Protozoan Interactions Protect from Mucosal Infections through Activation of the Inflammasome, Cell 167 (2) (2016) 444-456 e14.

[16] K. Hanevik, V. Dizdar, N. Langeland, T. Hausken, Development of functional gastrointestinal disorders after Giardia lamblia infection, BMC Gastroenterol 9 (2009) 27.

[17] M. A. Ghannoum, R. J. Jurevic, P. K. Mukherjee, F. Cui, M. Sikaroodi, A. Naqvi, P. M. Gillevet, Characterization of the oral fungal microbiome (mycobiome) in healthy individuals, PLoS Pathog 6 (1) (2010) e1000713.

[18] G. B. Huffnagle, M. C. Noverr, The emerging world of the fungal microbiome, Trends Microbiol 21 (7) (2013) 334-41.

[19] C. E. Huseyin, P. W. O'Toole, P. D. Cotter, P. D. Scanlan, Forgotten fungi-the gut mycobiome in human health and disease, FEMS Microbiol Rev (2017)

[20] L. D. Nguyen, E. Viscogliosi, L. Delhaes, The lung mycobiome: an emerging field of the human respiratory microbiome, Front Microbiol 6 (2015) 89.

[21] S. Minot, R. Sinha, J. Chen, H. Li, S. A. Keilbaugh, G. D. Wu, J. D. Lewis, F. D. Bushman, The human gut virome: inter-individual variation and dynamic response to diet, Genome Res 21 (10) (2011) 1616-25.

[22] G. D. Hannigan, Q. Zheng, J. S. Meisel, S. S. Minot, F. D. Bushman, E. A. Grice, Evolutionary and functional implications of hypervariable loci within the skin virome, PeerJ 5 (2017) e2959.

[23] H. W. Virgin, The virome in mammalian physiology and disease, Cell 157 (1) (2014) 142-50.

[24] J. M. Norman, S. A. Handley, M. T. Baldridge, L. Droit, C. Y. Liu, B. C. Keller, A. Kambal, C. L. Monaco, G. Zhao, P. Fleshner, T. S. Stappenbeck, D. P. McGovern, A. Keshavarzian, E. A. Mutlu, J. Sauk, D. Gevers, R. J. Xavier, D. Wang, M. Parkes, H. W. Virgin, Disease-specific alterations in the enteric virome in inflammatory bowel disease, Cell 160 (3) (2015) 447-60.

[25] C. Human Microbiome Project, Structure, function and diversity of the healthy human microbiome, Nature 486 (7402) (2012) 207-14.

[26] E. K. Costello, C. L. Lauber, M. Hamady, N. Fierer, J. I. Gordon, R. Knight, Bacterial community variation in human body habitats across space and time, Science 326 (5960) (2009) 1694-7.

[27] S. R. Gill, M. Pop, R. T. Deboy, P. B. Eckburg, P. J. Turnbaugh, B. S. Samuel, J. I. Gordon, D. A. Relman, C. M. Fraser-Liggett, K. E. Nelson, Metagenomic analysis of the human distal gut microbiome, Science 312 (5778) (2006) 1355-9.

[28] P. J. Turnbaugh, M. Hamady, T. Yatsunenko, B. L. Cantarel, A. Duncan, R. E. Ley, M. L. Sogin, W. J. Jones, B. A. Roe, J. P. Affourtit, M. Egholm, B. Henrissat, A. C. Heath, R. Knight, J. I. Gordon, A core gut microbiome in obese and lean twins, Nature 457 (7228) (2009) 480-4.

[29] A. Zhernakova, A. Kurilshikov, M. J. Bonder, E. F. Tigchelaar, M. Schirmer, T. Vatanen, Z. Mujagic, A. V. Vila, G. Falony, S. Vieira-Silva, J. Wang, F. Imhann, E. Brandsma, S. A. Jankipersadsing, M. Joossens, M. C. Cenit, P. Deelen, M. A. Swertz, s. LifeLines cohort, R. K. Weersma, E. J. Feskens, M. G. Netea, D. Gevers, D. Jonkers, L. Franke, Y. S. Aulchenko, C. Huttenhower, J. Raes, M. H. Hofker, R. J. Xavier, C. Wijmenga, J. Fu, Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity, Science 352 (6285) (2016) 565-9.

[30] P. D. Scanlan, Bacteria-Bacteriophage Coevolution in the Human Gut: Implications for Microbial Diversity and Functionality, Trends Microbiol (2017)

[31] K. Z. Coyte, J. Schluter, K. R. Foster, The ecology of the microbiome: Networks, competition, and stability, Science 350 (6261) (2015) 663-6.

[32] M. J. Kennedy, P. A. Volz, Ecology of Candida albicans gut colonization: inhibition of Candida adhesion, colonization, and dissemination from the gastrointestinal tract by bacterial antagonism, Infect Immun 49 (3) (1985) 654-63.

[33] C. S. Smillie, M. B. Smith, J. Friedman, O. X. Cordero, L. A. David, E. J. Alm, Ecology drives a global network of gene exchange connecting the human microbiome, Nature 480 (7376) (2011) 241-4.

[34] J. Walter, R. Ley, The human gut microbiome: ecology and recent evolutionary changes, Annu Rev Microbiol 65 (2011) 411-29.

[35] A. M. Walsh, F. Crispie, K. Kilcawley, O. O'Sullivan, M. G. O'Sullivan, M. J. Claesson, P. D. Cotter, Microbial Succession and Flavor Production in the Fermented Dairy Beverage Kefir, mSystems 1 (5) (2016)

[36] U. M. McCormack, T. Curiao, S. G. Buzoianu, M. L. Prieto, T. Ryan, P. Varley, F. Crispie, E. Magowan, B. U. Metzler-Zebeli, D. Berry, O. O'Sullivan, P. D. Cotter, G. E. Gardiner, P.

G. Lawlor, Exploring a Possible Link between the Intestinal Microbiota and Feed Efficiency in Pigs, Appl Environ Microbiol 83 (15) (2017)

[37] D. G. Bourne, K. M. Morrow, N. S. Webster, Insights into the Coral Microbiome: Underpinning the Health and Resilience of Reef Ecosystems, Annu Rev Microbiol 70 (2016) 317-40.

[38] C. Belzer, L. W. Chia, S. Aalvink, B. Chamlagain, V. Piironen, J. Knol, W. M. de Vos, Microbial Metabolic Networks at the Mucus Layer Lead to Diet-Independent Butyrate and Vitamin B12 Production by Intestinal Symbionts, MBio 8 (5) (2017)

[39] G. Clarke, R. M. Stilling, P. J. Kennedy, C. Stanton, J. F. Cryan, T. G. Dinan, Minireview: Gut microbiota: the neglected endocrine organ, Mol Endocrinol 28 (8) (2014) 1221-38.

[40] J. M. Evans, L. S. Morris, J. R. Marchesi, The gut microbiome: the role of a virtual organ in the endocrinology of the host, J Endocrinol 218 (3) (2013) R37-47.

[41] C. M. Guinane, P. D. Cotter, Role of the gut microbiota in health and chronic gastrointestinal disease: understanding a hidden metabolic organ, Therap Adv Gastroenterol 6 (4) (2013) 295-308.

[42] M. P. Hezel, E. Weitzberg, The oral microbiome and nitric oxide homoeostasis, Oral Dis 21 (1) (2015) 7-16.

[43] S. Possemiers, S. Bolca, W. Verstraete, A. Heyerick, The intestinal microbiome: a separate organ inside the body with the metabolic potential to influence the bioactivity of botanicals, Fitoterapia 82 (1) (2011) 53-66.

[44] A. G. Clooney, C. N. Bernstein, W. D. Leslie, K. Vagianos, M. Sargent, E. J. Laserna-Mendieta, M. J. Claesson, L. E. Targownik, A comparison of the gut microbiome between long-term users and non-users of proton pump inhibitors, Aliment Pharmacol Ther 43 (9) (2016) 974-84.

[45] F. Imhann, M. J. Bonder, A. Vich Vila, J. Fu, Z. Mujagic, L. Vork, E. F. Tigchelaar, S. A. Jankipersadsing, M. C. Cenit, H. J. Harmsen, G. Dijkstra, L. Franke, R. J. Xavier, D. Jonkers, C. Wijmenga, R. K. Weersma, A. Zhernakova, Proton pump inhibitors affect the gut microbiome, Gut 65 (5) (2016) 740-8.

[46] I. D. Wilson, J. K. Nicholson, Gut microbiome interactions with drug metabolism, efficacy, and toxicity, Transl Res 179 (2017) 204-222.

[47] J. W. Hegarty, C. M. Guinane, R. P. Ross, C. Hill, P. D. Cotter, Bacteriocin production: a relatively unharnessed probiotic trait?, F1000Res 5 (2016) 2587.

[48] S. Lebeer, J. Vanderleyden, S. C. De Keersmaecker, Host interactions of probiotic bacterial surface molecules: comparison with commensals and pathogens, Nat Rev Microbiol 8 (3) (2010) 171-84.

[49] C. J. Walsh, C. M. Guinane, P. W. O'Toole, P. D. Cotter, Beneficial modulation of the gut microbiota, FEBS Lett 588 (22) (2014) 4120-30.

[50] M. J. Claesson, I. B. Jeffery, S. Conde, S. E. Power, E. M. O'Connor, S. Cusack, H. M. Harris, M. Coakley, B. Lakshminarayanan, O. O'Sullivan, G. F. Fitzgerald, J. Deane, M. O'Connor, N. Harnedy, K. O'Connor, D. O'Mahony, D. van Sinderen, M. Wallace, L. Brennan, C. Stanton, J. R. Marchesi, A. P. Fitzgerald, F. Shanahan, C. Hill, R. P. Ross, P. W. O'Toole, Gut microbiota composition correlates with diet and health in the elderly, Nature 488 (7410) (2012) 178-84.

[51] D. A. Cobb-Clark, S. C. Kassenboehmer, S. Schurer, Healthy habits: The connection between diet, exercise, and locus of control, Journal of Economic Behavior & Organization 98 (2014) 1-28.

[52] A. Cotillard, S. P. Kennedy, L. C. Kong, E. Prifti, N. Pons, E. Le Chatelier, M. Almeida, B. Quinquis, F. Levenez, N. Galleron, S. Gougis, S. Rizkalla, J. M. Batto, P. Renault, A. N. R. M. consortium, J. Dore, J. D. Zucker, K. Clement, S. D. Ehrlich, Dietary intervention impact on gut microbial gene richness, Nature 500 (7464) (2013) 585-8.

[53] L. A. David, C. F. Maurice, R. N. Carmody, D. B. Gootenberg, J. E. Button, B. E. Wolfe, A. V. Ling, A. S. Devlin, Y. Varma, M. A. Fischbach, S. B. Biddinger, R. J. Dutton, P. J. Turnbaugh, Diet rapidly and reproducibly alters the human gut microbiome, Nature 505 (7484) (2014) 559-63.

[54] S. M. Finegold, H. R. Attebery, V. L. Sutter, Effect of diet on human fecal flora: comparison of Japanese and American diets, Am J Clin Nutr 27 (12) (1974) 1456-69.

[55] A. Koh, F. De Vadder, P. Kovatcheva-Datchary, F. Backhed, From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites, Cell 165 (6) (2016) 1332-1345.

[56] J. M. Allen, L. J. Mailing, G. M. Niemiro, R. Moore, M. D. Cook, B. A. White, H. D. Holscher, J. A. Woods, Exercise Alters Gut Microbiota Composition and Function in Lean and Obese Humans, Med Sci Sports Exerc (2017)

[57] W. Barton, N. C. Penney, O. Cronin, I. Garcia-Perez, M. G. Molloy, E. Holmes, F. Shanahan, P. D. Cotter, O. O'Sullivan, The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level, Gut (2017)

[58] S. F. Clarke, E. F. Murphy, O. O'Sullivan, A. J. Lucey, M. Humphreys, A. Hogan, P. Hayes, M. O'Reilly, I. B. Jeffery, R. Wood-Martin, D. M. Kerins, E. Quigley, R. P. Ross, P. W. O'Toole, M. G. Molloy, E. Falvey, F. Shanahan, P. D. Cotter, Exercise and associated dietary extremes impact on gut microbial diversity, Gut 63 (12) (2014) 1913-20.

[59] O. Cronin, W. Barton, P. Skuse, N. C. Penney, I. Garcia-Perez, E. F. Murphy, T. Woods, H. Nugent, A. Fanning, S. Melgar, E. C. Falvey, E. Holmes, P. D. Cotter, O. O'Sullivan, M. G. Molloy, F. Shanahan, J. A. Gilbert, A Prospective Metagenomic and Metabolomic Analysis

of the Impact of Exercise and/or Whey Protein Supplementation on the Gut Microbiome of Sedentary Adults, mSystems 3 (3) (2018) e00044-18.

[60] L. M. Petersen, E. J. Bautista, H. Nguyen, B. M. Hanson, L. Chen, S. H. Lek, E. Sodergren, G. M. Weinstock, Community characteristics of the gut microbiomes of competitive cyclists, Microbiome 5 (1) (2017) 98.

[61] X. Zhao, Z. Zhang, B. Hu, W. Huang, C. Yuan, L. Zou, Response of Gut Microbiota to Metabolite Changes Induced by Endurance Exercise, Frontiers in Microbiology 9 (2018)

[62] A. Harkin, Muscling in on depression, N Engl J Med 371 (24) (2014) 2333-4.

[63] F. B. Benatti, B. K. Pedersen, Exercise as an anti-inflammatory therapy for rheumatic diseases-myokine regulation, Nat Rev Rheumatol 11 (2) (2015) 86-97.

[64] J. A. Hawley, A. Krook, Metabolism: One step forward for exercise, Nat Rev Endocrinol 12 (1) (2016) 7-8.

[65] L. Schwingshackl, B. Missbach, S. Dias, J. Konig, G. Hoffmann, Impact of different training modalities on glycaemic control and blood lipids in patients with type 2 diabetes: a systematic review and network meta-analysis, Diabetologia 57 (9) (2014) 1789-97.

[66] R. D. Adam, Biology of Giardia lamblia, Clin Microbiol Rev 14 (3) (2001) 447-75.

[67] R. Fayer, U. Morgan, S. J. Upton, Epidemiology of Cryptosporidium: transmission, detection and identification, Int J Parasitol 30 (12-13) (2000) 1305-22.

[68] J. I. Ravdin, R. L. Guerrant, Role of adherence in cytopathogenic mechanisms of Entamoeba histolytica. Study with mammalian tissue culture cells and human erythrocytes, J Clin Invest 68 (5) (1981) 1305-13.

[69] D. S. Lindsay, J. P. Dubey, B. L. Blagburn, Biology of Isospora spp. from humans, nonhuman primates, and domestic animals, Clin Microbiol Rev 10 (1) (1997) 19-34.

[70] M. Usyk, C. P. Zolnik, H. Patel, M. H. Levi, R. D. Burk, Novel ITS1 Fungal Primers for Characterization of the Mycobiome, mSphere 2 (6) (2017)

[71] M. A. Pfaller, D. J. Diekema, Epidemiology of invasive candidiasis: a persistent public health problem, Clin Microbiol Rev 20 (1) (2007) 133-63.

[72] C. N. Allonsius, M. F. L. van den Broek, I. De Boeck, S. Kiekens, E. F. M. Oerlemans, F. Kiekens, K. Foubert, D. Vandenheuvel, P. Cos, P. Delputte, S. Lebeer, Interplay between Lactobacillus rhamnosus GG and Candida and the involvement of exopolysaccharides, Microb Biotechnol 10 (6) (2017) 1753-1763.

[73] A. Rizzo, A. Losacco, C. R. Carratelli, Lactobacillus crispatus modulates epithelial cell defense against Candida albicans through Toll-like receptors 2 and 4, interleukin 8 and human beta-defensins 2 and 3, Immunol Lett 156 (1-2) (2013) 102-9.

[74] M. C. Noverr, G. B. Huffnagle, Regulation of Candida albicans morphogenesis by fatty acid metabolites, Infect Immun 72 (11) (2004) 6206-10.

[75] H. H. Kong, B. Andersson, T. Clavel, J. E. Common, S. A. Jackson, N. D. Olson, J. A. Segre, C. Traidl-Hoffmann, Performing Skin Microbiome Research: A Method to the Madness, J Invest Dermatol 137 (3) (2017) 561-568.

[76] A. Fleming, On the antibacterial action of cultures of a penicillium, with special reference to their use in the isolation of B. influenzae. , British journal of experimental pathology 10 (3) (1929) 226.

[77] F. W. Twort, An investigation on the nature of ultra-microscopic viruses, The Lancet 186 (4814) (1915) 1241-1243.

[78] F. d'Herelle, Sur un microbe invisible antagoniste des bacilles dysentériques, CR Acad. Sci. Paris 165 (1917) 373-375.

[79] C. R. Merril, D. Scholl, S. L. Adhya, The prospect for bacteriophage therapy in Western medicine, Nat Rev Drug Discov 2 (6) (2003) 489-97.

[80] A. Sulakvelidze, Z. Alavidze, J. G. Morris, Jr., Bacteriophage therapy, Antimicrob Agents Chemother 45 (3) (2001) 649-59.

[81] C. L. Ventola, The antibiotic resistance crisis: part 1: causes and threats, P T 40 (4) (2015) 277-83.

[82] C. L. Ventola, The Antibiotic Resistance Crisis: Part 2: Management Strategies and New Agents, Pharmacy and Therapeutics 40 (5) (2015) 344-352.

[83] P. Manrique, M. Dills, M. J. Young, The Human Gut Phage Community and Its Implications for Health and Disease, Viruses 9 (6) (2017)

[84] M. Arumugam, J. Raes, E. Pelletier, D. Le Paslier, T. Yamada, D. R. Mende, G. R. Fernandes, J. Tap, T. Bruls, J. M. Batto, M. Bertalan, N. Borruel, F. Casellas, L. Fernandez, L. Gautier, T. Hansen, M. Hattori, T. Hayashi, M. Kleerebezem, K. Kurokawa, M. Leclerc, F. Levenez, C. Manichanh, H. B. Nielsen, T. Nielsen, N. Pons, J. Poulain, J. Qin, T. Sicheritz-Ponten, S. Tims, D. Torrents, E. Ugarte, E. G. Zoetendal, J. Wang, F. Guarner, O. Pedersen, W. M. de Vos, S. Brunak, J. Dore, H. I. T. C. Meta, M. Antolin, F. Artiguenave, H. M. Blottiere, M. Almeida, C. Brechot, C. Cara, C. Chervaux, A. Cultrone, C. Delorme, G. Denariaz, R. Dervyn, K. U. Foerstner, C. Friss, M. van de Guchte, E. Guedon, F. Haimet, W. Huber, J. van Hylckama-Vlieg, A. Jamet, C. Juste, G. Kaci, J. Knol, O. Lakhdari, S. Layec, K. Le Roux, E. Maguin, A. Merieux, R. Melo Minardi, C. M'Rini, J. Muller, R. Oozeer, J. Parkhill, P. Renault, M. Rescigno, N. Sanchez, S. Sunagawa, A. Torrejon, K. Turner, G. Vandemeulebrouck, E. Varela, Y. Winogradsky, G. Zeller, J. Weissenbach, S. D. Ehrlich, P. Bork, Enterotypes of the human gut microbiome, Nature 473 (7346) (2011) 174-80.

[85] P. B. Eckburg, E. M. Bik, C. N. Bernstein, E. Purdom, L. Dethlefsen, M. Sargent, S. R. Gill, K. E. Nelson, D. A. Relman, Diversity of the human intestinal microbial flora, Science 308 (5728) (2005) 1635-8.

[86] M. G. Dominguez-Bello, E. K. Costello, M. Contreras, M. Magris, G. Hidalgo, N. Fierer, R. Knight, Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns, Proc Natl Acad Sci U S A 107 (26) (2010) 11971-5.

[87] F. Backhed, J. Roswall, Y. Peng, Q. Feng, H. Jia, P. Kovatcheva-Datchary, Y. Li, Y. Xia, H. Xie, H. Zhong, M. T. Khan, J. Zhang, J. Li, L. Xiao, J. Al-Aama, D. Zhang, Y. S. Lee, D. Kotowska, C. Colding, V. Tremaroli, Y. Yin, S. Bergman, X. Xu, L. Madsen, K. Kristiansen, J. Dahlgren, J. Wang, Dynamics and Stabilization of the Human Gut Microbiome during the First Year of Life, Cell Host Microbe 17 (5) (2015) 690-703.

[88] D. M. Chu, J. Ma, A. L. Prince, K. M. Antony, M. D. Seferovic, K. M. Aagaard, Maturation of the infant microbiome community structure and function across multiple body sites and in relation to mode of delivery, Nat Med 23 (3) (2017) 314-326.

[89] M. Nuriel-Ohayon, H. Neuman, O. Koren, Microbial Changes during Pregnancy, Birth, and Infancy, Front Microbiol 7 (2016) 1031.

[90] K. Korpela, P. Costea, L. P. Coelho, S. Kandels-Lewis, G. Willemsen, D. I. Boomsma, N. Segata, P. Bork, Selective maternal seeding and environment shape the human gut microbiome, Genome Res 28 (4) (2018) 561-568.

[91] N. T. Mueller, R. Whyatt, L. Hoepner, S. Oberfield, M. G. Dominguez-Bello, E. M. Widen, A. Hassoun, F. Perera, A. Rundle, Prenatal exposure to antibiotics, cesarean section and risk of childhood obesity, Int J Obes (Lond) 39 (4) (2015) 665-70.

[92] A. Sevelsted, J. Stokholm, K. Bonnelykke, H. Bisgaard, Cesarean section and chronic immune disorders, Pediatrics 135 (1) (2015) e92-8.

[93] P. Bager, J. Wohlfahrt, T. Westergaard, Caesarean delivery and risk of atopy and allergic disease: meta-analyses, Clin Exp Allergy 38 (4) (2008) 634-42.

[94] K. Negele, J. Heinrich, M. Borte, A. Berg, B. Schaaf, I. Lehmann, H. E. Wichmann, G. Bolte, Mode of delivery and development of atopic disease during the first 2 years of life, Pediatric Allergy and Immunology 15 (1) (2004) 48-54.

[95] J. J. Goedert, X. Hua, G. Yu, J. Shi, Diversity and composition of the adult fecal microbiome associated with history of cesarean birth or appendectomy: Analysis of the American Gut Project, EBioMedicine 1 (2-3) (2014) 167-172.

[96] M. G. Dominguez-Bello, K. M. De Jesus-Laboy, N. Shen, L. M. Cox, A. Amir, A. Gonzalez, N. A. Bokulich, S. J. Song, M. Hoashi, J. I. Rivera-Vinas, K. Mendez, R. Knight, J. C. Clemente, Partial restoration of the microbiota of cesarean-born infants via vaginal microbial transfer, Nat Med 22 (3) (2016) 250-3.

[97] T. Haahr, J. Glavind, P. Axelsson, M. Bistrup Fischer, J. Bjurstrom, G. Andresdottir, D. Teilmann-Jorgensen, U. Bonde, N. Olsen Sorensen, M. Moller, J. Fuglsang, P. G. Ovesen, J. P. Petersen, J. Stokholm, T. D. Clausen, Vaginal seeding or vaginal microbial transfer from

the mother to the caesarean-born neonate: a commentary regarding clinical management, BJOG (2017)

[98] A. J. Cunnington, K. Sim, A. Deierl, J. S. Kroll, E. Brannigan, J. Darby, "Vaginal seeding" of infants born by caesarean section, BMJ 352 (2016) i227.

[99] R. Cabrera-Rubio, M. C. Collado, K. Laitinen, S. Salminen, E. Isolauri, A. Mira, The human milk microbiome changes over lactation and is shaped by maternal weight and mode of delivery, Am J Clin Nutr 96 (3) (2012) 544-51.

[100] N. T. Mueller, E. Bakacs, J. Combellick, Z. Grigoryan, M. G. Dominguez-Bello, The infant microbiome development: mom matters, Trends Mol Med 21 (2) (2015) 109-17.

[101] P. Vangay, T. Ward, J. S. Gerber, D. Knights, Antibiotics, pediatric dysbiosis, and disease, Cell Host Microbe 17 (5) (2015) 553-64.

[102] N. A. Bokulich, J. Chung, T. Battaglia, N. Henderson, M. Jay, H. Li, D. L. A, F. Wu, G. I. Perez-Perez, Y. Chen, W. Schweizer, X. Zheng, M. Contreras, M. G. Dominguez-Bello, M. J. Blaser, Antibiotics, birth mode, and diet shape microbiome maturation during early life, Sci Transl Med 8 (343) (2016) 343ra82.

[103] K. Korpela, A. Salonen, L. J. Virta, R. A. Kekkonen, K. Forslund, P. Bork, W. M. de Vos, Intestinal microbiome is related to lifetime antibiotic use in Finnish pre-school children, Nat Commun 7 (2016) 10410.

[104] V. L. Peterson, N. J. Jury, R. Cabrera-Rubio, L. A. Draper, F. Crispie, P. D. Cotter, T. G. Dinan, A. Holmes, J. F. Cryan, Drunk bugs: Chronic vapour alcohol exposure induces marked changes in the gut microbiome in mice, Behav Brain Res 323 (2017) 172-176.

[105] S. C. Campbell, P. J. Wisniewski, 2nd, Exercise is a Novel Promoter of Intestinal Health and Microbial Diversity, Exerc Sport Sci Rev 45 (1) (2017) 41-47.

[106] B. Cerda, M. Perez, J. D. Perez-Santiago, J. F. Tornero-Aguilera, R. Gonzalez-Soltero, M. Larrosa, Gut Microbiota Modification: Another Piece in the Puzzle of the Benefits of Physical Exercise in Health?, Front Physiol 7 (2016) 51.

[107] O. Cronin, M. G. Molloy, F. Shanahan, Exercise, fitness, and the gut, Curr Opin Gastroenterol 32 (2) (2016) 67-73.

[108] A. Mika, M. Fleshner, Early-life exercise may promote lasting brain and metabolic health through gut bacterial metabolites, Immunol Cell Biol 94 (2) (2016) 151-7.

[109] O. O'Sullivan, O. Cronin, S. F. Clarke, E. F. Murphy, M. G. Molloy, F. Shanahan, P. D. Cotter, Exercise and the microbiota, Gut Microbes 6 (2) (2015) 131-6.

[110] J. A. Paulsen, T. S. Ptacek, S. J. Carter, N. Liu, R. Kumar, L. Hyndman, E. J. Lefkowitz, C. D. Morrow, L. Q. Rogers, Gut microbiota composition associated with alterations in cardiorespiratory fitness and psychosocial outcomes among breast cancer survivors, Support Care Cancer 25 (5) (2017) 1563-1570.

[111] T. Yatsunenko, F. E. Rey, M. J. Manary, I. Trehan, M. G. Dominguez-Bello, M. Contreras, M. Magris, G. Hidalgo, R. N. Baldassano, A. P. Anokhin, A. C. Heath, B. Warner, J. Reeder, J. Kuczynski, J. G. Caporaso, C. A. Lozupone, C. Lauber, J. C. Clemente, D. Knights, R. Knight, J. I. Gordon, Human gut microbiome viewed across age and geography, Nature 486 (7402) (2012) 222-7.

[112] A. Spor, O. Koren, R. Ley, Unravelling the effects of the environment and host genotype on the gut microbiome, Nat Rev Microbiol 9 (4) (2011) 279-90.

[113] M. J. Claesson, S. Cusack, O. O'Sullivan, R. Greene-Diniz, H. de Weerd, E. Flannery, J. R. Marchesi, D. Falush, T. Dinan, G. Fitzgerald, C. Stanton, D. van Sinderen, M. O'Connor, N. Harnedy, K. O'Connor, C. Henry, D. O'Mahony, A. P. Fitzgerald, F. Shanahan, C. Twomey, C. Hill, R. P. Ross, P. W. O'Toole, Composition, variability, and temporal stability of the intestinal microbiota of the elderly, Proc Natl Acad Sci U S A 108 Suppl 1 (2011) 4586-91.

[114] N. Salazar, S. Arboleya, L. Valdes, C. Stanton, P. Ross, L. Ruiz, M. Gueimonde, C. G. de Los Reyes-Gavilan, The human intestinal microbiome at extreme ages of life. Dietary intervention as a way to counteract alterations, Front Genet 5 (2014) 406.

[115] G. T. Javan, S. J. Finley, Z. Abidin, J. G. Mulle, The Thanatomicrobiome: A Missing Piece of the Microbial Puzzle of Death, Front Microbiol 7 (2016) 225.

[116] G. T. Javan, S. J. Finley, I. Can, J. E. Wilkinson, J. D. Hanson, A. M. Tarone, Human Thanatomicrobiome Succession and Time Since Death, Sci Rep 6 (2016) 29598.

[117] C. M. Lloyd, B. J. Marsland, Lung Homeostasis: Influence of Age, Microbes, and the Immune System, Immunity 46 (4) (2017) 549-561.

[118] J. C. Madan, D. C. Koestler, B. A. Stanton, L. Davidson, L. A. Moulton, M. L. Housman, J. H. Moore, M. F. Guill, H. G. Morrison, M. L. Sogin, T. H. Hampton, M. R. Karagas, P. E. Palumbo, J. A. Foster, P. L. Hibberd, G. A. O'Toole, Serial analysis of the gut and respiratory microbiome in cystic fibrosis in infancy: interaction between intestinal and respiratory tracts and impact of nutritional exposures, MBio 3 (4) (2012)

[119] L. Millares, G. Bermudo, V. Perez-Brocal, C. Domingo, M. Garcia-Nunez, X. Pomares, A. Moya, E. Monso, The respiratory microbiome in bronchial mucosa and secretions from severe IgE-mediated asthma patients, BMC Microbiol 17 (1) (2017) 20.

[120] A. Morris, J. M. Beck, P. D. Schloss, T. B. Campbell, K. Crothers, J. L. Curtis, S. C. Flores, A. P. Fontenot, E. Ghedin, L. Huang, K. Jablonski, E. Kleerup, S. V. Lynch, E. Sodergren, H. Twigg, V. B. Young, C. M. Bassis, A. Venkataraman, T. M. Schmidt, G. M. Weinstock, H. I. V. M. P. Lung, Comparison of the respiratory microbiome in healthy nonsmokers and smokers, Am J Respir Crit Care Med 187 (10) (2013) 1067-75.

[121] M. T. Wilson, D. L. Hamilos, The nasal and sinus microbiome in health and disease, Curr Allergy Asthma Rep 14 (12) (2014) 485.

[122] C. M. Liu, L. B. Price, B. A. Hungate, A. G. Abraham, L. A. Larsen, K. Christensen, M. Stegger, R. Skov, P. S. Andersen, Staphylococcus aureus and the ecology of the nasal microbiome, Sci Adv 1 (5) (2015) e1400216.

[123] J. Wu, B. A. Peters, C. Dominianni, Y. Zhang, Z. Pei, L. Yang, Y. Ma, M. P. Purdue, E. J. Jacobs, S. M. Gapstur, H. Li, A. V. Alekseyenko, R. B. Hayes, J. Ahn, Cigarette smoking and the oral microbiome in a large study of American adults, ISME J 10 (10) (2016) 2435-46.

[124] R. Dybboe, J. Bandier, L. Skov, L. Engstrand, J. D. Johansen, The Role of the Skin Microbiome in Atopic Dermatitis: A Systematic Review, Br J Dermatol (2017)

[125] E. A. Grice, H. H. Kong, S. Conlan, C. B. Deming, J. Davis, A. C. Young, N. C. S. Program, G. G. Bouffard, R. W. Blakesley, P. R. Murray, E. D. Green, M. L. Turner, J. A. Segre, Topographical and temporal diversity of the human skin microbiome, Science 324 (5931) (2009) 1190-2.

[126] H. H. Kong, J. Oh, C. Deming, S. Conlan, E. A. Grice, M. A. Beatson, E. Nomicos, E. C. Polley, H. D. Komarow, N. C. S. Program, P. R. Murray, M. L. Turner, J. A. Segre, Temporal shifts in the skin microbiome associated with disease flares and treatment in children with atopic dermatitis, Genome Res 22 (5) (2012) 850-9.

[127] J. F. Meadow, A. C. Bateman, K. M. Herkert, T. K. O'Connor, J. L. Green, Significant changes in the skin microbiome mediated by the sport of roller derby, PeerJ 1 (2013) e53.

[128] Y. Wang, L. Zhang, J. Yu, S. Huang, Z. Wang, K. A. Chun, T. L. Lee, Y. T. Chen, R. L. Gallo, C. M. Huang, A Co-Drug of Butyric Acid Derived from Fermentation Metabolites of the Human Skin Microbiome Stimulates Adipogenic Differentiation of Adipose-Derived Stem Cells: Implications in Tissue Augmentation, J Invest Dermatol 137 (1) (2017) 46-56.

[129] N. R. Klatt, R. Cheu, K. Birse, A. S. Zevin, M. Perner, L. Noel-Romas, A. Grobler, G. Westmacott, I. Y. Xie, J. Butler, L. Mansoor, L. R. McKinnon, J. S. Passmore, Q. Abdool Karim, S. S. Abdool Karim, A. D. Burgener, Vaginal bacteria modify HIV tenofovir microbicide efficacy in African women, Science 356 (6341) (2017) 938-945.

[130] R. F. Lamont, J. D. Sobel, R. A. Akins, S. S. Hassan, T. Chaiworapongsa, J. P. Kusanovic, R. Romero, The vaginal microbiome: new information about genital tract flora using molecular based techniques, BJOG 118 (5) (2011) 533-49.

[131] J. Ravel, P. Gajer, Z. Abdo, G. M. Schneider, S. S. Koenig, S. L. McCulle, S. Karlebach, R. Gorle, J. Russell, C. O. Tacket, R. M. Brotman, C. C. Davis, K. Ault, L. Peralta, L. J. Forney, Vaginal microbiome of reproductive-age women, Proc Natl Acad Sci U S A 108 Suppl 1 (2011) 4680-7.

[132] S. Ying, D. N. Zeng, L. Chi, Y. Tan, C. Galzote, C. Cardona, S. Lax, J. Gilbert, Z. X. Quan, The Influence of Age and Gender on Skin-Associated Microbial Communities in Urban and Rural Human Populations, PLoS One 10 (10) (2015) e0141842.

[133] H. H. Kong, J. A. Segre, Skin microbiome: looking back to move forward, J Invest Dermatol 132 (3 Pt 2) (2012) 933-9.

[134] A. A. Ross, A. C. Doxey, J. D. Neufeld, T. Sharpton, The Skin Microbiome of Cohabiting Couples, mSystems 2 (4) (2017) e00043-17.

[135] N. Fierer, M. Hamady, C. L. Lauber, R. Knight, The influence of sex, handedness, and washing on the diversity of hand surface bacteria, Proc Natl Acad Sci U S A 105 (46) (2008) 17994-9.

[136] G. I. Perez Perez, Z. Gao, R. Jourdain, J. Ramirez, F. Gany, C. Clavaud, J. Demaude, L. Breton, M. J. Blaser, Body Site Is a More Determinant Factor than Human Population Diversity in the Healthy Skin Microbiome, PLoS One 11 (4) (2016) e0151990.

[137] D. Rothschild, O. Weissbrod, E. Barkan, A. Kurilshikov, T. Korem, D. Zeevi, P. I. Costea, A. Godneva, I. N. Kalka, N. Bar, S. Shilo, D. Lador, A. V. Vila, N. Zmora, M. Pevsner-Fischer, D. Israeli, N. Kosower, G. Malka, B. C. Wolf, T. Avnit-Sagi, M. Lotan-Pompan, A. Weinberger, Z. Halpern, S. Carmi, J. Fu, C. Wijmenga, A. Zhernakova, E. Elinav, E. Segal, Environment dominates over host genetics in shaping human gut microbiota, Nature 555 (7695) (2018) 210-215.

[138] E. A. Grice, J. A. Segre, The skin microbiome, Nat Rev Microbiol 9 (4) (2011) 244-53.

[139] J. Urban, D. J. Fergus, A. M. Savage, M. Ehlers, H. L. Menninger, R. R. Dunn, J. E. Horvath, The effect of habitual and experimental antiperspirant and deodorant product use on the armpit microbiome, PeerJ 4 (2016) e1605.

[140] J. Oh, A. L. Byrd, M. Park, N. C. S. Program, H. H. Kong, J. A. Segre, Temporal Stability of the Human Skin Microbiome, Cell 165 (4) (2016) 854-66.

[141] Y. Wang, A. Dai, S. Huang, S. Kuo, M. Shu, C. P. Tapia, J. Yu, A. Two, H. Zhang, R. L. Gallo, C. M. Huang, Propionic acid and its esterified derivative suppress the growth of methicillin-resistant Staphylococcus aureus USA300, Benef Microbes 5 (2) (2014) 161-8.

[142] A. SanMiguel, E. A. Grice, Interactions between host factors and the skin microbiome, Cell Mol Life Sci 72 (8) (2015) 1499-515.

[143] K. Findley, J. Oh, J. Yang, S. Conlan, C. Deming, J. A. Meyer, D. Schoenfeld, E. Nomicos, M. Park, N. I. H. I. S. C. C. S. Program, H. H. Kong, J. A. Segre, Topographic diversity of fungal and bacterial communities in human skin, Nature 498 (7454) (2013) 367-70.

[144] E. M. Gribbon, W. J. Cunliffe, K. T. Holland, Interaction of Propionibacterium acnes with skin lipids in vitro, J Gen Microbiol 139 (8) (1993) 1745-51.

[145] W. Francuzik, K. Franke, R. R. Schumann, G. Heine, M. Worm, Propionibacterium acnes Abundance Correlates Inversely with Staphylococcus aureus: Data from Atopic Dermatitis Skin Microbiome, Acta Derm Venereol (2018)

[146] I. Struzycka, The Oral Microbiome in Dental Caries, Polish Journal of Microbiology 63 (2) (2014) 127-135.

[147] K. Krishnan, T. Chen, B. J. Paster, A practical guide to the oral microbiome and its relation to health and disease, Oral Dis 23 (3) (2017) 276-286.

[148] M. Kilian, I. L. Chapple, M. Hannig, P. D. Marsh, V. Meuric, A. M. Pedersen, M. S. Tonetti, W. G. Wade, E. Zaura, The oral microbiome - an update for oral healthcare professionals, Br Dent J 221 (10) (2016) 657-666.

[149] B. Liu, L. L. Faller, N. Klitgord, V. Mazumdar, M. Ghodsi, D. D. Sommer, T. R. Gibbons, T. J. Treangen, Y. C. Chang, S. Li, O. C. Stine, H. Hasturk, S. Kasif, D. Segre, M. Pop, S. Amar, Deep sequencing of the oral microbiome reveals signatures of periodontal disease, PLoS One 7 (6) (2012) e37919.

[150] A. Edlund, N. Garg, H. Mohimani, A. Gurevich, X. He, W. Shi, P. C. Dorrestein, J. S. McLean, N. Segata, Metabolic Fingerprints from the Human Oral Microbiome Reveal a Vast Knowledge Gap of Secreted Small Peptidic Molecules, mSystems 2 (4) (2017) e00058-17.

[151] A. Gomez, K. E. Nelson, The Oral Microbiome of Children: Development, Disease, and Implications Beyond Oral Health, Microb Ecol 73 (2) (2017) 492-503.

[152] C. J. Walsh, C. M. Guinane, O. T. PW, P. D. Cotter, A Profile Hidden Markov Model to investigate the distribution and frequency of LanB-encoding lantibiotic modification genes in the human oral and gut microbiome, PeerJ 5 (2017) e3254.

[153] W. van 't Hof, E. C. Veerman, A. V. Nieuw Amerongen, A. J. Ligtenberg, Antimicrobial defense systems in saliva, Monogr Oral Sci 24 (2014) 40-51.

[154] N. Takahashi, Oral Microbiome Metabolism: From "Who Are They?" to "What Are They Doing?", J Dent Res 94 (12) (2015) 1628-37.

[155] E. S. Charlson, K. Bittinger, A. R. Haas, A. S. Fitzgerald, I. Frank, A. Yadav, F. D. Bushman, R. G. Collman, Topographical continuity of bacterial populations in the healthy human respiratory tract, Am J Respir Crit Care Med 184 (8) (2011) 957-63.

[156] G. Yu, M. H. Gail, D. Consonni, M. Carugno, M. Humphrys, A. C. Pesatori, N. E. Caporaso, J. J. Goedert, J. Ravel, M. T. Landi, Characterizing human lung tissue microbiota and its relationship to epidemiological and clinical features, Genome Biol 17 (1) (2016) 163.

[157] R. P. Dickson, G. B. Huffnagle, The Lung Microbiome: New Principles for Respiratory Bacteriology in Health and Disease, PLoS Pathog 11 (7) (2015) e1004923.

[158] R. P. Dickson, J. R. Erb-Downward, C. M. Freeman, L. McCloskey, N. R. Falkowski, G. B. Huffnagle, J. L. Curtis, Bacterial Topography of the Healthy Human Lower Respiratory Tract, MBio 8 (1) (2017)

[159] C. A. Lozupone, J. I. Stombaugh, J. I. Gordon, J. K. Jansson, R. Knight, Diversity, stability and resilience of the human gut microbiota, Nature 489 (7415) (2012) 220-30.

[160] A. Legatzki, B. Rosler, E. von Mutius, Microbiome diversity and asthma and allergy risk, Curr Allergy Asthma Rep 14 (10) (2014) 466.

[161] V. R. Ramakrishnan, L. J. Hauser, L. M. Feazel, D. Ir, C. E. Robertson, D. N. Frank, Sinus microbiota varies among chronic rhinosinusitis phenotypes and predicts surgical outcome, J Allergy Clin Immunol 136 (2) (2015) 334-42 e1.

[162] E. Goleva, L. P. Jackson, J. K. Harris, C. E. Robertson, E. R. Sutherland, C. F. Hall, J. T. Good, Jr., E. W. Gelfand, R. J. Martin, D. Y. Leung, The effects of airway microbiome on corticosteroid responsiveness in asthma, Am J Respir Crit Care Med 188 (10) (2013) 1193-201.

[163] M. J. Ege, M. Mayer, A. C. Normand, J. Genuneit, W. O. Cookson, C. Braun-Fahrlander, D. Heederik, R. Piarroux, E. von Mutius, G. T. S. Group, Exposure to environmental microorganisms and childhood asthma, N Engl J Med 364 (8) (2011) 701-9.

[164] M. M. Stein, C. L. Hrusch, J. Gozdz, C. Igartua, V. Pivniouk, S. E. Murray, J. G. Ledford, M. Marques Dos Santos, R. L. Anderson, N. Metwali, J. W. Neilson, R. M. Maier, J. A. Gilbert, M. Holbreich, P. S. Thorne, F. D. Martinez, E. von Mutius, D. Vercelli, C. Ober, A. I. Sperling, Innate Immunity and Asthma Risk in Amish and Hutterite Farm Children, N Engl J Med 375 (5) (2016) 411-421.

[165] M. Hilty, C. Burke, H. Pedro, P. Cardenas, A. Bush, C. Bossley, J. Davies, A. Ervine, L. Poulter, L. Pachter, M. F. Moffatt, W. O. Cookson, Disordered microbial communities in asthmatic airways, PLoS One 5 (1) (2010) e8578.

[166] J. Durack, S. V. Lynch, S. Nariya, N. R. Bhakta, A. Beigelman, M. Castro, A. M. Dyer, E. Israel, M. Kraft, R. J. Martin, D. T. Mauger, S. R. Rosenberg, T. Sharp-King, S. R. White, P. G. Woodruff, P. C. Avila, L. C. Denlinger, F. Holguin, S. C. Lazarus, N. Lugogo, W. C. Moore, S. P. Peters, L. Que, L. J. Smith, C. A. Sorkness, M. E. Wechsler, S. E. Wenzel, H. A. Boushey, Y. J. Huang, L. National Heart, A. Blood Institute's, Features of the bronchial bacterial microbiome associated with atopy, asthma, and responsiveness to inhaled corticosteroid treatment, J Allergy Clin Immunol 140 (1) (2017) 63-75.

[167] H. C. van Woerden, C. Gregory, R. Brown, J. R. Marchesi, B. Hoogendoorn, I. P. Matthews, Differences in fungi present in induced sputum samples from asthma patients and non-atopic controls: a community based case control study, BMC Infect Dis 13 (2013) 69.

[168] D. N. O'Dwyer, R. P. Dickson, B. B. Moore, The Lung Microbiome, Immunity, and the Pathogenesis of Chronic Lung Disease, J Immunol 196 (12) (2016) 4839-47.

[169] M. T. Salam, H. G. Margolis, R. McConnell, J. A. McGregor, E. L. Avol, F. D. Gilliland, Mode of delivery is associated with asthma and allergy occurrences in children, Ann Epidemiol 16 (5) (2006) 341-6.

[170] J. Kero, M. Gissler, M. M. Gronlund, P. Kero, P. Koskinen, E. Hemminki, E. Isolauri, Mode of delivery and asthma -- is there a connection?, Pediatr Res 52 (1) (2002) 6-11.

[171] B. H. K. Yip, H. Leonard, S. Stock, C. Stoltenberg, R. W. Francis, M. Gissler, R. Gross, D. Schendel, S. Sandin, Caesarean section and risk of autism across gestational age: a multi-national cohort study of 5 million births, Int J Epidemiol 46 (2) (2017) 429-439.

[172] S. A. Whiteside, H. Razvi, S. Dave, G. Reid, J. P. Burton, The microbiome of the urinary tract--a role beyond infection, Nat Rev Urol 12 (2) (2015) 81-90.

[173] S. L. Hillier, R. P. Nugent, D. A. Eschenbach, M. A. Krohn, R. S. Gibbs, D. H. Martin, M. F. Cotch, R. Edelman, J. G. Pastorek, 2nd, A. V. Rao, et al., Association between bacterial vaginosis and preterm delivery of a low-birth-weight infant. The Vaginal Infections and Prematurity Study Group, N Engl J Med 333 (26) (1995) 1737-42.

[174] A. Mitra, D. A. MacIntyre, Y. S. Lee, A. Smith, J. R. Marchesi, B. Lehne, R. Bhatia, D. Lyons, E. Paraskevaidis, J. V. Li, E. Holmes, J. K. Nicholson, P. R. Bennett, M. Kyrgiou, Cervical intraepithelial neoplasia disease progression is associated with increased vaginal microbiome diversity, Sci Rep 5 (2015) 16865.

[175] R. B. Ness, K. E. Kip, S. L. Hillier, D. E. Soper, C. A. Stamm, R. L. Sweet, P. Rice, H. E. Richter, A cluster analysis of bacterial vaginosis-associated microflora and pelvic inflammatory disease, Am J Epidemiol 162 (6) (2005) 585-90.

[176] N. Sewankambo, R. H. Gray, M. J. Wawer, L. Paxton, D. McNairn, F. Wabwire-Mangen, D. Serwadda, C. Li, N. Kiwanuka, S. L. Hillier, L. Rabe, C. A. Gaydos, T. C. Quinn, J. Konde-Lule, HIV-1 infection associated with abnormal vaginal flora morphology and bacterial vaginosis, The Lancet 350 (9077) (1997) 546-550.

[177] M. G. Gravett, Independent Associations of Bacterial Vaginosis and Chlamydia trachomatis Infection With Adverse Pregnancy Outcome, JAMA: The Journal of the American Medical Association 256 (14) (1986) 1899.

[178] B. Ma, L. J. Forney, J. Ravel, Vaginal microbiome: rethinking health and disease, Annu Rev Microbiol 66 (2012) 371-89.

[179] E. A. Miller, D. E. Beasley, R. R. Dunn, E. A. Archie, Lactobacilli Dominance and Vaginal pH: Why Is the Human Vaginal Microbiome Unique?, Front Microbiol 7 (2016) 1936.

[180] S. Boris, J. E. Suarez, F. Vazquez, C. Barbes, Adherence of human vaginal lactobacilli to vaginal epithelial cells and interaction with uropathogens, Infect Immun 66 (5) (1998) 1985-9.

[181] B. J. Moncla, C. A. Chappell, B. M. Debo, L. A. Meyn, The Effects of Hormones and Vaginal Microflora on the Glycome of the Female Genital Tract: Cervical-Vaginal Fluid, PLoS One 11 (7) (2016) e0158687.

[182] G. I. Gorodeski, U. Hopfer, C. C. Liu, E. Margles, Estrogen acidifies vaginal pH by up-regulation of proton secretion via the apical membrane of vaginal-ectocervical epithelial cells, Endocrinology 146 (2) (2005) 816-24.

[183] E. R. Boskey, K. M. Telsch, K. J. Whaley, T. R. Moench, R. A. Cone, Acid production by vaginal flora in vitro is consistent with the rate and extent of vaginal acidification, Infection and Immunity 67 (10) (1999) 5170-5175.

[184] G. I. Gorodeski, Effects of estrogen on proton secretion via the apical membrane in vaginal-ectocervical epithelial cells of postmenopausal women, Menopause 12 (6) (2005) 679-84.

[185] M. Aldunate, D. Srbinovski, A. C. Hearps, C. F. Latham, P. A. Ramsland, R. Gugasyan, R. A. Cone, G. Tachedjian, Antimicrobial and immune modulatory effects of lactic acid and short chain fatty acids produced by vaginal microbiota associated with eubiosis and bacterial vaginosis, Front Physiol 6 (2015) 164.

[186] I. M. Linhares, P. R. Summers, B. Larsen, P. C. Giraldo, S. S. Witkin, Contemporary perspectives on vaginal pH and lactobacilli, Am J Obstet Gynecol 204 (2) (2011) 120 e1-5.

[187] D. E. O'Hanlon, T. R. Moench, R. A. Cone, Vaginal pH and microbicidal lactic acid when lactobacilli dominate the microbiota, PLoS One 8 (11) (2013) e80074.

[188] P. Mirmonsef, M. R. Zariffard, D. Gilbert, H. Makinde, A. L. Landay, G. T. Spear, Short-chain fatty acids induce pro-inflammatory cytokine production alone and in combination with toll-like receptor ligands, Am J Reprod Immunol 67 (5) (2012) 391-400.

[189] L. B. Price, C. M. Liu, K. E. Johnson, M. Aziz, M. K. Lau, J. Bowers, J. Ravel, P. S. Keim, D. Serwadda, M. J. Wawer, R. H. Gray, The effects of circumcision on the penis microbiome, PLoS One 5 (1) (2010) e8422.

[190] C. M. Liu, B. A. Hungate, A. A. Tobian, D. Serwadda, J. Ravel, R. Lester, G. Kigozi, M. Aziz, R. M. Galiwango, F. Nalugoda, T. L. Contente-Cuomo, M. J. Wawer, P. Keim, R. H. Gray, L. B. Price, Male circumcision significantly reduces prevalence and load of genital anaerobic bacteria, MBio 4 (2) (2013) e00076.

[191] J. L. Prodger, R. Kaul, The biology of how circumcision reduces HIV susceptibility: broader implications for the prevention field, AIDS Res Ther 14 (1) (2017) 49.

[192] K. Yasuda, K. Oh, B. Ren, T. L. Tickle, E. A. Franzosa, L. M. Wachtman, A. D. Miller, S. V. Westmoreland, K. G. Mansfield, E. J. Vallender, G. M. Miller, J. K. Rowlett, D. Gevers, C. Huttenhower, X. C. Morgan, Biogeography of the intestinal mucosal and lumenal microbiome in the rhesus macaque, Cell Host Microbe 17 (3) (2015) 385-91.

[193] C. Tropini, K. A. Earle, K. C. Huang, J. L. Sonnenburg, The Gut Microbiome: Connecting Spatial Organization to Function, Cell Host Microbe 21 (4) (2017) 433-442.

[194] E. M. Bik, P. B. Eckburg, S. R. Gill, K. E. Nelson, E. A. Purdom, F. Francois, G. Perez-Perez, M. J. Blaser, D. A. Relman, Molecular analysis of the bacterial microbiota in the human stomach, Proc Natl Acad Sci U S A 103 (3) (2006) 732-7.

[195] J. M. Lang, J. A. Eisen, A. M. Zivkovic, The microbes we eat: abundance and taxonomy of microbes consumed in a day's worth of meals for three diet types, PeerJ 2 (2014) e659.

[196] J. L. Sonnenburg, F. Backhed, Diet-microbiota interactions as moderators of human metabolism, Nature 535 (7610) (2016) 56-64.

[197] E. C. Martens, A. G. Kelly, A. S. Tauzin, H. Brumer, The devil lies in the details: how variations in polysaccharide fine-structure impact the physiology and evolution of gut microbes, J Mol Biol 426 (23) (2014) 3851-65.

[198] B. V. Jones, M. Begley, C. Hill, C. G. Gahan, J. R. Marchesi, Functional and comparative metagenomic analysis of bile salt hydrolase activity in the human gut microbiome, Proc Natl Acad Sci U S A 105 (36) (2008) 13580-5.

[199] S. A. Joyce, J. MacSharry, P. G. Casey, M. Kinsella, E. F. Murphy, F. Shanahan, C. Hill, C. G. Gahan, Regulation of host weight gain and lipid metabolism by bacterial bile acid modification in the gut, Proc Natl Acad Sci U S A 111 (20) (2014) 7421-6.

[200] J. K. Nicholson, E. Holmes, J. Kinross, R. Burcelin, G. Gibson, W. Jia, S. Pettersson, Host-gut microbiota metabolic interactions, Science 336 (6086) (2012) 1262-7.

[201] M. W. van Passel, R. Kant, E. G. Zoetendal, C. M. Plugge, M. Derrien, S. A. Malfatti, P. S. Chain, T. Woyke, A. Palva, W. M. de Vos, H. Smidt, The genome of Akkermansia muciniphila, a dedicated intestinal mucin degrader, and its use in exploring intestinal metagenomes, PLoS One 6 (3) (2011) e16876.

[202] A. Vrieze, C. Out, S. Fuentes, L. Jonker, I. Reuling, R. S. Kootte, E. van Nood, F. Holleman, M. Knaapen, J. A. Romijn, M. R. Soeters, E. E. Blaak, G. M. Dallinga-Thie, D. Reijnders, M. T. Ackermans, M. J. Serlie, F. K. Knop, J. J. Holst, C. van der Ley, I. P. Kema, E. G. Zoetendal, W. M. de Vos, J. B. Hoekstra, E. S. Stroes, A. K. Groen, M. Nieuwdorp, Impact of oral vancomycin on gut microbiota, bile acid metabolism, and insulin sensitivity, J Hepatol 60 (4) (2014) 824-31.

[203] E. F. Enright, C. G. Gahan, S. A. Joyce, B. T. Griffin, The Impact of the Gut Microbiota on Drug Metabolism and Clinical Outcome, Yale J Biol Med 89 (3) (2016) 375-382.

[204] H. Wu, E. Esteve, V. Tremaroli, M. T. Khan, R. Caesar, L. Manneras-Holm, M. Stahlman, L. M. Olsson, M. Serino, M. Planas-Felix, G. Xifra, J. M. Mercader, D. Torrents, R. Burcelin, W. Ricart, R. Perkins, J. M. Fernandez-Real, F. Backhed, Metformin alters the gut microbiome of individuals with treatment-naive type 2 diabetes, contributing to the therapeutic effects of the drug, Nat Med 23 (7) (2017) 850-858.

[205] L. Maier, M. Pruteanu, M. Kuhn, G. Zeller, A. Telzerow, E. E. Anderson, A. R. Brochado, K. C. Fernandez, H. Dose, H. Mori, K. R. Patil, P. Bork, A. Typas, Extensive impact of non-antibiotic drugs on human gut bacteria, Nature 555 (7698) (2018) 623-628.

[206] N. L. Cluny, C. M. Keenan, R. A. Reimer, B. Le Foll, K. A. Sharkey, Prevention of Diet-Induced Obesity Effects on Body Weight and Gut Microbiota in Mice Treated Chronically with Delta9-Tetrahydrocannabinol, PLoS One 10 (12) (2015) e0144270.

[207] E. A. Mutlu, P. M. Gillevet, H. Rangwala, M. Sikaroodi, A. Naqvi, P. A. Engen, M. Kwasny, C. K. Lau, A. Keshavarzian, Colonic microbiome is altered in alcoholism, Am J Physiol Gastrointest Liver Physiol 302 (9) (2012) G966-78.

[208] J. Panee, M. Gerschenson, L. Chang, Associations Between Microbiota, Mitochondrial Function, and Cognition in Chronic Marijuana Users, J Neuroimmune Pharmacol 13 (1) (2018) 113-122.

[209] J. A. Fulcher, S. K. Hussain, R. Cook, F. Li, N. H. Tobin, A. Ragsdale, S. Shoptaw, P. M. Gorbach, G. M. Aldrovandi, Effects of Substance Use and Sex Practices on the Intestinal Microbiome During HIV-1 Infection, J Infect Dis 218 (10) (2018) 1560-1570.

[210] H. M. Hamer, D. M. Jonkers, A. Bast, S. A. Vanhoutvin, M. A. Fischer, A. Kodde, F. J. Troost, K. Venema, R. J. Brummer, Butyrate modulates oxidative stress in the colonic mucosa of healthy humans, Clin Nutr 28 (1) (2009) 88-93.

[211] E. Holmes, R. L. Loo, J. Stamler, M. Bictash, I. K. Yap, Q. Chan, T. Ebbels, M. De Iorio, I. J. Brown, K. A. Veselkov, M. L. Daviglus, H. Kesteloot, H. Ueshima, L. Zhao, J. K. Nicholson, P. Elliott, Human metabolic phenotype diversity and its association with diet and blood pressure, Nature 453 (7193) (2008) 396-400.

[212] E. G. Zoetendal, A. von Wright, T. Vilpponen-Salmela, K. Ben-Amor, A. D. L. Akkermans, W. M. de Vos, Mucosa-Associated Bacteria in the Human Gastrointestinal Tract Are Uniformly Distributed along the Colon and Differ from the Community Recovered from Feces, Applied and Environmental Microbiology 68 (7) (2002) 3401-3407.

[213] C. Chun, L. Zheng, S. P. Colgan, Tissue metabolism and host-microbial interactions in the intestinal mucosa, Free Radic Biol Med 105 (2017) 86-92.

[214] P. J. Turnbaugh, R. E. Ley, M. A. Mahowald, V. Magrini, E. R. Mardis, J. I. Gordon, An obesity-associated gut microbiome with increased capacity for energy harvest, Nature 444 (7122) (2006) 1027-31.

[215] C. T. Brown, A. G. Davis-Richardson, A. Giongo, K. A. Gano, D. B. Crabb, N. Mukherjee, G. Casella, J. C. Drew, J. Ilonen, M. Knip, H. Hyoty, R. Veijola, T. Simell, O. Simell, J. Neu, C. H. Wasserfall, D. Schatz, M. A. Atkinson, E. W. Triplett, Gut microbiome metagenomics analysis suggests a functional model for the development of autoimmunity for type 1 diabetes, PLoS One 6 (10) (2011) e25792.

[216] A. V. Hartstra, K. E. Bouter, F. Backhed, M. Nieuwdorp, Insights into the role of the microbiome in obesity and type 2 diabetes, Diabetes Care 38 (1) (2015) 159-65.

[217] J. Halfvarson, C. J. Brislawn, R. Lamendella, Y. Vazquez-Baeza, W. A. Walters, L. M. Bramer, M. D'Amato, F. Bonfiglio, D. McDonald, A. Gonzalez, E. E. McClure, M. F. Dunklebarger, R. Knight, J. K. Jansson, Dynamics of the human gut microbiome in inflammatory bowel disease, Nat Microbiol 2 (2017) 17004.

[218] X. C. Morgan, T. L. Tickle, H. Sokol, D. Gevers, K. L. Devaney, D. V. Ward, J. A. Reyes, S. A. Shah, N. LeLeiko, S. B. Snapper, A. Bousvaros, J. Korzenik, B. E. Sands, R. J. Xavier, C. Huttenhower, Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment, Genome Biol 13 (9) (2012) R79.

[219] P. J. Kennedy, A. B. Murphy, J. F. Cryan, P. R. Ross, T. G. Dinan, C. Stanton, Microbiome in brain function and mental health, Trends in Food Science & Technology 57 (2016) 289-301.

[220] M. Messaoudi, R. Lalonde, N. Violle, H. Javelot, D. Desor, A. Nejdi, J. F. Bisson, C. Rougeot, M. Pichelin, M. Cazaubiel, J. M. Cazaubiel, Assessment of psychotropic-like properties of a probiotic formulation (Lactobacillus helveticus R0052 and Bifidobacterium longum R0175) in rats and human subjects, Br J Nutr 105 (5) (2011) 755-64.

[221] A. D. Kostic, D. Gevers, C. S. Pedamallu, M. Michaud, F. Duke, A. M. Earl, A. I. Ojesina, J. Jung, A. J. Bass, J. Tabernero, J. Baselga, C. Liu, R. A. Shivdasani, S. Ogino, B. W. Birren, C. Huttenhower, W. S. Garrett, M. Meyerson, Genomic analysis identifies association of Fusobacterium with colorectal carcinoma, Genome Res 22 (2) (2012) 292-8.

[222] W. Chen, F. Liu, Z. Ling, X. Tong, C. Xiang, Human intestinal lumen and mucosa-associated microbiota in patients with colorectal cancer, PLoS One 7 (6) (2012) e39743.

[223] S. G. Daniel, C. L. Ball, D. G. Besselsen, T. Doetschman, B. L. Hurwitz, Functional Changes in the Gut Microbiome Contribute to Transforming Growth Factor beta-Deficient Colon Cancer, mSystems 2 (5) (2017)

[224] T. S. Postler, S. Ghosh, Understanding the Holobiont: How Microbial Metabolites Affect Human Health and Shape the Immune System, Cell Metab 26 (1) (2017) 110-130.

[225] P. Louis, H. J. Flint, Formation of propionate and butyrate by the human colonic microbiota, Environ Microbiol 19 (1) (2017) 29-41.

[226] E. Le Poul, C. Loison, S. Struyf, J. Y. Springael, V. Lannoy, M. E. Decobecq, S. Brezillon, V. Dupriez, G. Vassart, J. Van Damme, M. Parmentier, M. Detheux, Functional characterization of human receptors for short chain fatty acids and their role in polymorphonuclear cell activation, J Biol Chem 278 (28) (2003) 25481-9.

[227] G. Frost, M. L. Sleeth, M. Sahuri-Arisoylu, B. Lizarbe, S. Cerdan, L. Brody, J. Anastasovska, S. Ghourab, M. Hankir, S. Zhang, D. Carling, J. R. Swann, G. Gibson, A. Viardot, D. Morrison, E. Louise Thomas, J. D. Bell, The short-chain fatty acid acetate reduces appetite via a central homeostatic mechanism, Nat Commun 5 (2014) 3611.

[228] F. De Vadder, P. Kovatcheva-Datchary, D. Goncalves, J. Vinera, C. Zitoun, A. Duchampt, F. Backhed, G. Mithieux, Microbiota-generated metabolites promote metabolic benefits via gut-brain neural circuits, Cell 156 (1-2) (2014) 84-96.

[229] P. M. Smith, M. R. Howitt, N. Panikov, M. Michaud, C. A. Gallini, Y. M. Bohlooly, J. N. Glickman, W. S. Garrett, The microbial metabolites, short-chain fatty acids, regulate colonic Treg cell homeostasis, Science 341 (6145) (2013) 569-73.

[230] T. P. Bui, J. Ritari, S. Boeren, P. de Waard, C. M. Plugge, W. M. de Vos, Production of butyrate from lysine and the Amadori product fructoselysine by a human gut commensal, Nat Commun 6 (2015) 10062.

[231] G. Macfarlane, Estimation of short-chain fatty acid production from protein by human intestinal bacteria based on branched-chain fatty acid measurements, FEMS Microbiology Letters 101 (2) (1992) 81-88.

[232] M. Vital, A. C. Howe, J. M. Tiedje, Revealing the bacterial butyrate synthesis pathways by analyzing (meta)genomic data, MBio 5 (2) (2014) e00889.

[233] P. Louis, S. H. Duncan, S. I. McCrae, J. Millar, M. S. Jackson, H. J. Flint, Restricted Distribution of the Butyrate Kinase Pathway among Butyrate-Producing Bacteria from the Human Colon, Journal of Bacteriology 186 (7) (2004) 2099-2106.

[234] M. Thangaraju, G. Cresci, S. Itagaki, J. Mellinger, D. D. Browning, F. G. Berger, P. D. Prasad, V. Ganapathy, Sodium-coupled transport of the short chain fatty acid butyrate by SLC5A8 and its relevance to colon cancer, J Gastrointest Surg 12 (10) (2008) 1773-81; discussion 1781-2.

[235] D. R. Donohoe, L. B. Collins, A. Wali, R. Bigler, W. Sun, S. J. Bultman, The Warburg effect dictates the mechanism of butyrate-mediated histone acetylation and cell proliferation, Mol Cell 48 (4) (2012) 612-26.

[236] A. Trompette, E. S. Gollwitzer, K. Yadava, A. K. Sichelstiel, N. Sprenger, C. Ngom-Bru, C. Blanchard, T. Junt, L. P. Nicod, N. L. Harris, B. J. Marsland, Gut microbiota metabolism of dietary fiber influences allergic airway disease and hematopoiesis, Nat Med 20 (2) (2014) 159-66.

[237] A. N. Thorburn, C. I. McKenzie, S. Shen, D. Stanley, L. Macia, L. J. Mason, L. K. Roberts, C. H. Wong, R. Shim, R. Robert, N. Chevalier, J. K. Tan, E. Marino, R. J. Moore, L. Wong, M. J. McConville, D. L. Tull, L. G. Wood, V. E. Murphy, J. Mattes, P. G. Gibson, C. R. Mackay, Evidence that asthma is a developmental origin disease influenced by maternal diet and bacterial metabolites, Nat Commun 6 (2015) 7320.

[238] C. M. Lloyd, C. M. Hawrylowicz, Regulatory T cells in asthma, Immunity 31 (3) (2009) 438-49.

[239] A. Labbe, J. G. Ganopolsky, C. J. Martoni, S. Prakash, M. L. Jones, Bacterial bile metabolising gene abundance in Crohn's, ulcerative colitis and type 2 diabetes metagenomes, PLoS One 9 (12) (2014) e115175.

[240] J. M. Ridlon, D. J. Kang, P. B. Hylemon, J. S. Bajaj, Bile acids and the gut microbiome, Curr Opin Gastroenterol 30 (3) (2014) 332-8.

[241] P. Lefebvre, B. Cariou, F. Lien, F. Kuipers, B. Staels, Role of bile acids and bile acid receptors in metabolic regulation, Physiol Rev 89 (1) (2009) 147-91.

[242] J. M. Ridlon, S. C. Harris, S. Bhowmik, D. J. Kang, P. B. Hylemon, Consequences of bile salt biotransformations by intestinal bacteria, Gut Microbes 7 (1) (2016) 22-39.

[243] M. Kobayashi, H. Ikegami, T. Fujisawa, K. Nojima, Y. Kawabata, S. Noso, N. Babaya, M. Itoi-Babaya, K. Yamaji, Y. Hiromine, M. Shibata, T. Ogihara, Prevention and treatment of obesity, insulin resistance, and diabetes by bile acid-binding resin, Diabetes 56 (1) (2007) 239-47.

[244] N. C. Penney, J. Kinross, R. C. Newton, S. Purkayastha, The role of bile acids in reducing the metabolic complications of obesity after bariatric surgery: a systematic review, Int J Obes (Lond) 39 (11) (2015) 1565-74.

[245] M. E. Patti, S. M. Houten, A. C. Bianco, R. Bernier, P. R. Larsen, J. J. Holst, M. K. Badman, E. Maratos-Flier, E. C. Mun, J. Pihlajamaki, J. Auwerx, A. B. Goldfine, Serum bile acids are higher in humans with prior gastric bypass: potential contribution to improved glucose and lipid metabolism, Obesity (Silver Spring) 17 (9) (2009) 1671-7.

[246] S. M. Centuori, C. J. Gomes, J. Trujillo, J. Borg, J. Brownlee, C. W. Putnam, J. D. Martinez, Deoxycholic acid mediates non-canonical EGFR-MAPK activation through the induction of calcium signaling in colon cancer cells, Biochim Biophys Acta 1861 (7) (2016) 663-70.

[247] S. Yoshimoto, T. M. Loo, K. Atarashi, H. Kanda, S. Sato, S. Oyadomari, Y. Iwakura, K. Oshima, H. Morita, M. Hattori, K. Honda, Y. Ishikawa, E. Hara, N. Ohtani, Obesity-induced gut microbial metabolite promotes liver cancer through senescence secretome, Nature 499 (7456) (2013) 97-101.

[248] E. Bayerdorffer, G. A. Mannes, T. Ochsenkuhn, P. Dirschedl, B. Wiebecke, G. Paumgartner, Unconjugated secondary bile acids in the serum of patients with colorectal adenomas, Gut 36 (2) (1995) 268-273.

[249] J. P. Coppe, P. Y. Desprez, A. Krtolica, J. Campisi, The senescence-associated secretory phenotype: the dark side of tumor suppression, Annu Rev Pathol 5 (2010) 99-118.

Chapter 2

MATERIALS AND METHODS

The following chapter outlines the general materials and methods that were practiced in the experimental portions of this thesis. Any deviations or additional approaches have been described in corresponding chapters.

2.1 | Ethical approval

Investigations that focused on, or otherwise included human participants (chapters 3 - 8) were appropriately conducted under guidelines defined by the Helsinki Declaration. In all such studies, ethical approval was formally sanctioned by the Clinical Ethics Committee of the Cork Teaching Hospitals (CREC). All study participants provided written informed consent prior to commencement of the various studies. An adapted version of the American College of Sports Medicine's safe participation questionnaire was utilized to ensure participant safety during the course of the exercise intervention studies (chapters 6 and 7) [1].

2.2 | Study compliance and withdrawal

Exercise program compliance (chapters 6 and 7) was monitored remotely by the investigators using the FitLinxx® activity monitoring system (Shelton, Connecticut, USA). All volunteers were provided with a unique identification number for the FitLinxx® physical activity recording system and were required to login and record all activities undertaken at the Mardyke Arena gymnasium during the intervention period. Using this tracking system, compliance with the prescribed exercise program was monitored by the investigators. Similarly, the quantity of aerobic and resistance training performed by participants was

compiled. The Fitlinxx® software program enables accurate recording of the duration and frequency of training, and provides an estimate of energy expenditure during aerobic exercise. The facility's Fitlinxx® software and hardware was maintained and re-calibrated prior to commencement of the study. Volunteers noted not to be complying with the exercise regime for more than 7 consecutive days were withdrawn from participation.

2.3 | Participant data collection

2.3.1 | Diet & nutrition

In all studies, recruited volunteers were asked to maintain typical dietary habits throughout the observation periods. For chapters 3, 4, 5, 6, and 8, dietary data from participants was collected by means of a 146-item food frequency questionnaire (FFQ). The FFQ was an adapted version of the questionnaire used in the UK arm of the European Prospective Investigation into Cancer (EPIC) study [2], which was based on the original Willet FFQ [3]. Participants were asked to record their usual pattern of dietary intake over the duration of observation.

2.3.2 | Physical activity

Healthy individuals compared to athletes as controls for high-performance athleticism for studies described in chapters 3-5, were assessed for physical activity levels using an adapted version of the EPIC-Norfolk questionnaire [4]. Baseline physical activity levels of volunteers (chapters 6, 7 and 8) were characterized using the long form of the International Physical Activity Questionnaire (IPAQ) [5]. This self-reported questionnaire provided an estimate of the physical activity patterns of participants including an estimation of the time spent engaged in physical activity (walking, moderate, and vigorous activity) and the metabolic equivalents of task (MET) minutes and kilocalories expended per week by

physical activity. To prevent injury from unaccustomed vigorous exercise, a sub-maximal assessment of peak aerobic capacity was used in intervention studies (chapters 6 and 7). Baseline and post-intervention levels of cardio-respiratory fitness were measured using a validated sub-maximal fitness test [6]. The Rockport one-mile walk test was performed in a standardized temperature environment at the indoor running track of the Mardyke Arena, University College Cork, Ireland. This test was used to estimate maximal oxygen uptake (VO_{2max}).

2.3.3 | Body composition

Dual Energy X-ray Absorptiometry (DEXA) was used to assess body composition change in the volunteers from all cohorts. A GE Healthcare Lunar iDXA machine (Madison, Wisconsin, USA) at the Bone Densitometry Unit, Cork University Hospital, Ireland was used. The enCORE software (V.13.4, 2010) analysed body composition using a three-compartment body composition model (Fat mass, bone mass, lean tissue). Volunteers were scanned post-voiding and dressed in light clothing, with metal-wear removed where present. Quality control analysis was performed on the iDXA machine before use on each measurement day.

2.3.4 | Inflammatory cytokine measurement

Blood samples (4 ml) from participants were collected in serum separator clot activator blood collection tubes (Greiner Bio-One, Stonehouse, United Kingdom; reference no. 454071). The blood samples were allowed to rest upright on the laboratory bench for 30 min before centrifugation was performed at 1,000 × *g* for 20 min at room temperature. Approximately 2 ml of supernatant sera was then harvested by pipette, frozen, and stored at −80°C in polypropylene cryogenic vials. At a later date, following a complete thaw, resting levels of proinflammatory cytokines were measured using a mesoscale discovery

(MSD) platform (Meso Scale Discovery, Rockville, MD). The MSD system is an electrochem-iluminescence-based solid-phase multiplex assay. An ultrasensitive human proinflammatory I, V-Plex immunoassay panel containing interleukin-6 (IL-6), IL-8, IL-10, tumour necrosis factor alpha (TNF-α), and gamma interferon was used to measure serum cytokine levels. All such cytokines were measured in chapter 6, while gamma interferon was not measured in chapters 7 and 8. Samples were diluted 1:2 according to the manufacturer's protocol. The lower limit of detection was <1 pg/ml for all assays. All plasma samples were measured in duplicate, and the mean cytokine concentration of the duplicates (in picograms per millilitre) was used for analysis.

2.4 | Exercise intervention

Combined aerobic and resistance training (chapters 6 and 7) was performed at the Mardyke Arena gymnasium at University College Cork, Ireland, where all exercise sessions took place. Participants undergoing exercise intervention were instructed to adhere to the assigned exercise program and to avoid additional moderate to vigorous physical activity outside of that prescribed. Participants were instructed to train 3-times per week for the 8-week intervention periods. Participants were familiarized with the format of the required training program during a 90-minute induction session with designated gym instructors. This induction included demonstration of all aerobic and resistance training equipment allowing the opportunity to ask questions if required. Resistance training machines were customized for individual differences in range of motion and the participants were observed using all machines with instructor feedback and correction. For resistance machines, 1 repetition maximum (1RM) values were calculated using standardized methods [7].

The outline of the exercise sessions was as follows: After a 5-minute warm-up on the treadmill (brisk walking at approximately 4km/hour, modified Borg rating of perceived exertion (RPE) scale: 3-5/10) [8], participants underwent aerobic training of moderate intensity (modified Borg RPE: 5-7/10). To encourage compliance with the prescribed RPE scales, volunteers were reminded of the desired intensities on each of their weekly exercise training programs. The duration of aerobic exercise progressed on a weekly basis but remained of moderate intensity. Initially, aerobic exercise lasted approximately 18 minutes and by week 8 of the intervention period, aerobic exercise increased to approximately 32 minutes depending on the type of aerobic activity chosen by the volunteer. To allow variety and to maintain interest participants were provided with a choice of aerobic activities, including treadmill jogging/running, use of a cross-trainer device (with no added resistance), use of a stepper machine, and stationary cycling (with mild resistance). The duration for each of these activities was calculated based on the 2011 Compendium of Physical Activities [9] to ensure similar levels of energy expenditure across all activities. Participants were instructed not to change between aerobic exercises within a given training session.

Upon completion of aerobic activity, participants undertook machine-based resistance training. In summary, participants were required to perform a minimum of 3 sets of 8 repetitions on 7-different resistance machines (three upper body, three lower body, one core muscle exercise). The allowed resistance machine options were as follows: for the upper body, shoulder press, chest press, lateral pulldowns, and seated rowing; for the lower body, leg extension, leg curl, gluteal kick-back, and leg press; for the core muscles, abdominal curls and torso rotation. A minimum limit of 3 sets of 8 repetitions was instituted, with a maximum limit of 3 sets of 12 repetitions. Starting weights were calculated

at induction to correspond to 70% of the individual's one-repetition maximum (1RM) value. Resistance training was progressive, with the aim of increasing the resistance weight by 15% to 20% over an 8-week period. Free-weight use was not permitted.

2.5 | Extraction and sequencing of faecal microbiome DNA

2.5.1 | Processing of faecal samples

DNA was extracted from the donated fresh faecal samples received at the Teagasc Moorepark research facility using a QIAmp DNA stool minikit (Qiagen, Crawley, West Sussex, United Kingdom) [10]. Samples were provided by participants as partial evacuations into sterile containers and, when not immediately transported for DNA extraction, were held at 4°C for no more than 12 h. Samples were prepared for DNA extraction by manual homogenization of a portion of the sample representing all microenvironments (i.e., core and external surface) of the faeces. The provided manufacturer's protocol was used with modification whereby a zirconia bead (Stratech Scientific) cell disruption bead-beating step (performed three times for 30 s each time) was introduced in order to enhance homogenization of the extraction material. DNA extracts and the remaining faecal samples were subsequently stored at −80°C until they were prepared for sequencing.

2.5.2 | Metagenomic library preparation and sequencing

Metagenomic library preparation was performed with the Illumina Nextera XT DNA Library Preparation Kit (cat. no. FC-131-1096, Illumina Inc., USA) in explicit accordance with the manufacturer's protocol (15031942, Illumina). Normalisation of samples to the recommended 0.2 ng/μL per individual library was carried out with the ThermoFisher Qubit 2.0 Flurometric Quantitation system (ThermoFisher Scientific), using a combination of

broad-range and high-sensitivity assay kits as appropriate (cat. nos. Q32850 and Q32851 respectively, ThermoFisher Scientific). Tagmentation—the combined cleavage of sample DNA into fragments and application of universal overhang 'tag'—and amplification of DNA fragments was carried out with the G-STORM GS1 thermal cycler system (G-Storm Ltd. UK), using the thermal cycler programs provided by the Nextera XT protocol (15031942, Illumina). Following the combined enzymatic fragmentation and adapter sequence tagging—tagmentation—and the subsequent amplification of the tagmented DNA, libraries were purified with the AMPure magnetic bead system at a ratio of 1:1.8 (DNA:AMPure) (cat. no. 9A63880, Beckman Coulter). Subsequently, libraries were assessed for appropriate fragment size on the Agilent 2100 Bioanalyzer system (cat. no. G2939BA, Agilent Technologies), utilizing the accompanying Agilent DNA 1000 chip-assay kit (cat. no. 5067-1504, Agilent Technologies). With the libraries passing quality and fragment length requirements, the library preparation was continued on through library normalization, which was met with an additional assessment of suitable molar concentrations (~2 nM) with the KAPA Library Quantification Kit (KK4824, Kapabiosystems) run on a Roche Light-Cycler 480 (Roche Applied Science). Samples from the various studies were combined in equimolar concentration (2nM) pools for prior to sequencing.

Sequencing data used in chapters 3 – 5 were generated from the Illumina HiSeq 2500 (chemistry V.4.0) next generation sequencing (NGS) platform by Eurofins Genetic Services Ltd. (Ebersberg, Germany), while such data for chapter 6 was provided with the same instrument but by Beckman Coulter Genomics Inc. (Danvers, MA U.S.). NGS data for chapters 7 and 8 were produced at the Teagasc sequencing facility (Cork, Ireland) on the Illumina NextSeq 500 (chemistry V.2.0). Further specific details are outlined in the appropriate chapters.

2.6 | Bioinformatic processing of microbial sequencing data

Processing of metagenomic FASTQ sequence files proceeded with the removal of human-derived contaminant sequences with NCBI Best Match Tagger (BMTagger) software (Fig. 2.1), while trimming and removal of duplicate reads or of reads of substandard quality were performed with Picard and SAM tools. Functional profiling of high-quality processed reads was facilitated by use of the Human Microbiome Project (HMP) Unified Metabolic Analysis Network (HUMAnN2) pipeline [11]. Models of microbial metabolic pathways produced by HUMAnN2 were derived from the MetaCyc database [12] and were the basis for analyses performed on microbial metabolic profiling. Version 0.5.0 of HUMAnN2 was used in chapters 3-5, while a later edition, version 0.99 was used in subsequent chapters.

Taxonomic profiles were generated from the sequencing data using a combination of Metagenomic Phylogenetic Analysis (MetaPhlA2, V.2.0), Kaiju (V.1.5.0) [13] and Kraken (V.0.10.6) software packages [14, 15]. Additionally, for the comparison of sequencing approaches presented in chapter 3, the software tool GraPhlAn (V.0.9.7) was used to construct phylogenetic trees [16].

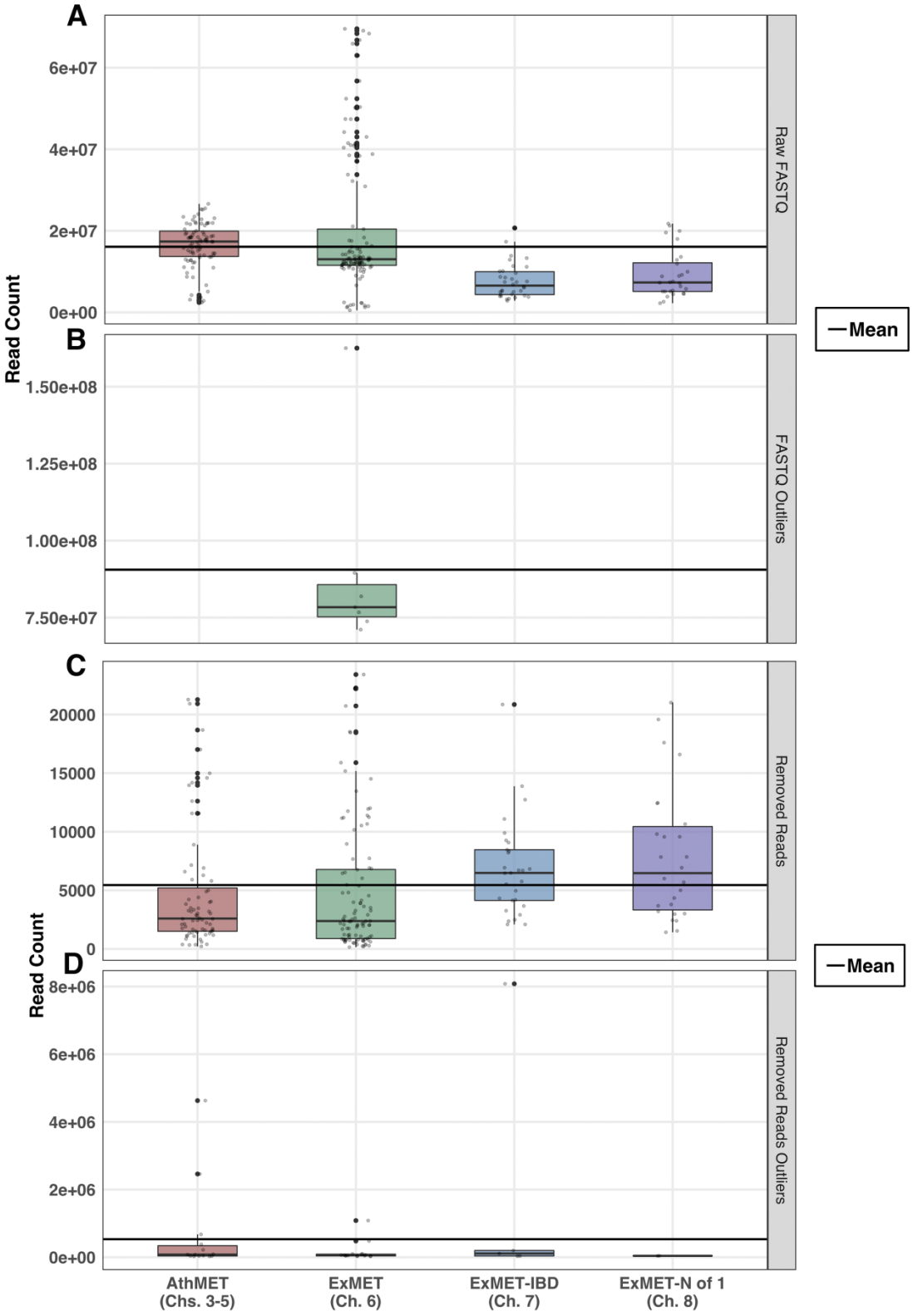


Figure 2.1 | Comparison of sequences from all studies. Quantification of sequencing reads yielded from all studies included in this thesis. **(A)** Raw FASTQ reads and **(B)** the outlying relatively high numbers of FASTQ reads. **(C)** FASTQ reads removed due to mapping to the human genome and **(D)** such reads that were particularly high.

2.7 | Metabolomic phenotyping

2.7.1 | Sample preparation

Faecal and urine samples donated by participants were stored at the Teagasc Moorepark research facility at -80°C before being shipped on dry ice for metabolomic analysis at the Centre for Computational and Systems Medicine in Imperial College London.

Frozen urine samples (-80°C) were thawed, vortexed and then centrifuged at 1600 × g for 10 minutes to remove particulates and precipitated proteins. Urine samples were prepared for metabolic profiling analysis by reversed phase (RP) and hydrophilic interaction chromatography (HILIC) ultra performance liquid chromatography – mass spectrometry (UPLC-MS) as follows: 200 µl of supernatant was diluted (1:1) with high purity (HPLC grade) water, vortexed, centrifuged at 2700 × g for 20 minutes and aliquoted for HILIC and RP methods. Quality control (QC) samples were prepared by pooling 50 µl volumes of each sample. During the analysis, the samples were maintained at 4°C in the autosampler. For ^{1}H NMR spectroscopy, 540 µL of urine samples were mixed with 60 µL of phosphate buffer (pH 7.4, 80% D2O) containing 1 mM of the internal standard, 3-(trimethylsilyl)-[[2,2,3,3,-2H4]]-propionic acid (TSP) and 2mM sodium azide ($Na^{3}N$), as described previously.[17]

Frozen faecal samples (-80°C) underwent x2 freeze thaw cycles to lyse the cells. After thoroughly defrosting, 100mg of homogenised sample was placed in a microtube containing 250 µl of 25% acetonitrile (1 ACN : 3 H_2O), 2mM sodium azide and ~0.05g 1mm Zirconia beads. The microtubes underwent 10 seconds in a Biospec bead beater. Samples were then centrifuged at 16000 x g for 20 mins. Following this the faecal water supernatant was centrifuged through centrifuge tube filters (cellulose acetate membrane, pore size 0.22 µm) to remove any remaining particular matter. The centrifuge tube filters were washed prior to

use three times with 25% acetonitrile. The resulting faecal water was prepared for metabolic profiling analysis by HILIC and bile acid profiling UPLC-MS as follows: 150ul of faecal water was diluted 3:1 with acetonitrile. Samples were vortexed and incubated at -20^{o}C for 1 hour. Following this, samples were centrifuged at 4^{o}C at 16000 x g for 1 hour. Quality control (QC) samples were prepared by pooling 20 µl volumes of each faecal water sample and then preparing as above. For ^{1}H nuclear magnetic resonance (NMR) spectroscopy, 50 µl of the filtered faecal water was added to a Pyrex glass tube, which was placed under Nitrogen gas flow for 30 mins or until all the liquid had evaporated. The dried sample was reconstituted with 540 µl of D2O and 60 µl of phosphate buffer solution as described above. The solution was mixed and sonicated for 5 minutes before undergoing further centrifugation at 14000 RPM for 10 mins before 600 µl supernatant was transferred to a NMR tube for ^{1}H-NMR spectral acquisition.

Faecal samples from participants in studies described in chapters 3-6 were prepared for targeted analysis of short-chain fatty acids (SCFA) using gas chromatography – mass spectrometry (GC-MS) as previously described.[18] In brief, 100mg of thawed faecal sample was suspended in 1ml of water with 0.5% phosphoric acid. After acidification, samples were vortexed for 2 min and centrifuged for 10 min at 16000 x g. 1ml of the resulting faecal water supernatant was added to 1 ml of ethyl acetate for 2 min and then centrifuged for 10 min at 16000 x g. Prior to analysis, a 600ul volume of the organic phase was transferred into a silanised vial with 4-methyl valeric acid added as the internal standard (IS) at a final concentration of 500uM. Samples were analysed in a random order with QCs every ten samples. Calibration curves of the measured SCFA were derived through analysis of duplicate dilution series of the purchased chemical standards at the beginning and end of the run.

2.7.2 | LC-MS Metabolic profiling analysis

Reversed-phased (RP), HILIC and bile acid UPLC-MS metabolic profiling experiments were performed using a Waters Acquity Ultra Performance LC system (Waters, Milford, MA, USA) coupled to Xevo G2 Q-TOF mass spectrometer (Waters, Milford, MA, USA) with an electrospray source. Samples were analysed in a random order, with QCs every ten samples.

Urine samples were first analysed using UPLC-MS, with a RP chromatographic method with both positive and negative MS ionisation modes. Secondly, to separate and detect more polar molecules, a HILIC chromatographic stage was used with positive MS ionisation modes. Faecal water samples underwent analysis using HILIC, and for samples used in chapters 3-5, bile acid profiling chromatographic methods in positive and negative ionisation modes respectively.

HILIC, Reversed-Phase and bile acid profiling liquid chromatographic separation was performed as previously described.[19, 20] Mass spectrometry was performed with the following settings: capillary and cone voltages were set at 1.5 kV and 30 V, respectively. The desolvation gas was set to 1000 L/hr at a temperature of 600°C; the cone gas was set to 50 L/hr and the source temperature was set to 120°C. For mass accuracy a lock-spray interface was used with leucine enkephalin [[556.27741 Da ([[M+H]]+), 554.2615 Da ([[M-H]]-)]] solution used as the lock mass at a concentration of 2000 ng/ml and at a flow rate of 15 μl/min.

2.7.3 | ^{1}H-NMR Metabolic profiling analysis

^{1}H-NMR spectroscopy was performed on the aqueous phase extracts at 300 K on a Bruker 600 MHz spectrometer (Bruker Biospin, Germany) using the following standard one-dimensional pulse sequence: RD – g_{z1} – 90° – t_1 – 90° – t_m – g_{z2} – 90° – ACQ.[17] The

relaxation delay (RD) was set at 4 s, 90° represents the applied 90° radio frequency pulse, interpulse delay (t_1) was set to an interval of 4 μs, mixing time (tm) was 10 ms, magnetic field gradients (g_{z1} and g_{z2}) were applied for 1 ms and the acquisition period (AQA) was 2.7 s. Water suppression was achieved through irradiation of the water signal during RD and t_m. For the urine samples, each spectrum was acquired using 4 dummy scans followed by 32 scans while faecal spectrum were acquired using 256 scans and 4 dummy scans and collected into 64K data points. A spectral width of 12 000Hz was used for all the samples. Prior to Fourier transformation, the FIDs were multiplied by an exponential function corresponding to a line broadening of 0.3 Hz.

2.7.4 | GC-MS SCFA targeted analysis

For samples used in chapters 3-6, the GC-MS targeted SCFA analysis was conducted on an Agilent 7890B GC system, coupled to an Agilent 5977A mass selective detector (Agilent Technologies, USA). The analysis was performed to detect levels of the SCFAs acetate, propionate, butyrate, valerate, isobutyrate, isovalerate, according to a previously described method.[18] The detector was operated in selected ion monitoring (SIM) mode (electron energy 70 eV), scanning the selected characteristic target ion for each measured SCFA (acetate, propionate, butyrate, valerate, isobutyrate, and isovalerate), at the corresponding retention times. Retention times were confirmed prior to analysis through analysis of authentic SCFAs in full scan mode.

2.7.5 | LC-MS data treatment

The raw mass spectrometric data acquired were pre-processed using xcms in R and the centwave peak picking methods were used to detect chromatographic peaks.[21] The xcms-centwave parameters were dataset specific. Feature grouping across samples was performed using the 'nearest' method within xcms. Peak filling, MinFrac (0.5) and CV (0.3)

filters were applied to the features. Data was normalised using median fold change normalisation using the median data set as the reference.[22]

2.7.6 | ^{1}H-NMR data treatment

^{1}H-NMR spectra were automatically corrected for phase and baseline distortions and referenced to the TSP singlet at δ 0.0 using TopSpin 3.1 software. Spectra were then digitized into 20,000 data points at a resolution of 0.0005ppm using an in-house MATLAB R2014a (Mathworks) script. Subsequently, spectral regions corresponding to the internal standard (δ -0.5 to 0.5) and water (δ 4.6 to 5) peaks were removed. In addition, urea (δ 5.4 to 6.3) was removed from the urinary spectra. All spectra were normalised using median fold change normalisation using the median spectrum as the reference.[22]

2.7.7 | GC-MS data treatment

GC-MS data was processed using MassHunter Quantitative Analysis (Agilent Technologies) software. Extracted ion chromatograms of the target ion selected for each SCFA were integrated and the peak area was normalised to the internal standard (4-methyl valeric acid) to correct for variability in the instrument response. Calibration curves were constructed by plotting the internal standard normalised area of authentic SCFA standards against the corresponding known SCFA concentrations and used to calculate the measured concentrations of SCFAs in the analysed samples.

2.7.8 | Metabolite ID

Confirmation of metabolite identities in the NMR data was obtained using 1D ^{1}H NMR sequence with water pre-saturation and 2D NMR experiments such as J-Resolved spectroscopy, ^{1}H-^{1}H TOtal Correlation SpectroscopY (TOCSY), ^{1}H-^{1}H COrrelation SpectroscopY (COSY), ^{1}H-^{13}C Hetero-nuclear Single Quantum Coherence (HSQC) and ^{1}H-

^{13}C Hetero-nuclear Multiple-Bond Correlation (HMBC) spectroscopy. In addition, statistical tools such as SubseT Optimization by Reference Matching (STORM) and Statistical TOtal Correlation SpectroscopY (STOCSY) were also applied. Confirmation of metabolites identities in the LC-MS data was obtained using Tandem MS (MS/MS) on selected target ions with an energy ramp 5-20eV to produce product ions.

Metabolite identification was characterized by a level of assignment (LoA) score that describes how the identification was made.[23] The levels used were as follows: LoA 1: Identified compound, confirmed by comparison to an authentic chemical reference. LoA 2: MS/MS precursor and product ions or 1D+2D NMR chemical shifts and multiplicity match to a reference database or literature to putatively annotate compound. LoA 3: Chemical shift (δ) and multiplicity matches a reference database to tentatively assign the compound.

2.8 | Quantification and statistical analysis

2.8.1 | Power analysis and sample size

For chapters related to examination of the athlete microbiome (chapters 3-5), sample size was defined by the number of players within the athletic team (n = 40). As the entire team was recruited, a number controls not less than the athletes were included for comparison (n = 46). *Post hoc* power analysis, was performed in R using the power.t.test function in the base *stats* package for pairwise comparisons, while pwr.anova.test from the *pwr* package was used for comparisons of 2 or more groups [24]. For the study described in chapter 3, at the given outcome standard deviation (σ) of 2.7%, effect size (δ) of 4.5%, significance level of 5%, and minimum sample size (n) of 23, power was calculated at 99.9% for the measurement of Shannon's alpha diversity index. For comparison, δ of 10% with the same σ, provides power of ~1 with a group size of 7. Comparing alpha diversity between all groups with a significance of 5%, and δ of 52% (given as ANOVA F value), power of 95.1% was

calculated with a group size of 20. The intervention studies described in chapters 6 and 7 had more explicit experimental aims, and accordingly utilised methods to determine appropriate sample sizes. For chapter 6, an estimation of required sample size was calculated with Mead's resource equation, yielding a minimum total sample size of 75 participants for the 3 interventions. Retrospective power analysis of this study illustrated that the detection of an 8.6% change in alpha diversity for the group size of 17 and significance level of 5% was powered at 12% for paired analysis. Balanced one-way analysis of variance power calculation on all 3 intervention groups showed that the effect size of 61%, with significance level of 5% was powered at 97% for the group size used (n = 17), while a power level of ~1 would be reached at 54 participants per group. Sample size for the study presented in chapter 7 was calculated to detect a 2% body fat percentage reduction with a two-sided significance level of 5% and 80% power, resulting in a minimum of 7 control patients and 14 patients in the exercise group. Additionally, analysis of alpha diversity showed that the testing of the microbiome was powered at 29% for a significance level of 5% and effect size 5% Due to the nature of the *N of 1* style study presented in chapter 8, statistical power calculations were inappropriate.

2.8.2 | Clinical data

Statistical analysis of participants' clinical measures (chapters 6 - 8) was carried out using the Statistical Package for the Social Sciences V.23 (SPSS, Inc., Chicago, Illinois). Due to the predominance of non-normally distributed data, nonparametric analyses were performed to compare clinical and demographic variables between experimental groups (chapters 6 and 7). Clinical data are presented as medians and interquartile ranges (IQR), unless stated otherwise. Between-group differences in baseline, follow-up, and post intervention changes (Δ) in clinical and demographic data were compared using the Kruskal-Wallis test. For

significantly different results, a Mann-Whitney *U* test was performed to determine the groups between which this difference applied. Where stated, the Wilcoxon signed-rank test was used to compare baseline and post intervention values within intervention groups. A type I error rate of ≤0.05 was considered significant in all cases.

2.8.3 | Microbiome analysis

Statistical analysis of microbiome data was conducted within the R statistical programing environment (V.3.2.2 for chapter 3 and V.3.3.2 for all other chapters). As with clinical variables, the predominance of non-normally distributed data from microbiome measurements was accommodated with the use of nonparametric statistical tests. Statistical variation of measures between 3 or more groups was carried out with the Kruskal-Wallis H test with compareGroups (V.2.0), while pairwise comparisons were performed with the Wilcoxon signed-rank test unless otherwise noted. Correlations of variables were done with Spearman's rank test. Correction of *P* values relating to microbiome and metabolomic analysis was performed using the Benjamini-Hochberg false-discovery rate (FDR) [25] in the base *stats* package in R.

Beta diversity, the measurement of variable diversity as it is defined by habitat type (i.e. experimental group), was assessed with the Bray-Curtis index of dissimilarity using a combination of principal coordinate analysis (PCoA) and non-metric multidimensional scaling (NMDS), both performed with the vegan R package (V.2.4-3) [26-29]. Statistical assessment of dissimilarity matrices (Bray-Curtis) derived from microbial data was facilitated with permutational multivariate analysis of variance (PERMANOVA) tests as implemented by the adonis2 function in the vegan package. Identification of statistically relevant taxonomic features was performed with the analysis of composition of microbiomes (ANCOM) test as implemented in the R package of the same name (V.1.1-3)

[30]. Detection of underlying features of metabolic pathways in chapters 5 and 6 was performed with unsupervised cross-validated partial-least-squares–discriminant analysis (PLS-DA) and the KODAMA algorithm from the R package of the same name (V.1.4) [31]. Diversity measurements of microbiome variables (e.g. species and metabolic pathways) defined by site (i.e. participant),alpha diversity, and calculations of relative abundances were also performed with the vegan package [27]. Relative-abundance data were generated separately for identified species within each phylogenetic domain (e.g., Bacteria).

For metabolomic analysis, the resulting ^{1}H-NMR and LC-MS data sets were imported into MatLab to conduct multivariate statistical analysis. Data were centred and scaled to account for the repeated-measures design and then modelled using partial-least-squares–discriminant analysis (PLS-DA) with Monte Carlo cross-validation (MCCV) [32]. The fit and predictability of the models obtained were determined and expressed as R2 and Q2 values, respectively.

2.8.4 | Metabonomic Statistical and Bioinformatics Analysis

The resulting ^{1}H-NMR and LC-MS data sets were imported into SIMCA 14.1 (Umetrics) to conduct multivariate statistical analysis. Principal Component Analysis (PCA), followed by an Orthogonal Partial Least Squares Discriminant Analysis (OPLS-DA) was performed to examine the data sets and to observe clustering in the results according to the predefined classes. The OPLS-DA models in the current study were established based on one PLS component and one orthogonal component. Unit variance scaling was applied to ^{1}H-NMR data, Pareto scaling was applied to MS data. The fit and predictability of the models obtained was determined by the R^2Y and Q^2Y values, respectively.

Significant metabolites were obtained from LC-MS OPLS-DA models through division of the regression coefficients by the jack-knife interval standard error to give an

estimate of the t-statistic. Variables with a t-statistic ≥ 1.96 (z-score, corresponding to the 97.5 percentile) were considered significant. Significant metabolites were obtained from 1H-NMR OPLS-DA models after investigating correlations with correlation coefficients values higher than 0.4.

Univariate statistical analysis was used to examine the SCFA data set. The data was not normally distributed; hence the Mann-Whitney U test was performed to examine differences between classes.

2.9 | References

[1] P. D. Thompson, R. Arena, D. Riebe, L. S. Pescatello, M. American College of Sports, ACSM's new preparticipation health screening recommendations from ACSM's guidelines for exercise testing and prescription, ninth edition, Curr Sports Med Rep 12 (4) (2013) 215-7.

[2] U. Ekelund, H. A. Ward, T. Norat, J. Luan, A. M. May, E. Weiderpass, S. J. Sharp, K. Overvad, J. N. Ostergaard, A. Tjonneland, N. F. Johnsen, S. Mesrine, A. Fournier, G. Fagherazzi, A. Trichopoulou, P. Lagiou, D. Trichopoulos, K. Li, R. Kaaks, P. Ferrari, I. Licaj, M. Jenab, M. Bergmann, H. Boeing, D. Palli, S. Sieri, S. Panico, R. Tumino, P. Vineis, P. H. Peeters, E. Monnikhof, H. B. Bueno-de-Mesquita, J. R. Quiros, A. Agudo, M. J. Sanchez, J. M. Huerta, E. Ardanaz, L. Arriola, B. Hedblad, E. Wirfalt, M. Sund, M. Johansson, T. J. Key, R. C. Travis, K. T. Khaw, S. Brage, N. J. Wareham, E. Riboli, Physical activity and all-cause mortality across levels of overall and abdominal adiposity in European men and women: the European Prospective Investigation into Cancer and Nutrition Study (EPIC), Am J Clin Nutr 101 (3) (2015) 613-21.

[3] W. C. Willett, L. Sampson, M. J. Stampfer, B. Rosner, C. Bain, J. Witschi, C. H. Hennekens, F. E. Speizer, Reproducibility and validity of a semiquantitative food frequency questionnaire, Am J Epidemiol 122 (1) (1985) 51-65.

[4] N. J. Wareham, R. W. Jakes, K. L. Rennie, J. Mitchell, S. Hennings, N. E. Day, Validity and repeatability of the EPIC-Norfolk Physical Activity Questionnaire, International Journal of Epidemiology 31 (1) (2002) 168-74.

[5] Y. Kim, I. Park, M. Kang, Convergent validity of the international physical activity questionnaire (IPAQ): meta-analysis, Public Health Nutr 16 (3) (2013) 440-52.

[6] G. M. Kline, J. P. Porcari, R. Hintermeister, P. S. Freedson, A. Ward, R. F. McCarron, J. Ross, J. M. Rippe, Estimation of VO2max from a one-mile track walk, gender, age, and body weight, Med Sci Sports Exerc 19 (3) (1987) 253-9.

[7] M. Brzycki, A practical approach to strength training., McGraw-Hill

[8] G. Borg, Borg's perceived exertion and pain scales, Human kinetics

[9] B. E. Ainsworth, W. L. Haskell, S. D. Herrmann, N. Meckes, D. R. Bassett, Jr., C. Tudor-Locke, J. L. Greer, J. Vezina, M. C. Whitt-Glover, A. S. Leon, 2011 Compendium of Physical Activities: a second update of codes and MET values, Med Sci Sports Exerc 43 (8) (2011) 1575-81.

[10] H. Mirsepasi, S. Persson, C. Struve, L. O. Andersen, A. M. Petersen, K. A. Krogfelt, Microbial diversity in fecal samples depends on DNA extraction method: easyMag DNA extraction compared to QIAamp DNA stool mini kit extraction, BMC Res Notes 7 (2014) 50.

[11] S. Abubucker, N. Segata, J. Goll, A. M. Schubert, J. Izard, B. L. Cantarel, B. Rodriguez-Mueller, J. Zucker, M. Thiagarajan, B. Henrissat, O. White, S. T. Kelley, B. Methe, P. D.

Schloss, D. Gevers, M. Mitreva, C. Huttenhower, Metabolic reconstruction for metagenomic data and its application to the human microbiome, PLoS Comput Biol 8 (6) (2012) e1002358.

[12] R. Caspi, H. Foerster, C. A. Fulcher, P. Kaipa, M. Krummenacker, M. Latendresse, S. Paley, S. Y. Rhee, A. G. Shearer, C. Tissier, The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases, Nucleic acids research 36 (suppl 1) (2008) D623-D631.

[13] P. Menzel, K. L. Ng, A. Krogh, Fast and sensitive taxonomic classification for metagenomics with Kaiju, Nat Commun 7 (2016) 11257.

[14] N. Segata, L. Waldron, A. Ballarini, V. Narasimhan, O. Jousson, C. Huttenhower, Metagenomic microbial community profiling using unique clade-specific marker genes, Nat Methods 9 (8) (2012) 811-4.

[15] D. E. Wood, S. L. Salzberg, Kraken: ultrafast metagenomic sequence classification using exact alignments, Genome Biol 15 (3) (2014) R46.

[16] F. Asnicar, G. Weingart, T. L. Tickle, C. Huttenhower, N. Segata, Compact graphical representation of phylogenetic data and metadata with GraPhlAn, PeerJ 3 (2015) e1029.

[17] A. C. Dona, B. Jimenez, H. Schafer, E. Humpfer, M. Spraul, M. R. Lewis, J. T. Pearce, E. Holmes, J. C. Lindon, J. K. Nicholson, Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping, Anal Chem 86 (19) (2014) 9887-94.

[18] R. Garcia-Villalba, J. A. Gimenez-Bastida, M. T. Garcia-Conesa, F. A. Tomas-Barberan, J. Carlos Espin, M. Larrosa, Alternative method for gas chromatography-mass spectrometry analysis of short-chain fatty acids in faecal samples, J Sep Sci 35 (15) (2012) 1906-13.

[19] E. J. Want, I. D. Wilson, H. Gika, G. Theodoridis, R. S. Plumb, J. Shockcor, E. Holmes, J. K. Nicholson, Global metabolic profiling procedures for urine using UPLC-MS, Nat Protoc 5 (6) (2010) 1005-18.

[20] M. H. Sarafian, M. R. Lewis, A. Pechlivanis, S. Ralphs, M. J. McPhail, V. C. Patel, M. E. Dumas, E. Holmes, J. K. Nicholson, Bile acid profiling and quantification in biofluids using ultra-performance liquid chromatography tandem mass spectrometry, Anal Chem 87 (19) (2015) 9662-70.

[21] R. Tautenhahn, C. Bottcher, S. Neumann, Highly sensitive feature detection for high resolution LC/MS, BMC Bioinformatics 9 (2008) 504.

[22] K. A. Veselkov, L. K. Vingara, P. Masson, S. L. Robinette, E. Want, J. V. Li, R. H. Barton, C. Boursier-Neyret, B. Walther, T. M. Ebbels, I. Pelczer, E. Holmes, J. C. Lindon, J. K. Nicholson, Optimized preprocessing of ultra-performance liquid chromatography/mass spectrometry urinary metabolic profiles for improved information recovery, Anal Chem 83 (15) (2011) 5864-72.

[23] L. W. Sumner, A. Amberg, D. Barrett, M. H. Beale, R. Beger, C. A. Daykin, T. W. Fan, O. Fiehn, R. Goodacre, J. L. Griffin, T. Hankemeier, N. Hardy, J. Harnly, R. Higashi, J. Kopka, A. N. Lane, J. C. Lindon, P. Marriott, A. W. Nicholls, M. D. Reily, J. J. Thaden, M. R. Viant, Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI), Metabolomics 3 (3) (2007) 211-221.

[24] Y. Xia, J. Sun, D.-G. Chen, Statistical Analysis of Microbiome Data with R,

[25] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, Journal of the royal statistical society. Series B (Methodological) (1995) 289-300.

[26] J. Oksanen, F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. O'Hara, G. L. Simpson, P. Solymos, M. Henry, H. Stevens, E. Szoecs, H. Wagner, Vegan: Community Ecology Package. R package version 2.4-3,

[27] R. H. Whittaker, Vegetation of the Siskiyou Mountains, Oregon and California, Ecological Monographs 30 (3) (1960) 279-338.

[28] P. R. Minchin, An evaluation of the relative robustness of techniques for ecological ordination, Vegetatio 69 (1) (1987) 89-107.

[29] J. C. Gower, Some distance properties of latent root and vector methods used in multivariate analysis, biometrika 53 (1966) 325-338.

[30] S. Mandal, W. Van Treuren, R. A. White, M. Eggesbo, R. Knight, S. D. Peddada, Analysis of composition of microbiomes: a novel method for studying microbial composition, Microb Ecol Health Dis 26 (2015) 27663.

[31] S. Cacciatore, L. Tenori, C. Luchinat, P. R. Bennett, D. A. MacIntyre, KODAMA: an R package for knowledge discovery and data mining, Bioinformatics 33 (4) (2017) 621-623.

[32] I. Garcia-Perez, J. M. Posma, R. Gibson, E. S. Chambers, T. H. Hansen, H. Vestergaard, T. Hansen, M. Beckmann, O. Pedersen, P. Elliott, J. Stamler, J. K. Nicholson, J. Draper, J. C. Mathers, E. Holmes, G. Frost, Objective assessment of dietary patterns by use of metabolic phenotyping: a randomised, controlled, crossover trial, Lancet Diabetes Endocrinol 5 (3) (2017) 184-195.

Chapter 3

THE MICROBIOME OF PROFESSIONAL ATHLETES DIFFERS FROM THAT OF CONTROLS NOT ONLY IN COMPOSITION BUT PARTICULARLY AT THE FUNCTIONAL METABOLIC LEVEL

3.1 | Abstract

It is evident that the gut microbiome, and therefore, factors driving its composition and activity, both influence human metabolic, immunological, and developmental processes. We previously reported that extreme physical activity and associated dietary adaptations, are associated with changes in faecal microbial diversity and composition relative to that of individuals with a more sedentary lifestyle. Here we address the impact of these factors on the functionality/metabolic activity of the gut microbiota in athletes, which reveals even greater separation between exercise and a more sedentary state. Athletes had relative increases in pathways (e.g. amino acid and antibiotic biosynthesis and carbohydrate metabolism) and faecal metabolites (e.g. microbial produced short chain fatty acids [SCFAs] acetate, propionate, and butyrate) associated with enhanced muscle turnover and overall health when compared to control groups. Differences in faecal microbiota between athletes and less active controls show even greater separation in metagenomic and metabolomics dynamics than at compositional levels, and provide added insight into the diet-exercise-gut microbiota paradigm.

In addition to the author, the original material subject of this chapter was contributed to accordingly:

Metabolomic processing and data analysis: Dr. Nicholas C. Penney, Dr. Isabel Garcia Perez, & Professor Elaine Holmes

Presented material has been published at the time of this thesis' submission.

The microbiome of professional athletes differs from that of controls not only in composition but particularly at the functional metabolic level

Wiley Barton[1,2,3], Nicholas C. Penney[4, 5], Owen Cronin[1,3], Isabel Garcia Perez[4], Michael G. Molloy[1,3], Elaine Holmes[4], Fergus Shanahan[1,3*], Paul D. Cotter[1,2], Orla O'Sullivan[1,2]

[1]Alimentary Pharmabiotic Centre Microbiome Institute, University College Cork, National University of Ireland, Cork, Republic of Ireland

[2]Teagasc Food Research Centre, Moorepark, Fermoy, Cork, Republic of Ireland

[3]Department of Medicine, University College Cork, National University of Ireland, Cork, Republic of Ireland

[4]Section of Biomolecular Medicine, Division of Computational Systems Medicine, Department of Surgery and Cancer, Imperial College London, UK

[5]Division of Surgery, Department of Surgery and Cancer, Imperial College London, UK

3.2 | Introduction

Regular exercise challenges systemic homeostasis resulting in a breadth of multi-organ molecular and physiological responses, including many that centre on immunity, metabolism and the microbiome-gut-brain axis [1-5]. Exercise exhibits systemic and end-organ anti-inflammatory effects, as well as contributing to more efficient carbohydrate metabolism, in addition to trophic effects at the level of the central nervous system [6, 7]. In fact, increasing physical activity offers an effective treatment and preventative strategy for many chronic conditions in which the gut microbiome has been implicated [8-10]. Conversely, a sedentary lifestyle is a major contributing factor to morbidity in developed Western society and is associated with heightened risk of numerous *diseases of affluence*, such as obesity, diabetes, asthma, and cardiovascular disease [11-14]. Recent evidence supports an influential role for the gut microbiome in these diseases [15-23].

The concept that regular exercise and sustained levels of increased physical activity foster or assist the maintenance of a preferential intestinal microbiome has recently gained momentum and interest [24-29]. Previously, using 16S rRNA amplicon sequencing, we demonstrated taxonomic differences in gut microbiota between an elite athlete cohort of international-level rugby players and a group of age-matched high (>28 kg/m^2) and low (<25 kg/m^2) BMI controls [26]. This analysis illustrated a significantly greater intestinal microbial diversity amongst the athletes compared to both control groups. This taxonomic diversity significantly correlated with exercise and dietary protein consumption. However, the possibility existed that these differences did not equate to differences at a functional level. Here, we re-examine the microbiome in these participants by whole metagenome shotgun sequencing to provide deeper insight into taxonomic composition and metabolic potential and by complementary metabolic phenotyping analyses of host- and microbial-derived

(urine and faecal respectively) metabolic profiles. This analysis shows that the differences in the gut microbiota between athletes and controls is even more pronounced at the functional metabolic level than at the compositional level as previously reported, and provides further rationale for prospective controlled studies to unravel the relationship between diet, exercise and the gut microbiome.

3.3 | Study objective

It is evident that the gut microbiota and factors that influence its composition and activity influence human metabolic, immunological, and developmental processes. We previously reported that extreme physical activity with associated dietary adaptations, such as that pursued by professional athletes, is associated with changes in faecal microbial diversity and composition relative to that of individuals with a more sedentary lifestyle. Here we address the impact of these factors on the functionality/metabolic activity of the microbiota which reveals even greater separation between exercise and a more sedentary state.

3.4 | Methods

3.4.1 | Study population

Elite professional male athletes (n = 40) and healthy controls (n = 46) matched for age and gender were enrolled in 2011 as previously described in the study [26]. Due to the range of physiques within a rugby team (player position dictates need for a variety of physical constitutions, i.e. forward players tend to have larger BMI values than backs, often in the overweight/obese range) the recruited control cohort was subdivided into two groups. In order to more completely include control participants, the BMI parameter for group inclusion was adjusted to BMI ≤ 25.2 and BMI ≥ 26.5 for the low BMI and high BMI

groups respectively. Approval for this study was granted by the Cork Clinical Research Ethics Committee.

3.4.2 | Acquisition of clinical, exercise and dietary data

Self-reported dietary intake information was accommodated by a research nutritionist within the parameters of a food frequency questionnaire (FFQ) in conjunction with a photographic food atlas as per the initial investigation [26]. Fasting blood samples were collected and analysed at the Mercy University Hospital clinical laboratories, Cork. As the athletes were involved in a rigorous training camp we needed to assess the physical activity levels of both control groups. To determine this we used an adapted version of the EPIC-Norfolk questionnaire [30]. Creatine kinase levels were used as a proxy for level of physical activity across all groups.

3.4.3 | Preparation of Metagenomic libraries

DNA derived from faecal samples was extracted and purified using the QIAmp DNA Stool Mini Kit (cat. no. 51504) prior to storage at -80°C. DNA libraries were prepared with the Nextera XT DNA Library Kit (cat. no. FC-131-1096) prior to processing on the Illumina HiSeq 2500 sequencing platform (refer to section 2.5 for further detail).

3.4.4 | Metagenomic statistical and bioinformatic analysis

Delivered raw FASTQ sequence files were quality checked as follows: contaminating sequences of human origin were first removed through the NCBI Best Match Tagger (BMTagger). Poor quality and duplicate read removal, as well as trimming was implemented using a combination of SAM and Picard tools. Processing of raw sequence data produced a total of 2,803,449,392 filtered reads with a mean read count of 32,598,248.74 (± 10,639,447 SD) per each of the 86 samples. These refined reads were then subjected to functional

profiling by the most recent iteration of the Human Microbiome Project (HMP) Unified Metabolic Analysis Network (HUMAnN2 V.0.5.0) pipeline [31]. The functional profiling performed by HUMAnN2 composed tabulated files of microbial metabolic pathway abundance and coverage derived from the Metacyc database [32]. Microbial pathway data was statistically analysed in the R software environment (V.3.2.2) (for further details see section 2.6) [33]. The GraPhlAn software tool was used to construct the circular phylogenetic trees that provided the basis for comparison of taxonomic profiles generated from original 16S sequencing and the shotgun metagenomic sequencing introduced in this present chapter. All presented p values were corrected for multiple comparisons using the Benjamini-Hochberg False Discovery Rate (pFDR) method [34].

3.4.5 | Metabolic profiling

Urine and faecal samples were prepared for metabonomic analysis as previously described [35, 36]. Utilising established methods, urine samples underwent ^{1}H-NMR, reversed-phased (RP) and hydrophilic interaction chromatography (HILIC) profiling experiments. Faecal samples underwent ^{1}H-NMR, hydrophilic interaction chromatography (HILIC) and bile acid UPLC-MS profiling experiments and GC-MS targeted SCFA analysis [36-38].

After data pre-processing,[39] the resulting ^{1}H-NMR and LC-MS data sets were imported into SIMCA 14.1 (Umetrics) to conduct multivariate statistical analysis. Principal Component Analysis (PCA), followed by Orthogonal Partial Least Squares Discriminant Analysis (OPLS-DA) was performed to examine the data sets and to observe clustering in the results according to the predefined classes. The OPLS-DA models in the current study were established based on one PLS component and one orthogonal component. Unit variance scaling was applied to ^{1}H-NMR data, Pareto scaling was applied to MS data. The

fit and predictability of the models obtained was determined by the R^2Y and Q^2Y values, respectively. Significant metabolites were obtained from LC-MS OPLS-DA models through division of the regression coefficients by the jack-knife interval standard error to give an estimate of the t-statistic. Variables with a t-statistic ≥ 1.96 (z-score, corresponding to the 97.5 percentile) were considered significant. Significant metabolites were obtained from ^{1}H-NMR OPLS-DA models after investigating correlations with correlation coefficients values higher than 0.4. Univariate statistical analysis (Mann-Whitney U test) was used to examine the SCFA data set. P-values were adjusted for multiple testing using the Benjamini-Hochberg False Discovery Rate (pFDR) method. Confirmation of metabolite identities in the NMR data was obtained using 1D ^{1}H NMR and 2D ^{1}H-^{1}H NMR and ^{1}H-^{13}C NMR experiments. In addition, statistical tools such as SubseT Optimization by Reference Matching (STORM) and Statistical TOtal Correlation SpectroscopY (STOCSY) were also applied [40, 41]. Confirmation of metabolites identities in the LC-MS data was obtained using Tandem MS (MS/MS) on selected target ions.

Metabolite identification was characterized by a level of assignment (LoA) score that describes how the identification was made [42]. The levels used were as follows: LoA 1: Identified compound, confirmed by comparison to an authentic chemical reference. LoA 2: MS/MS precursor and product ions or 1D+2D NMR chemical shifts and multiplicity match to a reference database or literature to putatively annotate compound. LoA 3: Chemical shift (δ) and multiplicity matches a reference database to tentatively assign the compound. (For further details see sections 2.7-2.8).

3.5 | Results

The study groups were comprised of professional male athletes (n = 40) and healthy controls (n = 46) [26]. To better represent the variability of BMI in the athletes, controls were classified as either low BMI (n = 22, BMI ≤ 25.2) or high BMI (n = 24, BMI ≥ 26.5). Participants made no report of gastrointestinal (GI) distress or alterations of GI transit time throughout the course of the initial study.

3.5.1 | Functional structure of the enteric microbiome correlates with athletic state

Functional metagenomic analysis of faecal samples allowed for the prediction of the operational potential of each individual's microbiota. In total, 19,300 taxonomically linked metabolic pathways were identified in at least one individual. Comparison of phylogenetic constructions derived from the 16S rRNA amplicon data of our previous study and the functional data of this present report revealed a greater level of identification at higher levels of taxonomy (e.g. phylum) for 16S sequences,[26] while the metagenomic data had greater fidelity and superior resolution of lower levels of taxonomy (e.g. species) (Figure 3.1). Consistent with previous results, the microbiota of the athletes were significantly more diverse than that of both the low and high BMI control groups at the functional level (Figure 3.2A). Furthermore, our previous findings of an enrichment of *Akkermansia* in athletes was corroborated by the presence of significantly higher proportions of metabolic pathways associated with this genus in athletes when compared to high BMI controls ($p < 0.001$). Correlation analysis revealed that, of the total 19,300 pathways, 98 were significantly altered between the three cohorts ($p < 0.05$) (Supplementary Table 3.1, see appendix A). Subsequently, large-scale functional dissimilarity between athletes and controls was determined, and distinct patterns of pathway composition between groups were revealed (Supplementary Figure 3.1A). This functional distinction remained true whether applied to total

pathway data or to the statistically significant subset of pathways (Supplementary Figure 3.1B). Correlation of pathways present in at least one member from both cohorts further exemplified the uniformity of the athletes and the division between the athletes and control groups (Supplementary Figure 3.1C). Separation according to group membership was further illustrated through Principal Coordinate Analysis (PCoA), with statistical support of the significant separation between the athletes and both control groups ($p < 0.05$) (Figure 3.2B). This was also the case for the statistically significant subset of pathways (Supplementary Figure 3.1D). Principal Component Analysis (PCA) supplemented with a Correspondence Analysis (CA) and k-Nearest Neighbour (k-NN) semi-supervised learning approach cast further light (i.e. visualization of robustly defined class associations of specific individuals within the groups) on the clustering of participants within and between cohorts (Supplementary Figure 3.1E).

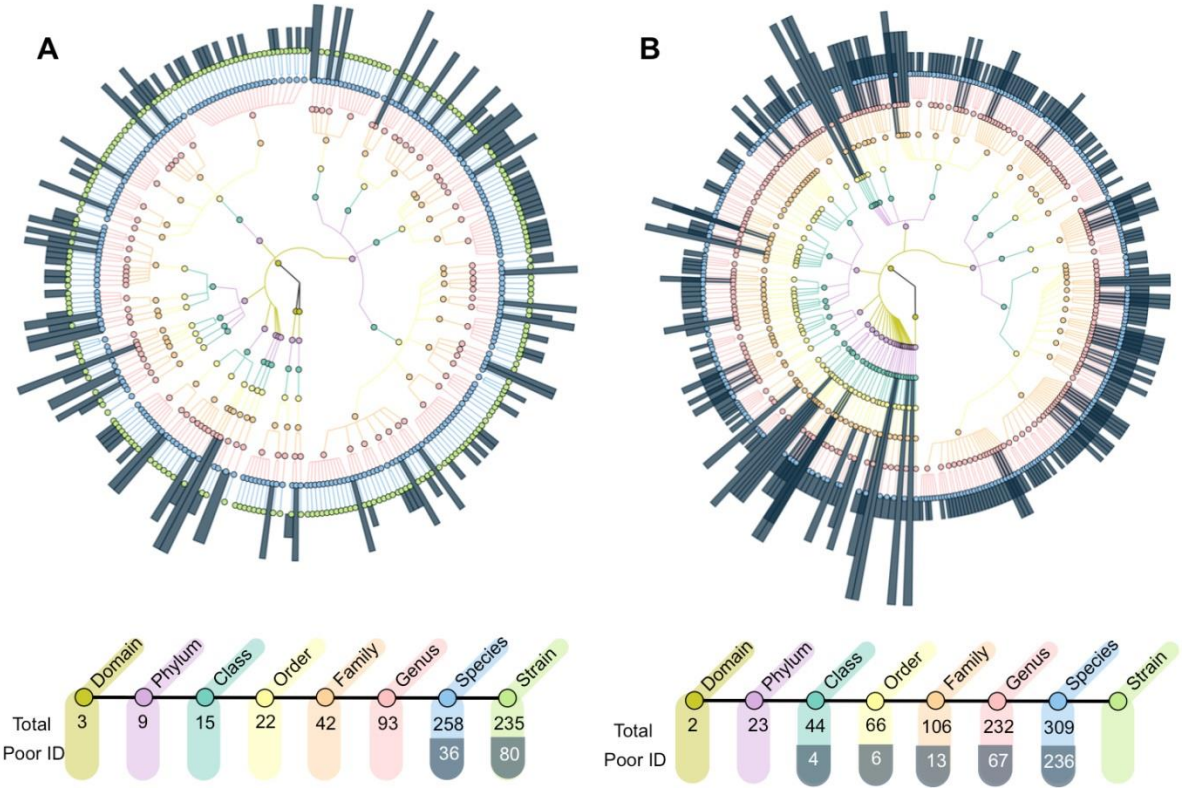


Figure 3.1 | Comparison of phylogenetic constructions from metagenomic and 16S rRNA gene sequencing sourced from all participants. Phylogenetic trees derived from **(A)** metagenomic sequencing and **(B)** 16S rRNA amplicon sequencing. Taxonomic levels are assigned from centre out with kingdom level assignment in centre and strain level assignment in outer most ring. Dark blue radial highlights correspond to poorly identified taxonomies (i.e. 'unknown' and 'unassigned' database entries). Number of assignments at each level of phylogeny is displayed below the respective graph. Taxonomic trees derived from the two sequencing approaches illustrate an advantage of metagenomic sequencing in the number of predictions of lower taxonomic levels and the frequency of full identification of taxa, while 16S rRNA sequencing grants greater insight of high level phylogenies within the population.

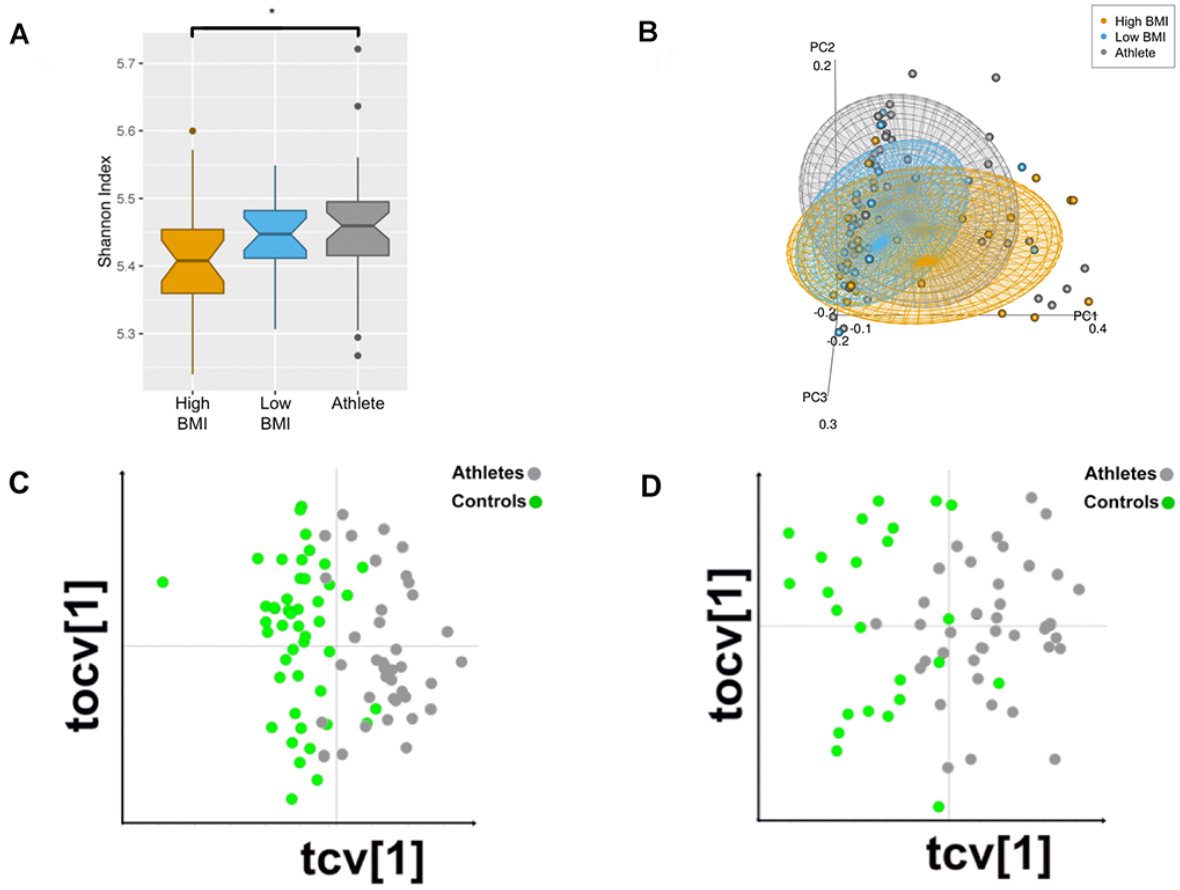


Figure 3.2 | Group-wise comparison of microbial metagenomic and metabolomic profiles. (A) Alpha diversity for metabolic pathways from all three groups presented as Shannon diversity index. Pathway diversity is increased in the athlete group when compared to low BMI and high BMI controls. Diversity measures are statistically significant between high BMI controls and athletes ($p < 0.049$), with statistical significance between all groups (Kruskal Wallis $p < 0.05$). **(B)** Principal coordinate analysis (PCoA) of Bray-Curtis compiled distance matrix illustrates beta diversity of all microbial metabolic pathway relative abundances between the three groups. Ordination of the pathways between the groups shows significant variation from one another (Adonis PERMANOVA $p < 0.05$). **(C & D)** Cross validated orthogonal partial least squares regression discriminant analysis (OPLS-DA) of full Nuclear Magnetic Resonance (^{1}H-NMR) spectra from urine ($R^2Y=0.86$, $Q^2Y=0.60$)**(C)** and faecal water ($R^2Y=0.86$, $Q^2Y=0.52$)**(D)** samples. OPLS-DA displays robust separation between the detected metabolic profiles of athletes and controls. Models are comprised of 1 predictive (tcv[1]) and 1 orthogonal (tocv[1]) principal component.

Pathways exhibiting statistically significant variation between the athletes and both control groups were organised according to MetaCyc metabolic pathway hierarchy classification (34 metabolic categories), highlighting a number of differences (Figure 3.3A, Supplementary Table 3.2 located in Appendix A). Distinct clustering patterns were observed within each cohort, with the high BMI control group having the lowest average abundance scores across 31 metabolic pathway categories (the exceptions being Vitamin Biosynthesis (VB), Lipid Biosynthesis (LB), and Amino Acid Biosynthesis (AAB) categories). The athlete group had the highest mean abundance across 29 of the 34 metabolic categories (e.g. Carbohydrate Biosynthesis [CB], Cofactor Biosynthesis [CfB], and Energy Metabolism [EM]) (Supplementary Table 3.2, Appendix A). Numerous statistically significant ($p < 0.05$) associations were identified between pathway abundances and serum Creatine Kinase—an enzymatic marker of muscle activity (CK, IU/L), total bilirubin (IU/L) and dietary macronutrient intake of protein (g/day), fibre (g/day), carbohydrates (g/day), sugars (g/day), starch (g/day), fat (g/day), and total energy (KJ/day) (Figure 3.3B). Each group was represented by distinct association profiles of the correlation between clinical measurements and metagenomic pathways. Dietary factors, sugars and other carbohydrates, as well as energy intake, provide the majority of the correlation for the control groups whereas the athlete group was predominantly correlated with CK, total bilirubin, and total energy intake. Of the total number of metabolic pathways with associations to the clinical data from all three groups (10,760; data not shown), relevant pathways related to the production of secondary metabolites, co-factors, and SCFAs were identified (e.g. biotin biosynthesis and pyruvate fermentation to butanoate).

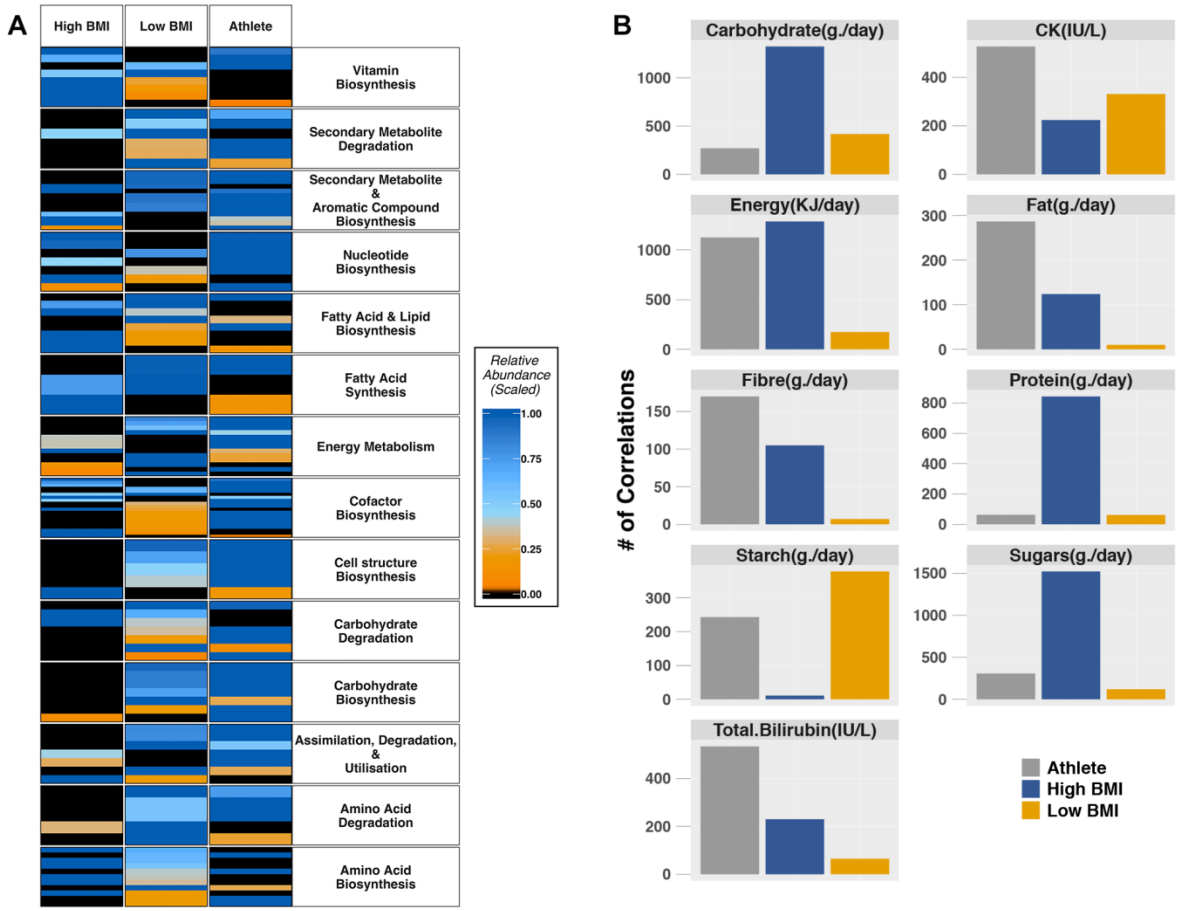


Figure 3.3 | Group variation of microbial metabolic function and associations between pathways and clinical and dietary variables. (A) Mean relative abundance values of statistically significant (Kruskal Wallis $p < 0.05$) metabolic pathways binned according to categories of metabolic function. The relative abundance values of each group (columns) and statistically detected pathways (rows) demonstrates variability in the metabolic potential of microbial communities from the separate groups. **(B)** Number of metabolic pathways significantly (Benjamini-Hochberg corrected $p < 0.05$) correlated with dietary constituents and blood serum metabolites. The relatively large number of significant associations between metabolic pathways and clinical measurements are presented as counts of associations for the three groups, and illustrates variation in the potential influence of different factors on the microbiome of participants in the separate groups.

3.5.2 | Distinct differences between host and microbial metabolites in athletes and controls

A combination of multi-platform metabolic phenotyping and multivariate analysis based on Orthogonal Partial Least Squares Discriminant Analysis (OPLS-DA) was used to compare urinary and faecal samples from athletes and controls. The cross-validated (CV) OPLS-DA models show strong differences between athletes and controls in urine samples by proton nuclear magnetic resonance (^{1}H-NMR) analysis (R^2Y=0.86, Q^2Y=0.60, Figure 3.2C), hydrophilic interaction ultra-performance liquid chromatography mass spectroscopy (HILIC UPLC-MS) positive mode analysis (R^2Y=0.85, Q^2Y=0.74, Supplementary Figure 3.2A) and reverse phase ultra-performance liquid chromatography mass spectroscopy (RP UPLC-MS) in both positive and negative mode analysis (R^2Y=0.83, Q^2Y=0.73, and R^2Y=0.83, Q^2Y=0.67, Supplementary Figure 3.2B and 3.2C respectively). Likewise, the CV-OPLS-DA models comparing faecal samples, although weaker than the urine models, reveal significant differences between athletes and controls by ^{1}H-NMR analysis (R^2Y=0.86, Q^2Y=0.52, Figure 3.2D) and HILIC UPLC-MS positive mode analysis (R^2Y=0.65, Q^2Y=0.34, Supplementary Figure 3.2D).

The loadings of the pairwise OPLS-DA models were used to identify metabolites discriminating between the two classes. Athletes' ^{1}H-NMR metabolic phenotypes were characterised by higher levels of trimethylamine-*N*-oxide (TMAO), L-carnitine, dimethyl-glycine, O-acetyl carnitine, proline betaine, creatinine, acetoacetate, 3-hydroxy-isovaleric acid, acetone, *N*-methylnicotinate, *N*-methylnicotinamide, phenylacetylglutamine (PAG) and 3-methylhistidine in urine samples and higher levels of propionate, acetate, butyrate, trimethylamine (TMA), lysine, and methylamine in faecal samples, relative to controls. Beta-alanine betaine was higher in both faecal and urine samples of athletes. Athletes were

further characterised by lower levels of glycerate, allantoin and succinate and lower levels of glycine and tyrosine relative to controls in urine and faecal samples, respectively (Supplementary Table 3.3).

While numerous metabolites discriminated significantly between athletes and controls with RP UPLC-MS positive (490) and negative (434) modes for urine, as well as with HILIC UPLC-MS positive mode for urine (196) and faecal water (3), key metabolites were structurally identified using the strategy described below. UPLC-MS analyses revealed higher urinary excretion of *N*-formylanthranilic acid, hydantoin-5-propionic acid, 3-Carboxy-4-methyl-5-propyl-2-furanpropionic acid (CMPF), CMPF glucuronide, trimetaphosphoric acid, acetylcarnitine - C2, propionylcarnitine - C3, isobutyrylcarnitine – C4, 2-Methylbutyroylcarnitine - C5, Hexanoylcarnitine - C6, C9:1-carnitine, L-valine, nicotinuric acid, 4-pyridoxic acid and creatinine in athletes relative to controls. Levels of glutamine, 7-methylxanthine, imidazoleacetic acid, isoquinoline / quinolone were lower in athletes' urinary samples relative to controls. Additionally, 16 unknown glucuronides were lower in the athlete samples (Supplementary Table 3.4).

SCFA levels in faeces measured by targeted GC-MS showed significantly higher levels of acetate, propionate, butyrate (all with $p < 0.001$) and valerate ($p = 0.011$) in athletes relative to controls. Isobutyrate and isovalerate did not differ significantly between the groups (Figure 3.4B, Supplementary Table 3.5). Furthermore, concentrations of propionate strongly correlated to protein intake while butyrate was shown to have a strong association with intake of dietary fibre (Supplementary Table 3.6).

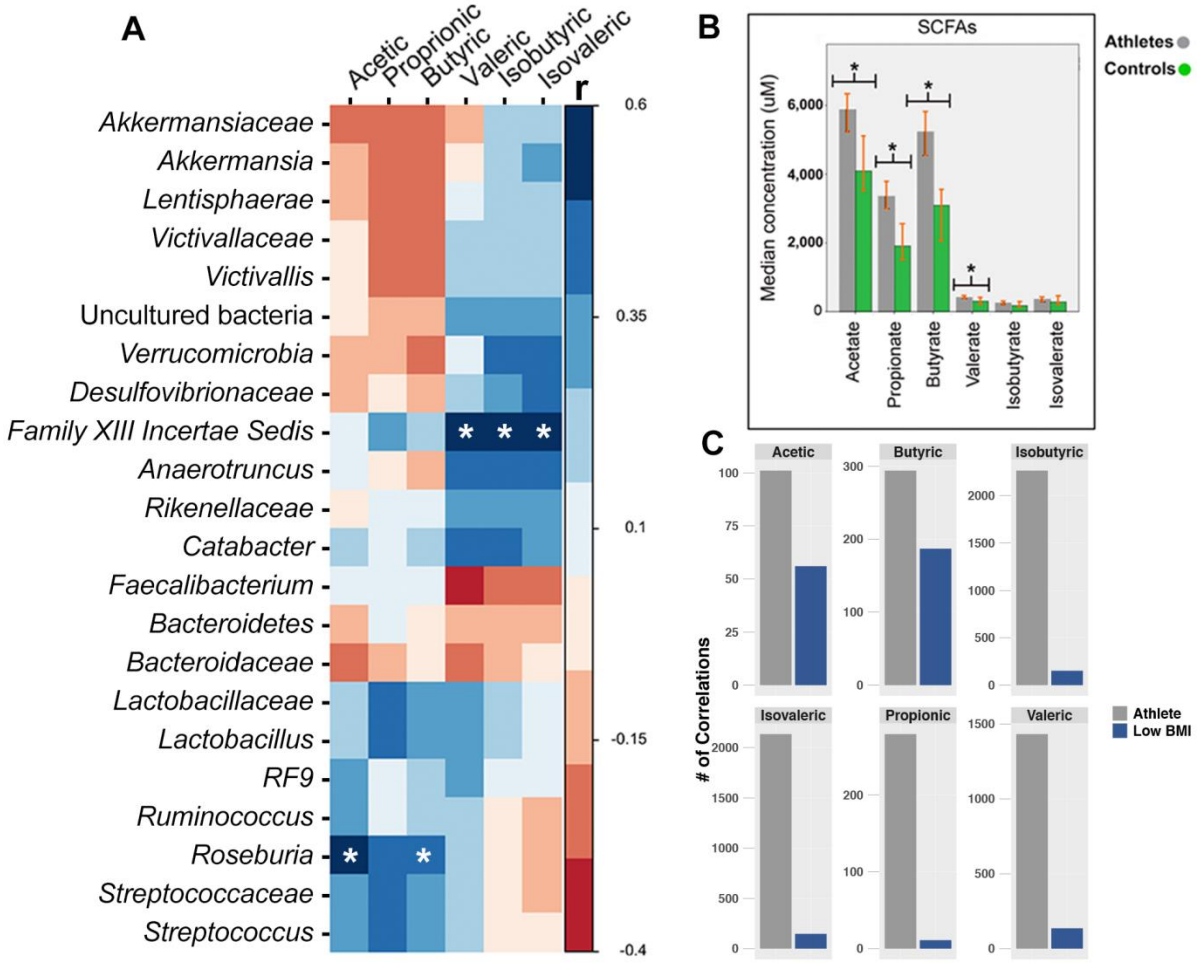


Figure 3.4 | Athletes display a profile of Short Chain Fatty Acids that alters from that of the controls. (A) Heat map of bacterial taxa (family, genus, and species level) that correlate with faecal short-chain fatty acid levels using Spearman's correlation. Cool colours represent positive correlations; hot colours represent negative correlations (r). All taxa shown had a correlation p-value < 0.01. Those marked * represent correlations with a pFDR < 0.01 after Benjamini-Hochberg multiple testing corrections. **(B)** Median concentrations of GC-MS derived faecal short chain fatty acid. Quantitative analysis of SCFAs in faecal samples shows significant increase in measured concentrations of acetate, propionate, butyrate, and valerate in athletes. Error bars represent 95% confidence intervals. * Data statistically significant ($p < 0.05$ after Benjamini-Hochberg corrections). **(C)** Quantification of statistically relevant correlations of metabolic pathways to GC-MS derived faecal SCFA concentrations (μM). Athletes consistently present great numbers of significant associations between pathways and SCFA concentrations.

3.5.3 | Correlating metabonomic and metagenomic results

Correlation analysis between targeted measurements of SCFAs and taxonomic data from 16S rRNA sequencing revealed a number of correlations that remained significant following correction; *Roseburia* was positively correlated with acetate ($p = 0.004$) and butyrate ($p = 0.018$) while *Family XIII Incertae Sedis* was positively correlated with isobutyrate ($p < 0.001$), isovaleric acid ($p < 0.001$) and valeric acid ($p = 0.008$) (Figure 3.4A, Supplementary Table 3.7).

SCFAs were also correlated with pathway relative abundances, with all SCFAs associating with considerably more pathways in the athletes versus the controls (Figure 3.4C). Multiple statistically significant (7,948) ($p < 0.05$) correlations between the metabolic pathways and SCFAs were identified (Supplementary Table 3.8, refer to Appendix A). Two distinct blocks of proportionately discriminant correlations were observed with isobutyric and isovaleric acids, which were more abundant in the athletes while acetic and butyric acids were more proportionately abundant in controls. Correlations of the SCFA concentrations to pathways related to fermentation, biosynthesis, or modification of fatty acids were identified among the numerous other associations (see supplementary Table 3.8 for complete list). Additional correlations of metabolic pathways against well-identified metabolites detected from both faecal water (Figure 3.5A and 3.5C) and urine (Figure 3.5B and 3.5D) presented numerous significant associations (6,186 and 13,412, respectively; $p < 0.05$, data not shown) (). It was also observed that 16 genera correlated with 12 metabolites (Supplementary Table 3.9).

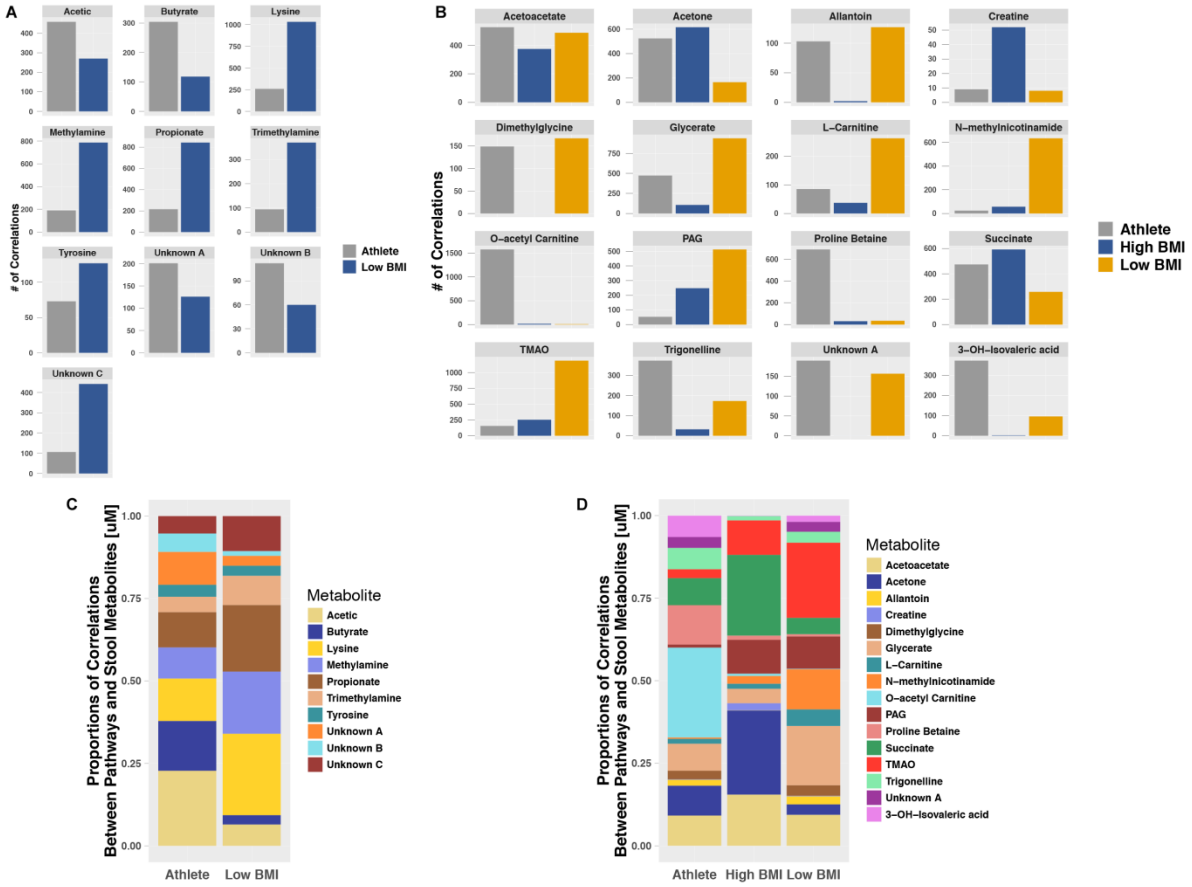


Figure 3.5 | Distinctive association profiles of metabolic pathways to metabolites in athletes and controls. (A) Significant correlations of faecal water derived metabolites and metabolic pathways, represented by number of correlations for each metabolite. **(B)** Urine metabolites significantly correlated to pathways and displayed as number of correlations. **(C)** Significant correlations shown in (**A**) displayed as proportions of total associations. **(D)** Correlations presented in (**B**) given as proportions of total associations. All panels illustrate potential variation in microbial influence on metabolic profiles from the athletes and controls.

3.6 | Discussion

The results confirm enhancement of microbial diversity in athletes compared with controls. Supporting previous insights into the beneficial influence of physical exercise and associated diet on the compositional structure of the gut microbiota [25, 26, 43], this study has extended the paradigm to include links between physical fitness and the functional potential of the gut microbiota and its metabolites. It must be conceded that some athletes, although fit, may not necessarily be more healthy [44].

Athletes have an increased abundance of pathways that—giving an equivalent amount of expression activity—could be exploited by the host for potential health benefit, including biosynthesis of organic cofactors and antibiotics, as well as carbohydrate degradation and secondary metabolite metabolism compared to both control groups [45]. Furthermore, athletes have an enriched profile of SCFAs, previously associated with numerous health benefits and a lean phenotype [46-48]. While interpretation of SCFA data can be difficult as levels represent a combination of SCFA production and host-absorption rates, it is notable that as previously presented, the athletes' diet maintained significantly higher quantities of fibre intake [26]. This, along with an increased number of detected SCFA pathways in the athletes would be conducive to an enhanced rate of SCFA production [49].

It was noted that athletes excreted proportionately higher levels of the metabolite TMAO, an end product metabolite of dietary protein degradation. Elevated TMAO has been observed in patients with cardiovascular disease and atherosclerosis, highlighting a potential downside to increased protein intake [15-17, 22, 50]. However, TMAO is also found in high levels in the urine of Japanese populations [51], who do not have high risk for CVD. Similarly to these populations, the athletes' diet contained a significantly greater

proportion of fish. Our current understanding of the implications of this result remains limited and requires elaboration in future studies. Furthermore, pathway abundance in a metagenome merely reflects functional potential and not necessarily increased expression *in situ*.

Variance of metagenomic composition between athletes and controls was exemplified with unique pathway-pathway correlations between the two groups. Analysis of categorically arranged pathway abundances within the separate cohorts provided additional insight into the previously described dichotomy between the microbiota of athletes and high BMI controls. The two groups displayed distinct structures of functional capacity, separately oriented to operate under the different physiological milieu of the two groups. Notably, from a functional perspective, the microbiota of the low BMI group was more similar to the athletes. The low BMI controls were generally engaged in a modestly active-lifestyle, reflected by their leanness and increased levels of CK. It is speculative but not implausible, that moderate improvements in physical activity, for overweight and obese individuals may confer the beneficial metabolic functions observed within the athlete microbiome.

Dietary contributions to the functional composition of the enteric microbial system are also evident in our study. The relative abundances of pathways related to fundamental metabolic function—amino acid biosynthesis, vitamin biosynthesis, and lipid biosynthesis—were higher on average within the high BMI control group when compared to the athlete group. The mechanisms behind these differences are unclear and might reflect chronic adaptation of the athlete gut microbiome; possibly due to a reduced reliance on the corresponding biosynthetic capacities of their gut microbiota. On the contrary, the athlete microbiome presents a functional capacity that is primed for tissue repair and to harness

energy from the diet with increased capacity for carbohydrate, cell structure and nucleotide biosynthesis, reflecting the significant energy demands and high cell-turnover evident in elite sport.

Remarkably, our examination of pathway correlation to dietary macronutrients and plasma CK, as a biomarker of exercise [52], is suggestive of an impact of physical activity upon the utilization of dietary nutrients by the microbiota of the gut. Comparing athletes to both high and low BMI controls, a greater number of pathways correlating to specific macronutrients with the controls suggests a shift in the dynamics of these varied metabolic functions. The impact of the athletes' increased protein intake compared to both control groups was evident in the metabolomic phenotyping results. By-products of dietary protein metabolism (mostly by microbes) including TMAO, carnitines, trimethylamine, 3-Carboxy-4-methyl-5-propyl-2-furanpropionic acid, and 3-hydroxy-isovaleric acid are all elevated in the athlete cohort. Of particular interest is 3-hydroxy-isovaleric acid (potentially from egg consumption) which has been demonstrated to have efficacy for inhibiting muscle wasting when used in conjunction of physical exercise [53]. The compound is also commonly used as a supplement by athletes to increase exercise-induced gains in muscle size, muscle strength, and lean body mass, reduce exercise-induced muscle damage, and speed recovery from high-intensity exercise [54]. Numerous metabolites associated with muscle turnover—creatine, 3-methylhistidine, and L-valine—and host metabolism—carnitine—are elevated in the athlete groups. Metabolites derived from vitamins and recovery supplements common in professional sports, including glutamine, lysine, 4-pyridoxic acid, and nicotinamide, are also raised in the athlete group.

It is notable that PAG—a microbial conversion product of phenylalanine—has been associated with a lean phenotype, and is increased in the athletes [55]. Furthermore, PAG

positively correlates with the genus *Erysipelotrichaceae Incertae Sedis*, which we have previously noted to be present in relatively higher proportions in the athlete group compared to both control groups. PAG is the strongest biomarker post bariatric surgery, where it is associated with an increase in the relative proportions of Proteobacteria as observed here in the athlete group. Within the SCFAs, two distinct clusters were observed; acetic acid, propionic acid and butyric acid correlate with dietary contributors (fibre and protein), while isobutyric acid, isovaleric acid and valeric acid correlate with microbial diversity. The same clusters are observed when correlating with individual taxa, in support of previously observed links between SCFAs and numerous metabolic benefits and a lean phenotype [46-48].

Our on-going work in this area with non-athletes engaging in a structured exercise regime looks to further explore components of the exercise and diet-microbiome paradigm, which, along with the present study may inform the design of exercise and fitness programs, including diet design in the context of optimizing microbiota functionality for both athletes and the general population.

3.7 | References

[1] A. Harkin, Muscling in on depression, N Engl J Med 371 (24) (2014) 2333-4.

[2] F. B. Benatti, B. K. Pedersen, Exercise as an anti-inflammatory therapy for rheumatic diseases-myokine regulation, Nat Rev Rheumatol 11 (2) (2015) 86-97.

[3] J. A. Hawley, A. Krook, Metabolism: One step forward for exercise, Nat Rev Endocrinol 12 (1) (2016) 7-8.

[4] L. Hoffman-Goetz, N. Pervaiz, N. Packer, J. Guan, Freewheel training decreases pro- and increases anti-inflammatory cytokine expression in mouse intestinal lymphocytes, Brain Behav Immun 24 (7) (2010) 1105-15.

[5] W. Barton, F. Shanahan, P. D. Cotter, O. O'Sullivan, The metabolic role of the microbiota, Clinical Liver Disease 5 (4) (2015) 91-93.

[6] K. L. Szuhany, M. Bugatti, M. W. Otto, A meta-analytic review of the effects of exercise on brain-derived neurotrophic factor, J Psychiatr Res 60 (2015) 56-64.

[7] S. M. Ryan, Y. M. Nolan, Neuroinflammation negatively affects adult hippocampal neurogenesis and cognition: can exercise compensate?, Neurosci Biobehav Rev 61 (2016) 121-31.

[8] E. Johannesson, M. Simren, H. Strid, A. Bajor, R. Sadik, Physical activity improves symptoms in irritable bowel syndrome: a randomized controlled trial, Am J Gastroenterol 106 (5) (2011) 915-22.

[9] T. E. Robsahm, B. Aagnes, A. Hjartaker, H. Langseth, F. I. Bray, I. K. Larsen, Body mass index, physical activity, and colorectal cancer by anatomical subsites: a systematic review and meta-analysis of cohort studies, Eur J Cancer Prev 22 (6) (2013) 492-505.

[10] L. Schwingshackl, B. Missbach, S. Dias, J. Konig, G. Hoffmann, Impact of different training modalities on glycaemic control and blood lipids in patients with type 2 diabetes: a systematic review and network meta-analysis, Diabetologia 57 (9) (2014) 1789-97.

[11] A. Biswas, P. I. Oh, G. E. Faulkner, R. R. Bajaj, M. A. Silver, M. S. Mitchell, D. A. Alter, Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis, Ann Intern Med 162 (2) (2015) 123-32.

[12] R. V. Same, D. I. Feldman, N. Shah, S. S. Martin, M. Al Rifai, M. J. Blaha, G. Graham, H. M. Ahmed, Relationship Between Sedentary Behavior and Cardiovascular Risk, Curr Cardiol Rep 18 (1) (2016) 6.

[13] E. G. Wilmot, C. L. Edwardson, F. A. Achana, M. J. Davies, T. Gorely, L. J. Gray, K. Khunti, T. Yates, S. J. Biddle, Sedentary time in adults and the association with diabetes, cardiovascular disease and death: systematic review and meta-analysis, Diabetologia 55 (11) (2012) 2895-905.

[14] Y. C. Chen, Y. K. Tu, K. C. Huang, P. C. Chen, D. C. Chu, Y. L. Lee, Pathway from central obesity to childhood asthma. Physical fitness and sedentary time are leading factors, Am J Respir Crit Care Med 189 (10) (2014) 1194-203.

[15] R. A. Koeth, Z. Wang, B. S. Levison, J. A. Buffa, E. Org, B. T. Sheehy, E. B. Britt, X. Fu, Y. Wu, L. Li, J. D. Smith, J. A. DiDonato, J. Chen, H. Li, G. D. Wu, J. D. Lewis, M. Warrier, J. M. Brown, R. M. Krauss, W. H. Tang, F. D. Bushman, A. J. Lusis, S. L. Hazen, Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis, Nat Med 19 (5) (2013) 576-85.

[16] W. H. W. Tang, Z. E. Wang, B. S. Levison, R. A. Koeth, E. B. Britt, X. M. Fu, Y. P. Wu, S. L. Hazen, Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk, New England Journal of Medicine 368 (17) (2013) 1575-1584.

[17] W. H. W. Tang, S. L. Hazen, Microbiome, trimethylamine N-oxide, and cardiometabolic disease, Translational Research (2016)

[18] A. Woting, N. Pfeiffer, G. Loh, S. Klaus, M. Blaut, Clostridium ramosum promotes high-fat diet-induced obesity in gnotobiotic mouse models, MBio 5 (5) (2014) e01530-14.

[19] K. M. Utzschneider, M. Kratz, C. J. Damman, M. Hullar, Mechanisms linking the gut microbiome and glucose metabolism., The Journal of clinical endocrinology and metabolism 101 (4) (2016) jc20154251.

[20] P. J. Turnbaugh, R. E. Ley, M. A. Mahowald, V. Magrini, E. R. Mardis, J. I. Gordon, An obesity-associated gut microbiome with increased capacity for energy harvest, Nature 444 (7122) (2006) 1027-31.

[21] N. C. Williams, M. A. Johnson, D. E. Shaw, I. Spendlove, J. Vulevic, G. R. Sharpe, K. A. Hunter, A prebiotic galactooligosaccharide mixture reduces severity of hyperpnoea-induced bronchoconstriction and markers of airway inflammation, Br J Nutr 116 (5) (2016) 798-804.

[22] Z. Wang, E. Klipfell, B. J. Bennett, R. Koeth, B. S. Levison, B. Dugar, A. E. Feldstein, E. B. Britt, X. Fu, Y. M. Chung, Y. Wu, P. Schauer, J. D. Smith, H. Allayee, W. H. Tang, J. A. DiDonato, A. J. Lusis, S. L. Hazen, Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease, Nature 472 (7341) (2011) 57-63.

[23] A. Zhernakova, A. Kurilshikov, M. J. Bonder, E. F. Tigchelaar, M. Schirmer, T. Vatanen, Z. Mujagic, A. V. Vila, G. Falony, S. Vieira-Silva, J. Wang, F. Imhann, E. Brandsma, S. A. Jankipersadsing, M. Joossens, M. C. Cenit, P. Deelen, M. A. Swertz, s. LifeLines cohort, R. K. Weersma, E. J. Feskens, M. G. Netea, D. Gevers, D. Jonkers, L. Franke, Y. S. Aulchenko, C. Huttenhower, J. Raes, M. H. Hofker, R. J. Xavier, C. Wijmenga, J. Fu, Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity, Science 352 (6285) (2016) 565-9.

[24] O. Cronin, M. G. Molloy, F. Shanahan, Exercise, fitness, and the gut, Curr Opin Gastroenterol 32 (2) (2016) 67-73.

[25] M. Estaki, J. Pither, P. Baumeister, J. P. Little, S. K. Gill, S. Ghosh, Z. Ahmadi-Vand, K. R. Marsden, D. L. Gibson, Cardiorespiratory fitness as a predictor of intestinal microbial diversity and distinct metagenomic functions, Microbiome 4 (1) (2016) 42.

[26] S. F. Clarke, E. F. Murphy, O. O'Sullivan, A. J. Lucey, M. Humphreys, A. Hogan, P. Hayes, M. O'Reilly, I. B. Jeffery, R. Wood-Martin, D. M. Kerins, E. Quigley, R. P. Ross, P. W. O'Toole, M. G. Molloy, E. Falvey, F. Shanahan, P. D. Cotter, Exercise and associated dietary extremes impact on gut microbial diversity, Gut 63 (12) (2014) 1913-20.

[27] O. O'Sullivan, O. Cronin, S. F. Clarke, E. F. Murphy, M. G. Molloy, F. Shanahan, P. D. Cotter, Exercise and the microbiota, Gut Microbes 6 (2) (2015) 131-6.

[28] O. Cronin, O. O'Sullivan, W. Barton, P. D. Cotter, M. G. Molloy, F. Shanahan, Gut microbiota: implications for sports and exercise medicine, Br J Sports Med (2017)

[29] A. Rankin, C. O'Donavon, S. M. Madigan, O. O'Sullivan, P. D. Cotter, 'Microbes in sport' – The potential role of the gut microbiota in athlete health and performance, British Journal of Sports Medicine (2017) bjsports-2016-097227.

[30] N. J. Wareham, R. W. Jakes, K. L. Rennie, J. Mitchell, S. Hennings, N. E. Day, Validity and repeatability of the EPIC-Norfolk Physical Activity Questionnaire, International Journal of Epidemiology 31 (1) (2002) 168-74.

[31] S. Abubucker, N. Segata, J. Goll, A. M. Schubert, J. Izard, B. L. Cantarel, B. Rodriguez-Mueller, J. Zucker, M. Thiagarajan, B. Henrissat, O. White, S. T. Kelley, B. Methe, P. D. Schloss, D. Gevers, M. Mitreva, C. Huttenhower, Metabolic reconstruction for metagenomic data and its application to the human microbiome, PLoS Comput Biol 8 (6) (2012) e1002358.

[32] R. Caspi, T. Altman, R. Billington, K. Dreher, H. Foerster, C. A. Fulcher, T. A. Holland, I. M. Keseler, A. Kothari, A. Kubo, M. Krummenacker, M. Latendresse, L. A. Mueller, Q. Ong, S. Paley, P. Subhraveti, D. S. Weaver, D. Weerasinghe, P. Zhang, P. D. Karp, The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases, Nucleic Acids Res 42 (Database issue) (2014) D459-71.

[33] R Development Core Team, R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna, 2012), URL: http:// www. R-project. org (2015)

[34] Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing, Journal of the Royal Statistical Society Series B-Methodological 57 (1) (1995) 289-300.

[35] A. C. Dona, B. Jimenez, H. Schafer, E. Humpfer, M. Spraul, M. R. Lewis, J. T. Pearce, E. Holmes, J. C. Lindon, J. K. Nicholson, Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping, Anal Chem 86 (19) (2014) 9887-94.

[36] R. Garcia-Villalba, J. A. Gimenez-Bastida, M. T. Garcia-Conesa, F. A. Tomas-Barberan, J. Carlos Espin, M. Larrosa, Alternative method for gas chromatography-mass spectrometry analysis of short-chain fatty acids in faecal samples, J Sep Sci 35 (15) (2012) 1906-13.

[37] E. J. Want, I. D. Wilson, H. Gika, G. Theodoridis, R. S. Plumb, J. Shockcor, E. Holmes, J. K. Nicholson, Global metabolic profiling procedures for urine using UPLC-MS, Nat Protoc 5 (6) (2010) 1005-18.

[38] M. H. Sarafian, M. R. Lewis, A. Pechlivanis, S. Ralphs, M. J. McPhail, V. C. Patel, M. E. Dumas, E. Holmes, J. K. Nicholson, Bile acid profiling and quantification in biofluids using ultra-performance liquid chromatography tandem mass spectrometry, Anal Chem 87 (19) (2015) 9662-70.

[39] K. A. Veselkov, L. K. Vingara, P. Masson, S. L. Robinette, E. Want, J. V. Li, R. H. Barton, C. Boursier-Neyret, B. Walther, T. M. Ebbels, I. Pelczer, E. Holmes, J. C. Lindon, J. K. Nicholson, Optimized preprocessing of ultra-performance liquid chromatography/mass spectrometry urinary metabolic profiles for improved information recovery, Anal Chem 83 (15) (2011) 5864-72.

[40] J. M. Posma, I. Garcia-Perez, M. De Iorio, J. C. Lindon, P. Elliott, E. Holmes, T. M. Ebbels, J. K. Nicholson, Subset optimization by reference matching (STORM): an optimized statistical approach for recovery of metabolic biomarker structural information from 1H NMR spectra of biofluids, Anal Chem 84 (24) (2012) 10694-701.

[41] O. Cloarec, M. E. Dumas, A. Craig, R. H. Barton, J. Trygg, J. Hudson, C. Blancher, D. Gauguier, J. C. Lindon, E. Holmes, J. Nicholson, Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H NMR data sets, Anal Chem 77 (5) (2005) 1282-9.

[42] L. W. Sumner, A. Amberg, D. Barrett, M. H. Beale, R. Beger, C. A. Daykin, T. W. Fan, O. Fiehn, R. Goodacre, J. L. Griffin, T. Hankemeier, N. Hardy, J. Harnly, R. Higashi, J. Kopka, A. N. Lane, J. C. Lindon, P. Marriott, A. W. Nicholls, M. D. Reily, J. J. Thaden, M. R. Viant, Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI), Metabolomics 3 (3) (2007) 211-221.

[43] B. A. Petriz, A. P. Castro, J. A. Almeida, C. P. Gomes, G. R. Fernandes, R. H. Kruger, R. W. Pereira, O. L. Franco, Exercise induction of gut microbiota modifications in obese, non-obese and hypertensive rats, BMC Genomics 15 (1) (2014) 511.

[44] P. B. Maffetone, P. B. Laursen, Athletes: Fit but Unhealthy?, Sports Med Open 2 (1) (2016) 24.

[45] H. J. Flint, K. P. Scott, S. H. Duncan, P. Louis, E. Forano, Microbial degradation of complex carbohydrates in the gut, Gut Microbes 3 (4) (2012) 289-306.

[46] A. Koh, F. De Vadder, P. Kovatcheva-Datchary, F. Backhed, From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites, Cell 165 (6) (2016) 1332-1345.

[47] V. K. Ridaura, J. J. Faith, F. E. Rey, J. Cheng, A. E. Duncan, A. L. Kau, N. W. Griffin, V. Lombard, B. Henrissat, J. R. Bain, M. J. Muehlbauer, O. Ilkayeva, C. F. Semenkovich, K. Funai, D. K. Hayashi, B. J. Lyle, M. C. Martini, L. K. Ursell, J. C. Clemente, W. Van Treuren, W. A. Walters, R. Knight, C. B. Newgard, A. C. Heath, J. I. Gordon, Gut microbiota from twins discordant for obesity modulate metabolism in mice, Science 341 (6150) (2013) 1241214.

[48] H. M. Hamer, D. M. Jonkers, A. Bast, S. A. Vanhoutvin, M. A. Fischer, A. Kodde, F. J. Troost, K. Venema, R. J. Brummer, Butyrate modulates oxidative stress in the colonic mucosa of healthy humans, Clin Nutr 28 (1) (2009) 88-93.

[49] G. den Besten, K. van Eunen, A. K. Groen, K. Venema, D. J. Reijngoud, B. M. Bakker, The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism, J Lipid Res 54 (9) (2013) 2325-40.

[50] B. J. Bennett, T. Q. de Aguiar Vallim, Z. Wang, D. M. Shih, Y. Meng, J. Gregory, H. Allayee, R. Lee, M. Graham, R. Crooke, P. A. Edwards, S. L. Hazen, A. J. Lusis, Trimethylamine-N-oxide, a metabolite associated with atherosclerosis, exhibits complex genetic and dietary regulation, Cell Metab 17 (1) (2013) 49-60.

[51] E. Holmes, R. L. Loo, J. Stamler, M. Bictash, I. K. Yap, Q. Chan, T. Ebbels, M. De Iorio, I. J. Brown, K. A. Veselkov, M. L. Daviglus, H. Kesteloot, H. Ueshima, L. Zhao, J. K. Nicholson, P. Elliott, Human metabolic phenotype diversity and its association with diet and blood pressure, Nature 453 (7193) (2008) 396-400.

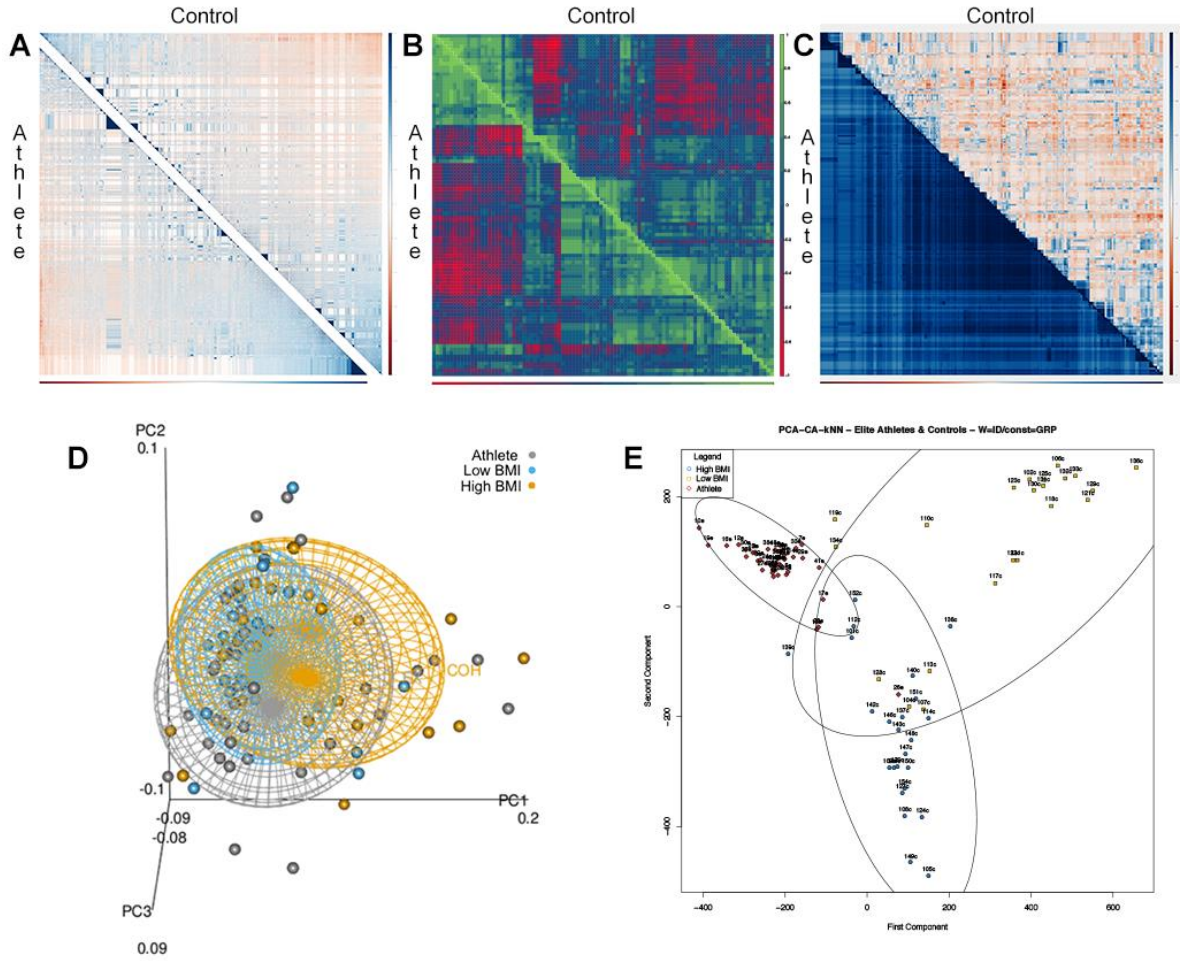
[52] P. Brancaccio, F. M. Limongelli, N. Maffulli, Monitoring of serum enzymes in sport, Br J Sports Med 40 (2) (2006) 96-7.

[53] S. L. Stratton, A. Bogusiewicz, M. M. Mock, N. I. Mock, A. M. Wells, D. M. Mock, Lymphocyte propionyl-CoA carboxylase and its activation by biotin are sensitive indicators of marginal biotin deficiency in humans, Am J Clin Nutr 84 (2) (2006) 384-8.

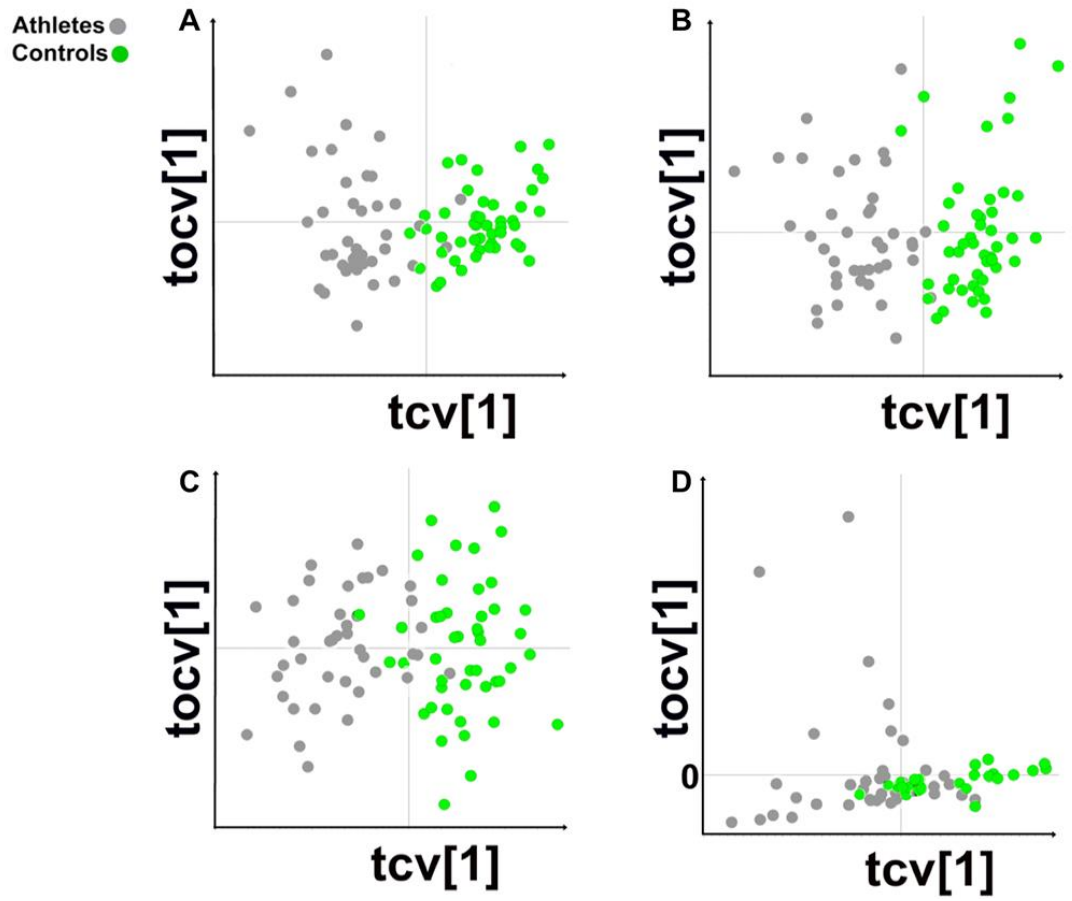
[54] G. J. Wilson, J. M. Wilson, A. H. Manninen, Effects of beta-hydroxy-beta-methylbutyrate (HMB) on exercise performance and body composition across varying levels of age, sex, and training experience: A review, Nutr Metab (Lond) 5 (2008) 1.

[55] E. Holmes, J. V. Li, T. Athanasiou, H. Ashrafian, J. K. Nicholson, Understanding the role of gut microbiome-host metabolic signal disruption in health and disease, Trends Microbiol 19 (7) (2011) 349-59.

3.8 | Supplementary Content



Supplementary Figure 3.1 | Comparison of metabolic pathway composition between groups. Comparative correlation plots of total metabolic pathways **(A)** pathways of statistical significance **(B)** and pathways of which each is represented in at least one subject from each group **(C)**. The respective subsets of metabolic pathways are correlated separately for the athlete group (left and bottom) and all controls (right and top), after which the correlation plots are merged. The joined correlation plots illustrate distinct dynamics of pathway interrelation between athletes and controls. **(D)** Principal coordinate analysis (PCoA) of Bray-Curtis compiled distance matrix of the 98 significant pathway relative abundances. Groups show significant variation from one another (Adonis PERMANOVA P <.001). **(E)** Semi-supervised learning supplemented Principal Component Analysis (PCA-CA-kNN) of total pathway abundance values from all three groups. Outliers were detected through deviation of the defined grouping structure.



Supplementary Figure 3.2 | OPLS-DA cross-validated scores plots comparing athletes to controls. Models are comprised of 1 predictive (tcv[1]) and 1 orthogonal (tocv[1]) principal component. **(A)** UPLC-MS (HILIC, positive mode) urine sample analysis (R2Y=0.85, Q2Y=0.74), **(B)** UPLC-MS (reversed phase, positive mode) urine sample analysis (R2Y=0.83, Q2Y=0.73), **(C)** UPLC-MS (reversed phase, negative mode) urine sample analysis (R2Y=0.83, Q2Y=0.67), **(D)** UPLC-MS (HILIC, positive mode) faecal sample analysis (R2Y=0.65, Q2Y=0.34). Models are comprised of 1 predictive (tcv[1]) and 1 orthogonal (tocv[1]) principal component.

Supplementary Table 3.1 | Statistically significant metagenomic pathways. Pathways found to be statistically significant between the athlete group and both low and high BMI control groups. See appendix A.

Supplementary Table 3.2 | Significant pathways organised according to MetaCyc metabolic classification. Pathways of statistical significance were organised into metabolic classes according to the MetaCyc database. See appendix A

Metabolite	^{1}H chemical shift δ (multiplicity) [a] [Linked with STOCSY / TOCSY]	LoA[b]	Association	Biofluid
TMAO	3.27 (s)	1	↑	Urine
L-Carnitine	3.23 (s), 2.45 (dd), 3.43 (m)	1	↑	Urine
Dimethylglycine	2.93 (s), 3.72 (s)	1	↑	Urine
O-Acetyl Carnitine	3.19 (s), 2.15 (s)	2	↑	Urine
Proline Betaine	3.30 (s), 3.11 (s)	2	↑	Urine
Creatinine	3.05 (s), 4.06 (s)	2	↑	Urine
Acetoacetate	2.29 (s), 3.45 (s)	2	↑	Urine
3-hydroxy-isovaleric acid	1.27 (s), 2.37 (s)	2	↑	Urine
Acetone	2.24 (s)	2	↑	Urine
N-methylnicotinate	4.44 (s), 8.84 (d), 9.12 (s), 8.10 (t)	2	↑	Urine
N-methylnicotinamide	4.486 (s), 8.19 (t), 8.90 (d), 8.96 (d), 9.29 (s)	2	↑	Urine
Phenylacetylglutamine (PAG)	2.11 (m), 2.27 (m), 3.67 (m), 4.19 (m), 7.36 (t), 7.43 (t)	2	↑	Urine
3-Methylhistidine	3.32 (dd), 3.34 (dd), 3.76 (s), 3.98 (dd), 7.16 (s), 8.27 (s)	2	↑	Urine
Beta alanine betaine (3-N-trimethyl-aminoproprionic acid)*	3.12 (s)	-	↑	Urine
Unknown B	6.67 (d)	-	↑	Urine
Unknown C	2.90 (t), 4.42 (s), 4.43 (d)	-	↑	Urine
Unknown D	2.13 (d)	-	↑	Urine
Glycerate	3.70(m), 3.87 (dd)	1	↓	Urine
Allantoin	5.39 (s)	1	↓	Urine
Succinate	2.41 (s)	1	↓	Urine
Propionate	1.06 (t), 2.19 (q)	2	↑	Faeces
Acetate	1.92 (s)	2	↑	Faeces
Butyrate	0.90 (t), 1.56 (m), 2.16 (t)	2	↑	Faeces
Trimethylamine	2.88 (s)	1	↑	Faeces
Lysine	1.46 (m), 1.73 (q), 3.02 (t)	1	↑	Faeces
Methylamine	2.60 (s)	1	↑	Faeces
Beta alanine betaine	3.12 (s)	-	↑	Faeces

Metabolite	^{1}H chemical shift δ (multiplicity) [a] [Linked with STOCSY / TOCSY]	LoA[b]	Association	Biofluid
(3-N-trimethyl-aminoproprionic acid)*				
Unknown E**	4.43 (d), 5.49 (d)	3	↑	Faeces
Glycine	3.56 (s)	2	↓	Faeces
Tyrosine	6.91 (d), 7.20 (d), 3.94 (dd), 3.06 (dd), 3.20 (dd)	2	↓	Faeces

*Tentative identification under the process of spiking the standard.

**overlapping region – multiple candidate saccharides

[a]Multiplicity key is as follows: s=singlet, d=doublet, t=triplet, q=quartet, dd=doublet of doublets, m=multiplet.

[b]The Level of Assignment (LoA) used for the molecules identified by NMR are LoA 1: Identified compound, confirmed by adding the authentic chemical compound to the urine samples (spike-in experiments). LoA 2: ^{1}H and ^{13}C NMR chemical shifts and their multiplicity matched to database or literature to putatively annotate compound. LoA 3: ^{1}H NMR chemical shifts and their multiplicity matched to database or literature to putatively annotate compound.

Supplementary Table 3.3 | List of significantly associated metabolites whose excretion discriminates between athletes and controls when using 1H-NMR metabolic profiling models of urine (Figure 3.2C) and faecal water (Figure 3.2D). Lists the metabolite name, significant chemical shift values and multiplicity, LoA and direction of association (↑ indicates higher excretion in the athletes group, ↓ indicates lower excretion in the athletes group) for urine and faecal biofluids.

Metabolite	LC-platform	m/z	MSMS product ions (5-20eV)	RT	LoA	Association
N-formylanthranilic acid	RP -	164.034	120.04, 136.04, 92.05	4.52	2	↑
Hydantoin-5-propionic acid	RP -	171.040	100.04, 127.06, 128.04	1.43	2	↑
3-Carboxy-4-methyl-5-propyl-2-furanpropionic acid (CMPF)	RP -	239.091	195.10, 151.11	6.16	2	↑
CMPF glucuronide x1	RP -	-	Aglycone: CMPF *(239.09 – see above)*	-	*	↑
Trimetaphosphoric acid	RP -	238.891	96.96, 220.90, 78.96	0.42	2	↑
Glucuronide x1	RP -	-	Aglycone: 239.09	-	*	↑
Acetylcarnitine - C2	RP +	204.125	145.05, 85.03	0.85	2	↑
Propionylcarnitine - C3	RP +	218.139	159.07, 85.03	1.99	2	↑
(Iso)Butyrylcarnitine - C4	RP +	232.155	173.08. 85.03	2.83	2	↑
2-Methyl-butyroylcarnitine - C5	RP +	246.171	187.10, 85.03	3.60	2	↑
C9:1-carnitine	RP +	300.217	241.14, 139.11, 85.03	5.32	2	↑
L-Valine	RP +	118.086	72.08	2.43	2	↑
Nicotinuric acid	RP +	181.061	135.06	3.33	2	↑
Hexanoylcarnitine - C6	HILIC +	260.185	201.11, 85.03	5.72	2	↑
4-Pyridoxic acid	HILIC +	184.062	166.05, 148.04, 138.06	1.14	2	↑
Creatinine	HILIC +	114.067	87.06	6.49	2	↑
Glutamine	RP -	145.061	127.05, 102.95, 101.94, 128.04, 84.05, 109.04	2.89	2	↓
Unknown glucuronides x9	RP -	-	Aglycones: 195.06, 271.07, 285.12, 237.14, 209.12, 143.07, 269.18, 257.17, 217.06	-	*	↓
7-Methylxanthine	RP +	167.059	124.05, 150.03, 142.06	1.98	2	↓

Metabolite	LC-platform	m/z	MSMS product ions (5-20eV)	RT	LoA	Association
Imidazoleacetic acid	RP +	127.051	81.05	0.67	2	↓
Isoquinoline / quinoline	RP +	130.065	103.06	4.79	2	↓
Unknown glucuronides x7	RP +	-	Aglycones: 303.15, 219.06, 297.14, 383.14, 281.17, 239.16, 265.18	-	*	↓

The Level of Assignment (LoA) used for the molecules identified by MS are LoA 1: Identified compound, confirmed by comparison to an authentic chemical reference. LoA 2: MS/MS spectrum matched to database or literature to putatively annotate compound.
*Glucuronides characterised by a neutral loss of 176.03 Da (loss of monodehydrated glucuronic acid) in positive and negative modes. Further characterised by 175 m/z (anhydrous glucuronic acid) + 113 m/z (fragment of glucuronic acid) product ions in negative mode. If unable to unambiguously identify resulting aglycone, aglycone m/z stated.

Supplementary Table 3.4 | List of significantly associated metabolites whose excretion discriminates between athletes and controls when using LC-MS metabolic profiling models of urine (Supplementary Figure 3.2A-D). Lists the metabolite name, liquid chromatography (LC)- platform, mass charge ratio (m/z), MSMS fragments, retention time (RT), level of assignment (LoA) and direction of association (↑ indicates higher excretion in the athletes group, ↓ indicates lower excretion in the athletes group) in urine.

SCFA	Subjects	n	Mean (uM)	SD	Median (uM)	IQR	Mann-Whitney U test p value (2-tailed)	pFDR (BH corrected)
Acetate	Athlete	40	5822.3	1169	5883.5	1628	0.00002	0.00013
	Control	23	4377	1200.7	4092.1	1704		
Propionate	Athlete	40	3419.3	909.8	3359.3	1231	0.00000	0.00000
	Control	23	2077.1	729.4	1910.5	1188		
Butyrate	Athlete	40	5561.5	2449.1	5234.3	2051	0.00003	0.00013
	Control	23	3397	1963.8	3091.2	1639		
Valerate	Athlete	40	429.4	222.1	421.5	237	0.00676	0.01120
	Control	23	283.6	189.9	302.9	260		
Isobutyrate	Athlete	40	277.9	170.4	255.6	175	0.18338	0.21800
	Control	23	217.3	120.6	176.1	180		
Isovalerate	Athlete	40	407.8	270.2	367.6	267	0.36746	0.36500
	Control	23	344.8	207.1	278.5	356		

Supplementary Table 3.5 | Short Chain Fatty Acid concentrations and statistical. analysis. Concentrations of SCFAs detected in faecal water samples. Statistical testing was performed on concentrations, comparing the athlete and control group.

		Lean mass	Protein	Fibre	Simpson diversity	Shannon diversity	PD whole tree diversity
Acetic acid	Spearman's Correlation	0.54					
	Sig. (2-tailed)	0					
Propioninc acid	Spearman's Correlation	0.6	0.45	0.34			
	Sig. (2-tailed)	0	0.0003	0.0069			
Butyric acid	Spearman's Correlation	0.53	0.38	0.39			
	Sig. (2-tailed)	0	0.0024	0.0016			
Isobutyric acid	Spearman's Correlation				0.32	0.4	0.41
	Sig. (2-tailed)				0.0116	0.0015	0.0009
Isovaleric acid	Spearman's Correlation				0.37	0.38	0.35
	Sig. (2-tailed)				0.0035	0.0021	0.0049
Valeric acid	Spearman's Correlation	0.33			0.3	0.37	0.33
	Sig. (2-tailed)	0.0088			0.0169	0.0035	0.0096

Supplementary Table 3.6 | Short Chain Fatty Acid correlation with diet and diversity. SCFA concentrations correlated to measures of gut microbial diversity and participant dietary and body composition measurements.

	Acetic acid			Propionic acid			Butyric acid		
	Spearman Corr	*p value*	*pFDR*	Spearman Corr	*p value*	*pFDR*	Spearman Corr	*p value*	*pFDR*
Rikenellaceae									
Verrucomicrobia									
Desulfovibrionaceae									
Family XIII									
Incertae Sedis									
Enterobacteriaceae									
Anaerotruncus									
Catabacter									
(Uncultured)									
Roseburia	**0.53**	***0***	***0.004***	0.4	*0.001*	*0.109*	**0.48**	***0***	***0.018***
Ruminococcus	0.35	*0.005*	*0.21*						
Veillonellaceae				0.34	*0.006*	*0.205*			
Streptococcus				0.36	*0.004*	*0.183*			
Streptoccaceae				0.42	*0.001*	*0.109*	0.34	*0.006*	*0.248*
Lactobacillus				0.41	*0.001*	*0.109*			
Lactobacillaceae				0.41	*0.001*	*0.109*			
Pseudodubutyrivibrio							0.32	*0.01*	*0.308*
Victivallis				-0.38	*0.002*	*0.138*			
Lentisphaerae				-0.38	*0.002*	*0.138*	-0.33	*0.009*	*0.297*
Faecalibacterium									
RF3				-0.37	*0.003*	*0.142*	-0.33	*0.008*	*0.272*
Akkermansia							-0.35	*0.005*	*0.23*
Akkermansiaceae	-0.32	*0.01*	*0.276*	-0.35	*0.005*	*0.183*	-0.38	*0.002*	*0.135*

Supplementary Table 3.7 (Part A)

	Isobutyric acid			Isovaleric acid			Valeric acid		
	Spearman Corr	*p value*	*pFDR*	Spearman Corr	*p value*	*pFDR*	Spearman Corr	*p value*	*pFDR*
Rikenellaceae	0.33	*0.009*	*0.295*						
Verrucomicrobia	0.37	*0.003*	*0.147*	0.4	*0.001*	*0.109*			
Desulfovibrionaceae	0.38	*0.002*	*0.138*	0.4	*0.001*	*0.109*			
Family XIII									
Incertae Sedis	**0.6**	***0***	***0***	**0.58**	***0***	***0***	**0.51**	***0***	***0.008***
Enterobacteriaceae	0.33	*0.007*	*0.272*	0.34	*0.006*	*0.248*	0.3	*0.019*	*0.401*
Anaerotruncus	0.4	*0.001*	*0.109*	0.41	*0.001*	*0.109*	0.41	*0.001*	*0.109*
Catabacter	0.39	*0.001*	*0.119*	0.38	*0.002*	*0.138*	0.42	*0.001*	*0.109*
(Uncultured)	0.35	*0.005*	*0.211*	0.34	*0.007*	*0.253*			
Roseburia									
Ruminococcus									
Veillonellaceae									
Streptococcus									
Streptoccaceae									
Lactobacillus									
Lactobacillaceae									
Pseudodubutyrivibrio									
Victivallis									
Lentisphaerae									
Faecalibacterium							-0.38	*0.002*	*0.138*
RF3									
Akkermansia									
Akkermansiaceae									

Supplementary Table 3.7 (Part B) | Short Chain Fatty Acid correlation with taxa. Statistically significant correlations between faecal water derived concentrations of SCFAs and 16S identified gut bacteria.

Supplementary Table 3.8 | Short Chain Fatty Acid correlation with metagenomic pathways. List of statistically significant correlations between faecal water derived concentrations of SCFAs and metagenomic pathways. See appendix A.

Genus	Metabolite	p value	PPMC	P correct BH
Erysipelotrichaceae_Incertae_Sedis	Acetoacetate	8.27E-05	4.14E-01	1.32E-03
Howardella	Glycerate	4.08E-04	-3.75E-01	6.52E-03
Anaerotruncus	PAG	6.82E-04	3.61E-01	1.09E-02
Erysipelotrichaceae_Incertae_Sedis	PAG	1.48E-03	3.39E-01	1.19E-02
Parasutterella	N.methylnicotinamide	7.82E-04	3.58E-01	1.25E-02
Succinivibrio	O.Acetyl.Carnitine	8.47E-04	3.55E-01	1.35E-02
Succinivibrio	TMAO	1.95E-03	3.31E-01	1.56E-02
Erysipelotrichaceae_Incertae_Sedis	O.Acetyl.Carnitine	3.59E-03	3.13E-01	1.92E-02
Erysipelotrichaceae_Incertae_Sedis	Unkown.A..3.12.	5.38E-03	2.99E-01	2.15E-02
Phascolarctobacterium	L.Carnitine	2.10E-03	3.29E-01	2.24E-02
Phascolarctobacterium	N.methylnicotinamide	2.80E-03	3.20E-01	2.24E-02
Erysipelotrichaceae_Incertae_Sedis	TMAO	7.36E-03	2.89E-01	2.35E-02
Lactobacillus	O.Acetyl.Carnitine	1.53E-03	3.38E-01	2.45E-02
Escherichia_higella	Glycerate	1.54E-03	-3.38E-01	2.46E-02
Phascolarctobacterium	Glycerate	7.10E-03	-2.90E-01	2.85E-02
Phascolarctobacterium	Acetoacetate	7.12E-03	2.90E-01	2.85E-02
Escherichia_higella	O.Acetyl.Carnitine	4.14E-03	3.08E-01	3.32E-02
Parasutterella	Creatinine	4.91E-03	3.02E-01	3.93E-02
Succinivibrio	L.Carnitine	8.42E-03	2.84E-01	4.49E-02
Anaerotruncus	Acetoacetate	5.66E-03	2.98E-01	4.52E-02
Escherichia_higella	X3.OH.Isovaleric.acid	8.80E-03	2.83E-01	4.69E-02
Mogibacterium	Acetoacetate	6.51E-03	2.93E-01	5.97E-02
Mogibacterium	PAG	7.47E-03	2.88E-01	5.97E-02
Catabacter	Proline.Betaine	4.06E-03	3.09E-01	6.50E-02
Coriobacteriaceae_bacterium_WAL_18889	TMAO	5.28E-03	3.00E-01	8.44E-02
Clostridium	O.Acetyl.Carnitine	5.82E-03	2.97E-01	8.76E-02
Ruminococcus	Acetone	5.48E-03	2.99E-01	8.77E-02
Allisonella	L.Carnitine	6.41E-03	2.93E-01	9.43E-02
Bacteroides	O.Acetyl.Carnitine	8.04E-03	-2.86E-01	9.91E-02
Turicibacter	Proline.Betaine	6.89E-03	-2.91E-01	1.10E-01

Supplementary Table 3.9 | Correlation of gut bacterial taxa with metabolites. Statistically significant correlations of metabolites derived from both urine and faecal water to genera of gut microbiota

.

Chapter 4

EXTENDING THE DIET-EXERCISE-MICROBIOME PARADIGM

4.1 | Abstract

Strong influences have been illustrated for the gut microbiome's role in numerous aspects of human health. Previously we demonstrated that the gut microbiome of high-performance athletes is distinct from comparatively less active but otherwise healthy controls in terms of both composition and functional potential. While the robust differences in the athlete microbiome are intriguing, the dietary habits associated with high-intensity physical activity obscure potential influences on the gut microbiome arising directly from exercise. Here we address the disentangling of exercise from diet in concern to the gut microbiome with an elaboration on dietary influence of metabolomic activity with the intestinal environment via a novel mathematical modelling strategy

In addition to the author, the original material subject of this chapter was contributed to accordingly:

Metabolomic processing and data analysis: Dr. Nicholas C. Penney, Dr. Isabel Garcia Perez, & Professor. Elaine Holmes

Presented material has not been published at the time of this thesis' submission.

4.2 | Introduction

The human gut microbiome has been implicated in a number of different pathologies, including obesity [1, 2], diabetes [3], cardiovascular disease [4], and cancer [5, 6]. Metabolic diseases in particular have been strongly associated with reduced bacterial diversity [7-10]. A high fat, low fibre "Western diet" and sedentary lifestyle are well known risk factors for these morbidities. Conversely, both healthy diet (high fibre, fruit, vegetable content with low sodium, saturated fats and sugars) and exercise are inversely associated with cardiovascular disease (CVD) risk, and there is evidence that they also promote increased gut microbial diversity [11, 12]. Indeed, recent studies have shown that increased physical activity and aerobic fitness may also mediate some of its known health benefits through modulation of gut bacteria [12-15]. Human studies have shown differences in the gut microbiome, including increased diversity, in both habitual exercisers [16, 17] and professional athletes [18-20] compared to less active individuals.

Acquiring a true understanding of the relationship between the gut microbiome, diet and exercise remains elusive since, in humans, the relationship between increased physical activity and microbial changes is confounded by the dietary adaptations that often accompany physical activity. This is further compounded by the lack of accuracy in assessing dietary habits, with an estimated 30-80% misreporting of food intake using traditional self-reported assessment tools [21, 22]. In our previous analysis, elite athletes' increased taxonomic and functional diversity positively correlated with dietary protein intake, in addition to increased physical activity [18, 19].

Here, in order to unravel the contribution of diet and exercise to modulation of the gut microbiota, we apply metabolic profiling combined with a novel mathematical modelling strategy to provide objective evidence of adherence to healthy dietary habits [23],

thereby facilitating the removal of the confounding effects of individuals' dietary misreporting. This enables individuals to be characterised more accurately according to both their physical activity and dietary status. Thus, the current study provides novel functional insights into the distinct influences of each of these factors on host metabolism, the gut microbiome and subsequent gut microbiome–host co-metabolism. Lastly, we further explore the functional and metabolic implications of increased bacterial diversity on the host through multivariate regression analyses.

4.3 | Materials and Methods

4.3.1 | Study population

Described in section 3.4.1, elite professional male athletes (n=40) and healthy but less active controls (n=46) matched for age and gender were enrolled in this study in 2011 as previously described [19]. Approval for the study was granted prior to commencing by the Cork Clinical Research Ethics Committee.

4.3.2 | Acquisition of clinical exercise and dietary data

Activity levels were validated using the EPIC-Norfolk questionnaire [24] and creatine kinase levels. Dietary intake was calculated from food frequency questionnaires (FFQ) administered by a research nutritionist [19].

4.3.3 | Preparation of metagenomic libraries and Sequencing

Detailed previously [18], and in section 2.5 metagenomic library preparation was originally performed with the Illumina Nextera XT DNA Library Preparation Kit (cat. no. FC-131-1096, Illumina Inc., USA) in explicit accordance with the manufacturer's protocol (15031942, Illumina).

Metagenomic libraries were sequenced on the Illumina Hiseq 2500 (chemistry v4.0) next generation sequencing (NGS) platform by Eurofins Genetic Services Ltd. (Ebersberg, Germany), using the high-output run mode for 2 x 125 bp paired-end reads with the addition of a PhiX library (1%) to estimate sequence quality.

4.3.4 | Metagenomic statistical and bioinformatics analysis

Following the approach detailed in section 2.6, raw FASTQ sequence files were quality checked and processed before being subjected to functional profiling by the Human Microbiome Project (HMP) Unified Metabolic Analysis Network (HUMAnN2 V.0.5.0) pipeline [25]. The functional profiling performed by HUMAnN2 composed tabulated files of microbial metabolic pathway abundance and coverage derived from the Metacyc database [26]. Taxonomic profiling was regenerated with the Kraken software package [27]. Microbial pathway and taxonomy data were statistically analysed in the R software environment (V.3.3.2) [28].

Ordination with nonmetric multidimensional scaling (nmds) of relative abundance values for pathways and phylogeny was compiled with the Bray-Curtis index of dissimilarity using the R package Vegan (V.2.3-1), as detailed in section 2.8.3 [29]. Statistical assessment of microbiome variables between participant groups was performed with Wilcoxon signed-rank test of the base stat package in R. All graphical components related to the metagenomic analysis of microbiome pathways and taxonomy were generated using ggplot2 (V.2.2.1) [30]. All presented p values were corrected for multiple comparisons using the Benjamini-Hochberg (BH) False Discovery Rate (pFDR) method [31].

4.3.5 | Metabolic profiling

^{1}H-NMR urine and faecal metabolic profiling analysis and GC-MS targeted SCFA analysis was conducted using established methods [32, 33] as previously described [18]. Additional details can be found in section 2.7

4.3.6 | Objective assessment of dietary intake

Here, we have applied a novel mathematical tool capable of objectively assessing free-living individuals' dietary patterns based on their urine composition, without the need to collect dietary data. The tool is based on a Monte Carlo cross-validated PLS-DA model built from urinary metabolic profiles derived from a highly controlled environment (in-patient randomised controlled clinical trial) to assure healthy participants were fully adherent to four dietary interventions that reflected four different levels of adherence to W.H.O. healthy eating recommendations [23]. The application of this model has been successfully validated in free-living populations.

^{1}H-NMR urinary metabolic profiles from athletes and controls were projected into the MCCV-PLS-DA model to calculate a predicted score for each participant that reflected their adherence to healthy eating. This score was used to cluster participants in the three groups: i) controls with a ≤35% adherence to healthy eating; ii) controls with a ≥100% adherence to healthy eating; iii) athletes with a ≥100% adherence to healthy eating. These groups were compared using OPLS-DA.

4.3.7 | Statistical Analysis

^{1}H-NMR spectra were phased and digitised over the range δ0·5–9·5 and imported into MATLAB to undergo median fold change normalisation [34]. Multivariate statistical analysis was performed using SIMCA 14.1 (Umetrics) [34]. The three comparative groups previously defined by extreme dietary habits were modelled using Orthogonal Partial Least

Squares – Discriminant Analysis (OPLS-DA): i) controls with a ≤35% adherence to healthy eating versus controls with a ≥100% adherence to healthy eating (Model 1); ii) controls with a ≥100% adherence to healthy eating versus athletes with a ≥100% adherence to healthy eating (Model 2); iii) controls with a ≤35% adherence to healthy eating versus athletes with a ≥100% adherence to healthy eating (Model 3). The OPLS-DA models in this study were established based on one predictive component and one orthogonal component. Unit variance scaling was applied to ^{1}H-NMR spectral data. The fit and predictability of the models obtained were determined by the R^2Y and Q^2Y values, respectively. Significant metabolites differentiating between exercise and/or diet groups were obtained from ^{1}H-NMR OPLS-DA models after investigating correlations with correlation coefficients values higher than 0.4.

Univariate statistical analysis (Mann-Whitney U test) was used to identify discriminatory metabolites in SCFA quantitative datasets. P values were adjusted for multiple testing with the BH False Discovery Rate method (pFDR).

Linear Regression analysis of ^{1}H-NMR global spectra was performed against Shannon, Simpson, and Whole tree phylogenetic diversity (PD) diversity indices, corrected for confounding factors (age, lean mass and fat mass), using an in-house MATLAB (release 2104a) script. Metabolites with adjusted p-values <0.01, after correcting for multiple testing using the BH False Discovery Rate method, were considered significant and were subsequently visualised in a Manhattan plot.

Finally, heatmaps showing significant Spearman's correlations between datasets were generated using the ComplexHeatmap script in R [35]. Spearman's correlations were first calculated in MATLAB, with p-values corrected for multiple testing using the BH False Discovery Rate (pFDR) method. Correlations between metabolite-metabolite data sets with

a pFDR <0.05 were included. Due to the large number of variables, a false discovery rate of 1% was set for correlations including microbiome taxa and pathway datasets. Hierarchical clustering of correlations was performed using Euclidean distances.

4.3.8 | Metabolite identification

A combination of data-driven strategies such as such as SubseT Optimization by Reference Matching (STORM) [36] and Statistical TOtal Correlation SpectroscopY (STOCSY) [37] and analytical identification strategies were used to aid structural identification of significant discriminatory metabolites. Specifically, a catalogue of 1D ^{1}H-NMR sequence with water pre-saturation and 2D NMR experiments such as J-Resolved spectroscopy, ^{1}H–^{1}H TOtal Correlation SpectroscopY (TOCSY), ^{1}H–^{1}H COrrelation SpectroscopY (COSY), ^{1}H–^{13}C Hetero-nuclear Single Quantum Coherence (HSQC) and ^{1}H–^{13}C Hetero-nuclear Multiple-Bond Correlation (HMBC) spectroscopy were performed. Finally, metabolites were confirmed by in situ spiking experiments using authentic chemical standards. See section 2.7.8 for additional details.

4.3.9 | Metabolite quantification

Semi-quantification of identified metabolites was performed using an in-house MATLAB script to annotate maximal intensity of median fold change normalised spectral peaks. The highest intensity spectral peak was identified for each measured metabolite and precisely quantified for each individual. Manual annotation enabled precise marking of peaks, allowing for variations in chemical shift in individual samples. Where spectral peaks overlapped with other metabolites the highest intensity unambiguous spectral region was used.

4.4 | Results

4.4.1 | Model generation through clustering of individuals according to objective dietary healthiness scores and exercise levels.

Scores corresponding to participants' adherence to healthy eating habits were generated from ^{1}H-NMR urinary profiles using a validated metabolic profiling dietary assessment tool [23] (see Figure 4.1). As expected, the non-professional athlete controls had diverse dietary healthiness scores, whilst professional athletes were predominantly assigned to scores with ≥100% adherence to healthy eating, reflecting the known association between habitual exercise and healthy eating [38]. Individuals were then clustered according to their score for dietary healthiness to construct comparative groups, with subsequent comparisons made between the groups (see Figure 4.1): i) Model 1 was constructed comparing controls with a ≤35% adherence to healthy eating versus controls with a ≥100% adherence to healthy eating; ii) Model 2 compares sedentary controls with a ≥100% adherence to healthy eating versus athletes with a ≥100% adherence to healthy eating; iii) Model 3 compares controls with a ≤35% adherence to healthy eating versus athletes with a ≥100% adherence to healthy eating. The models were generated to assess both the combined [Model 3] and discrete effects of a healthy diet [Model 1] and exercise [Model 2] on host metabolism, the gut microbiome and subsequent gut microbiome – host co-metabolism.

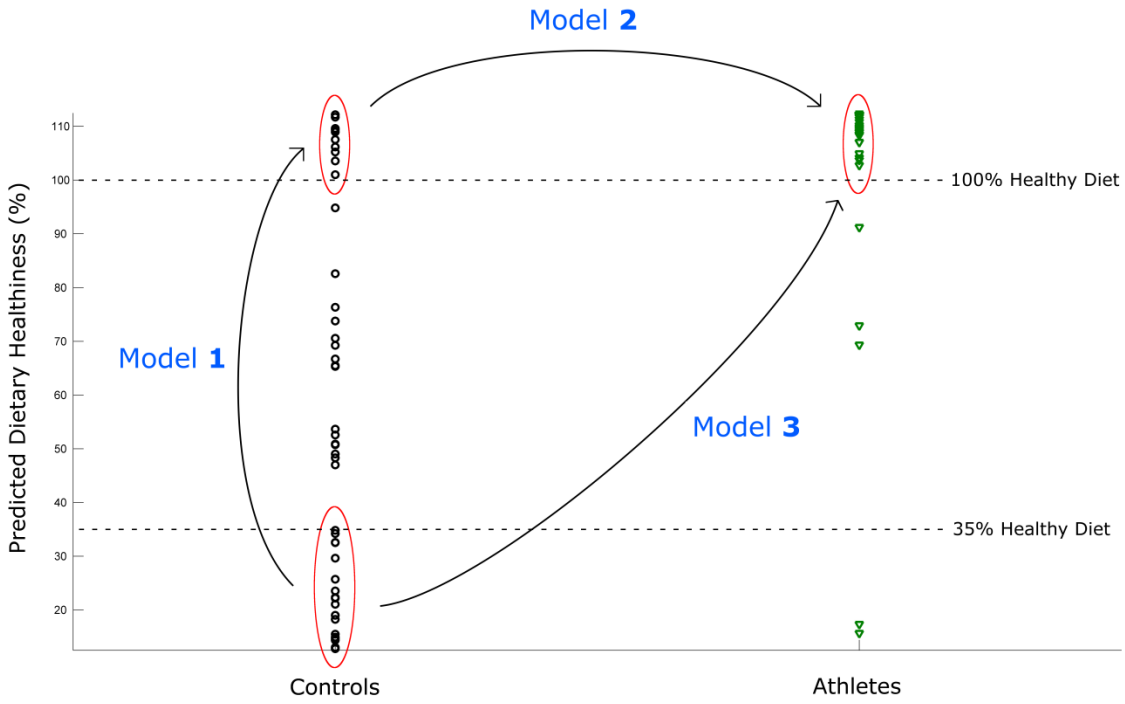


Figure 4.1 | Clustering of individuals according to exercise status and predicted adherence to healthy dietary habits, calculated from urinary 1H-NMR spectra of individuals, to form comparative models. Model 1 (Dietary Healthiness): controls with a ≤35% adherence to healthy diet versus controls with a ≥100% adherence to healthy diet. Model 2 (Exercise): controls with a ≥100% adherence to healthy diet versus athletes with a ≥100% adherence to a healthy diet. Model 3 (Combined Healthy Diet and Exercise): controls with a ≤35% adherence to a healthy diet versus athletes with a ≥100% adherence to a healthy diet.

4.4.2 | Metabolic characterisation of athletes and less active controls stratified by predicted dietary healthiness.

Based on the urine spectral data, three robust Orthogonal Partial Least Squares Discriminant Analysis (OPLS-DA) models were obtained for assessment of the metabolic impact of diet, exercise and the combination of diet and exercise. The model to investigate the combined effect of diet healthiness and exercise (Model 3a) was the strongest (R^2Y=0.94, Q^2Y=0.74), with robust models also defining the impact of diet healthiness (R^2Y=0.97, Q^2Y=0.55; Model 1a) or exercise alone (R^2Y=0.92, Q^2Y=0.45; Model 2a). With regard to the faecal profiles, the strongest model was that defining the effect of exercise (R^2Y=0.92, Q^2Y=0.45, Model 2b) whilst the combination of diet and exercise healthiness produced a relatively robust model (R^2Y=0.91, Q^2Y=0.39; Model 3b) with diet showing the least, but nevertheless significant, impact (R^2Y=0.95, Q^2Y=0.21; Model 1b).

The loadings of the pairwise OPLS-DA models were used to identify metabolites discriminating between the two classes in each model. Metabolites discriminating between classes are listed in Tables 4.1 (urine) and 4.2 (faeces). As expected, in Model 1 (Healthy Diet effect) urinary metabolites with well-known dietary associations such as proline-betaine (associated with citrus intake) and hippurate (associated with the consumption of fruits and vegetables) were significantly higher in the group with healthier eating scores, whereas markers of oxidative stress (2 hydroxybutyrate) and anaerobic metabolism (lactate) were found in lower concentrations within the healthier eating group. Faecal metabolite patterns associated with healthier diet scores included lower concentrations of 2-aminobutyrate and 2-hydroxybutyrate and higher levels of markers of cruciferous vegetable intake (SMCSO derivatives). In the Exercise effect model (Model 2), the athlete group was characterised by a number of urinary metabolites derived from the gut microbiome, such as phenylacetylglu-

tamine (PAG) and 3-indoxyl sulphate, as well as the ketone body acetoacetate and markers of red meat intake and fatty-acid metabolism (carnitine, o-acetyl carnitine), while markers of oxidative stress (allantoin) and TCA cycle intermediates (succinate) were found to be higher in the control group. Analysis of faecal metabolites revealed higher concentrations of the SCFAs acetate, propionate, butyrate and valerate and lower levels of the amino acids glycine, phenylalanine and tyrosine in the athlete group compared to controls consuming a comparatively healthy diet. Model 3 (Combined effect) consisted predominantly of metabolites derived from models 1 and 2. Healthy eating, exercise, and the combined effect of both (Models 1-3), were all characterised by higher markers of choline metabolism; with higher urinary TMAO. In the exercise effect model (Model 2), this corresponded with higher faecal methylamine, and trimethylamine and lower dimethylamine in athletes relative to controls.

4.4.3 | Quantitative analysis of short chain fatty acids and organic acids.

Targeted GC-MS of SCFA level in faeces revealed significantly higher levels of acetate ($pFDR<0.001$), propionate ($pFDR<0.001$) and butyrate ($pFDR=0.006$) in healthy eating athletes relative to healthy eating controls (Model 2 – Exercise effect). In addition to propionate ($pFDR=0.004$) and butyrate ($pFDR=0.019$), valerate ($pFDR=0.043$) was higher in healthy eating athletes relative to controls with unhealthy eating, while acetate was significantly higher before but not after BH multiple testing corrections ($pFDR=0.056$) (Model 3 – Combined effect). There were no significant differences in the diet healthiness model (Model 1) and no differences in isobutyrate or isovalerate concentrations in any model. (Figure 4.2)

Targeted analysis of ten organic acids in urine by GC-MS demonstrated lower levels of lactate ($pFDR=0.002$) and 2-hydroxybutyrate ($pFDR=0.014$) in controls with ≥100%

adherence to healthy eating versus controls with a ≤35% adherence to healthy eating (Model 1 – Healthy diet effect). Levels of 2-Methylbutyrate were lower in healthy eating athletes versus healthy eating controls (pFDR=0.043; Model 2 – Exercise effect). In contrast to the faecal SCFA model, acetate (pFDR=0.006), propionate (pFDR=0.003), 2-methylbutyrate (pFDR=0.001), isovalerate (pFDR=0.001) and lactate (pFDR<0.001) were all lower in athletes with ≥100% adherence to healthy eating versus controls with ≤35% adherence to healthy eating (Model 3 – Combined Effect).

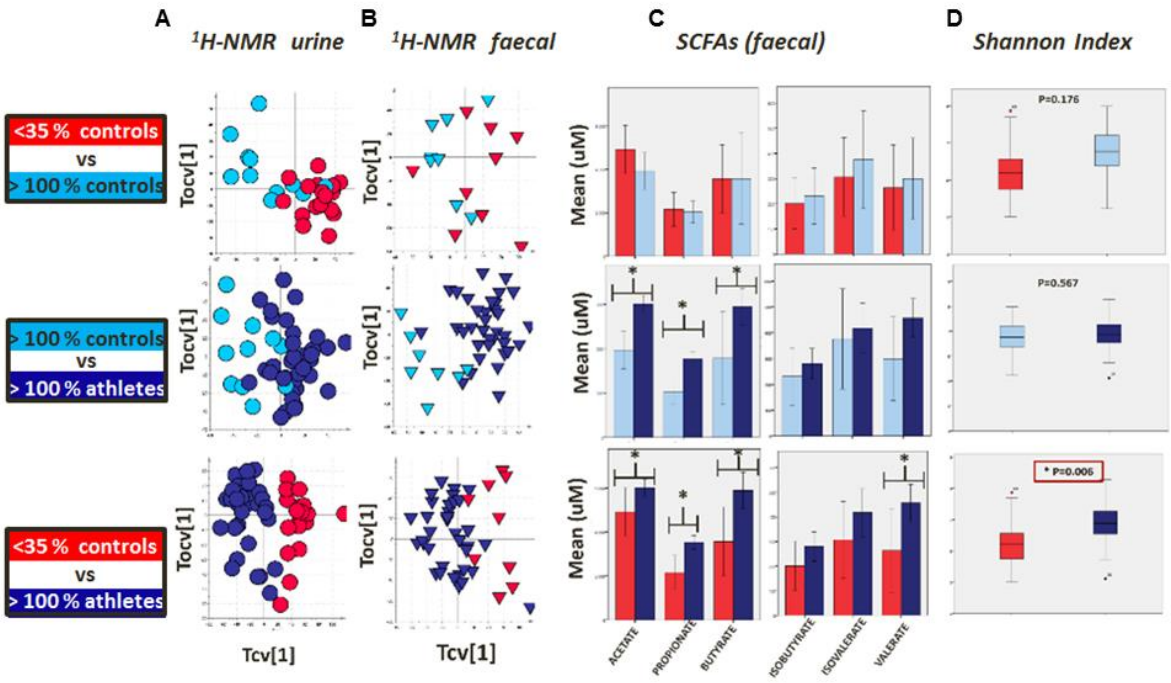


Figure 4.2 | ¹H-NMR, SCFA and bacterial diversity results. Cross-validated OPLS-DA score plots of (**A**) 1H-NMR urinary profiles, (**B**) 1H-NMR faecal profiles. These OPLS models demonstrate the divergence of metabolomic profiles between the participants included in the different dietary models. Bar charts of (**C**) mean SCFA levels measured through quantitative GC-MS and (**D**) mean Shannon diversity index levels, with 95% CI shown. Detected concentrations of SCFAs show that levels of the compounds do not significantly differ between controls, while athletes have significantly elevated levels when compared to both control groups. Significant results (pFDR<0.05) marked with *. Rows 1, 2, and 3 represent Model 1 (Diet), Model 2 (Exercise) and Model 3 (Combined Diet Healthiness and Exercise) respectively.

4.4.4 | Bacterial diversity

No difference in gut bacterial diversity was found between classes in any of the Models. Each class had equivalent levels of α-diversity (Shannon H-index; Supplementary Figure 4.1). Species diversity between participants (β-diversity) as delineated by high-level phylogenetic classification (i.e. Bacteria) demonstrated little class variability for Archaea species, while Bacteria and virus species showed greater class separation (Figure 4.3). Metabolic pathway models generated via metagenomic sequencing, revealed moderate variability of β-diversity between classes for subsets of general (e.g. glycogen biosynthesis) and species-specific (e.g. glycogen biosynthesis in *R. intestinalis*) pathways, while corresponding α-diversity was slightly varied (Figure 4.4).

Linear regression analysis of ^{1}H-NMR faecal water global spectra against Shannon, Simpson and PD bacterial diversity indices, corrected for confounding factors (age, lean mass and fat mass), revealed multiple metabolites associated with bacterial diversity (see Figure 4.5 and Supplementary Table 4.2). Conversely, no urinary ^{1}H-NMR global spectra were found to significantly correlate with measures of gut bacterial diversity. Faecal metabolites associated with increased bacterial diversity include elevated levels of short/medium chain FAs valerate and caproate, branched SCFAs isobutyrate, isovalerate and 2-methylbutyrate, branched chain amino acid (BCAA) degradation products 2-oxoisocaproate and 2-oxoisovalerate as well as products of phenylalanine (phenylacetate and 3-phenylpropionate), choline (dimethylamine), proline (2-methylproline) and uracil (ureidopropionate) metabolism. Increased diversity was also associated with lower glucose levels in addition to reduced isoleucine, asparagine and histidine.

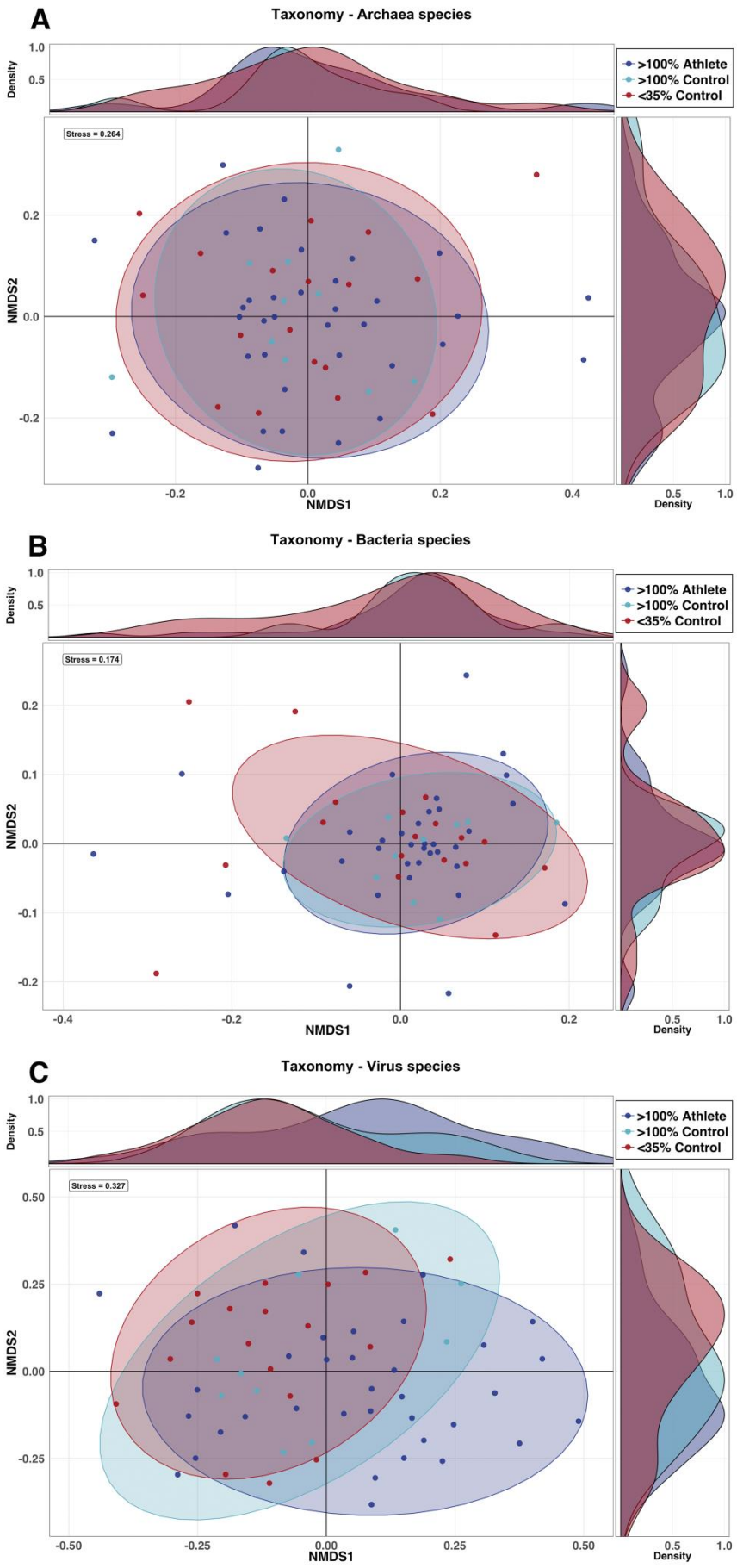


Figure 4.3 | Taxonomic β-diversity of participant classes.

Figure 4.3 continued | Taxonomic β-diversity of participant classes. Ordination of classes corresponding to species from Archaea (**A**), Bacteria (**B**), and viruses (**C**) demonstrates little alteration between classes for Archaea species, while healthy athletes and healthy controls are more similar for bacterial species. Participant clustering with virus species shows a divergence of the athletes from controls. Ordination performed with non-metric multidimensional scaling (NMDS). Density plots along the axes of all panels display the concentrations of data points, and were generated from kernel density estimates and scaled to a maximum value of 1.

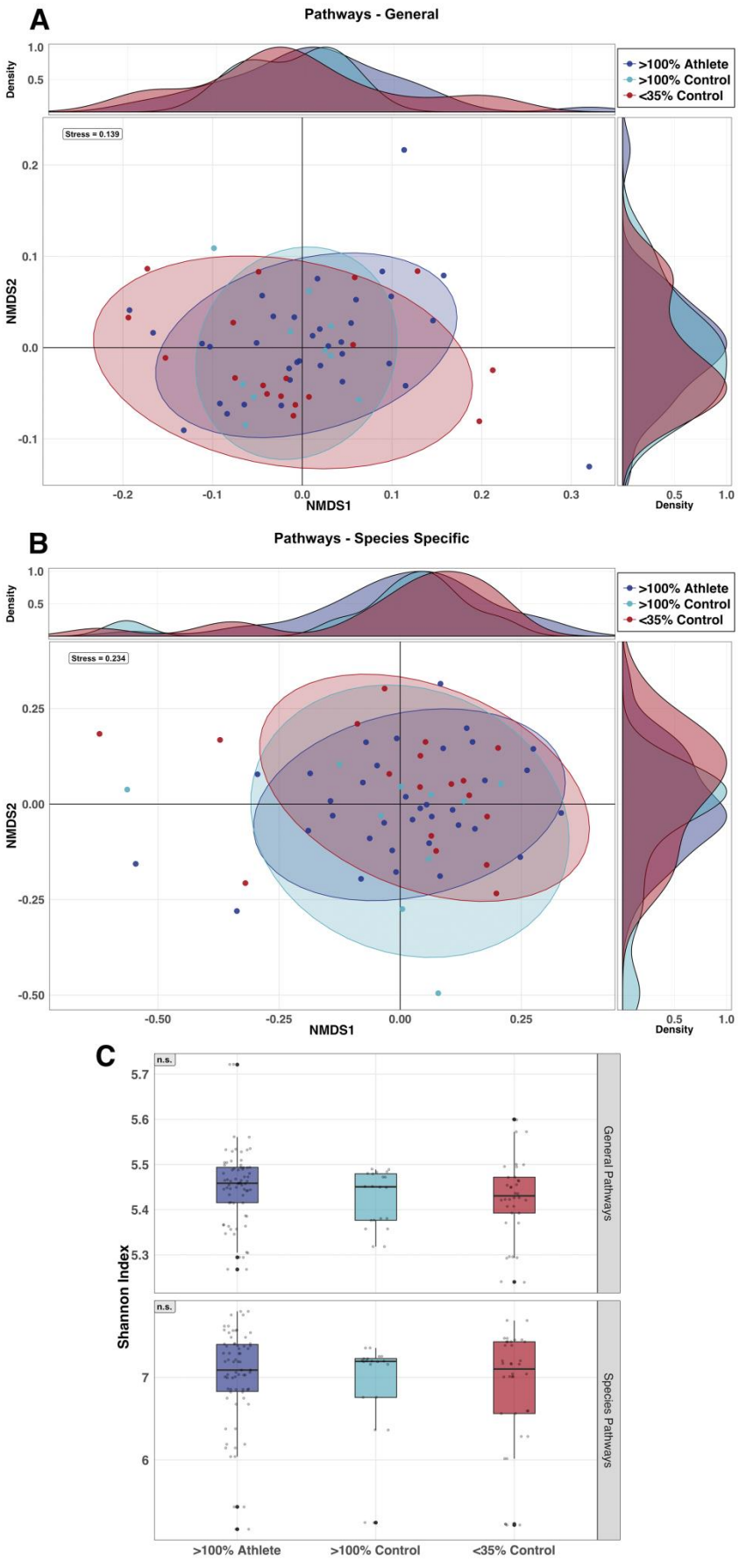


Figure 4.4 | Diversity of metabolic pathways between classes.

Figure 4.4 continued | Diversity of metabolic pathways between classes. β-diversity of metabolic pathways, separated according to general (**A**) and species-specific (**B**) reveals modest separation between the classes. α-diversity of metabolic pathways (**C**), presented as Shannon H-index demonstrate, slight variation in the participants. Ordination performed with non-metric multidimensional scaling (NMDS). Density plots along the axes of all panels display the concentrations of data points, and were generated from kernel density estimates and scaled to a maximum value of 1.

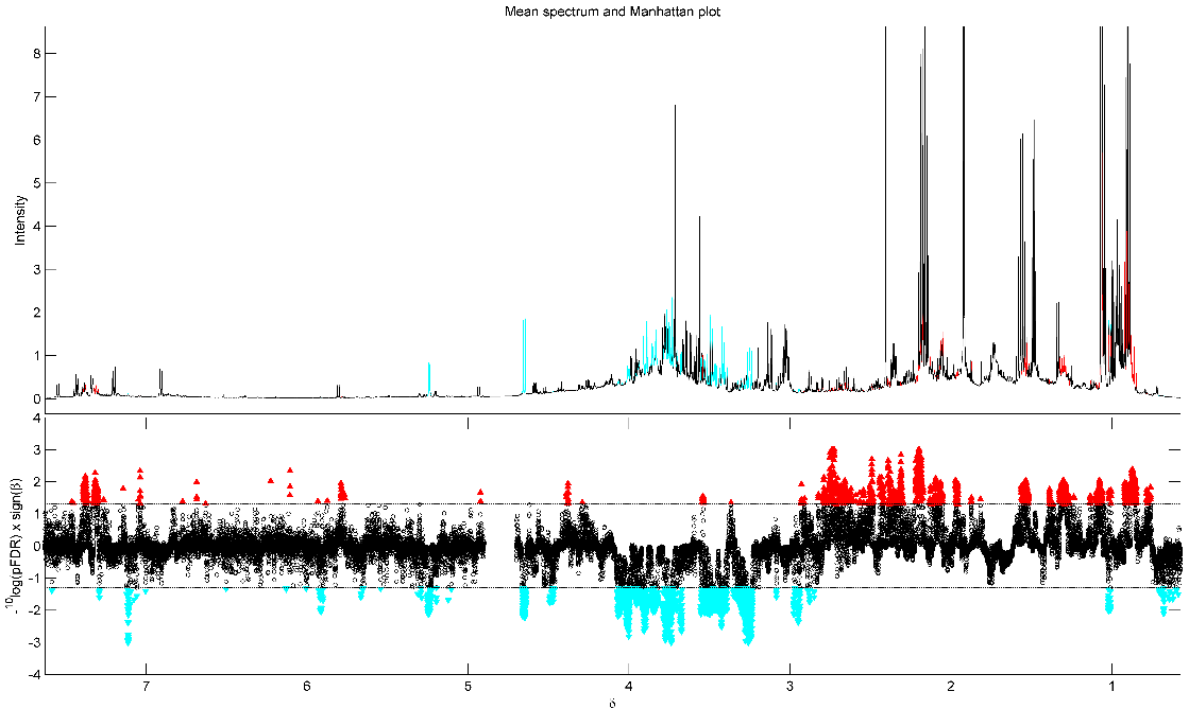


Figure 4.5 | Linear regression of faecal water ^{1}H-NMR spectra against bacterial diversity (Shannon index), corrected for confounding factors (age, lean mass and fat mass). Metabolites with a pFDR of <0.01 were considered significant and are highlighted in the Manhattan plot and mean ^{1}H-NMR spectrum. Spectra highlighted red increase with bacterial diversity, spectra highlighted blue decrease with increased bacterial diversity. Together, these associations illustrate the broad range of metabolites that may be influenced by the microbiome of participants. Further detail of specific metabolites identified with this approach are presented in subsequent sections of this chapter.

4.4.5 | Diet – Gut microbiome – Gut metabolome – Host metabolome interactions

A total of 47 faecal and 27 urinary metabolites of interest, based on discriminatory capacity from Models 1-3 and/or metabolites associated with bacterial diversity, were quantified. Gut metabolome–host metabolome interactions (faecal vs urinary metabolites respectively) were derived from Spearman's correlations between the two datasets (see Figure 4.7). Diet – Gut bacterial metabolome and Diet – host metabolome relationships were derived from Spearman's correlations between dietary food groups calculated from food frequency questionnaires and faecal and urinary metabolite levels (see Supplementary Figure 4.2). Furthermore, Spearman's correlations of metabolite levels against bacterial species (Supplementary Figure 4.3) and microbial metabolic superpathway relative abundances (Figure 4.7) revealed numerous significant correlations with one or more urinary or faecal metabolites (Species pFDR<0.01 = 257, Metabolic Superpathways pFDR<0.01 = 58). Similarly, relative abundance of 206 microbial metabolic superpathways was significantly correlated to measures of bacterial diversity (pFDR<0.01, Supplementary Figure 4.4). Lastly, correlations within faecal-faecal and urinary-urinary metabolite datasets were used to highlight known or unknown metabolic pathways (see Supplementary Figures 4.5 & 4.6).

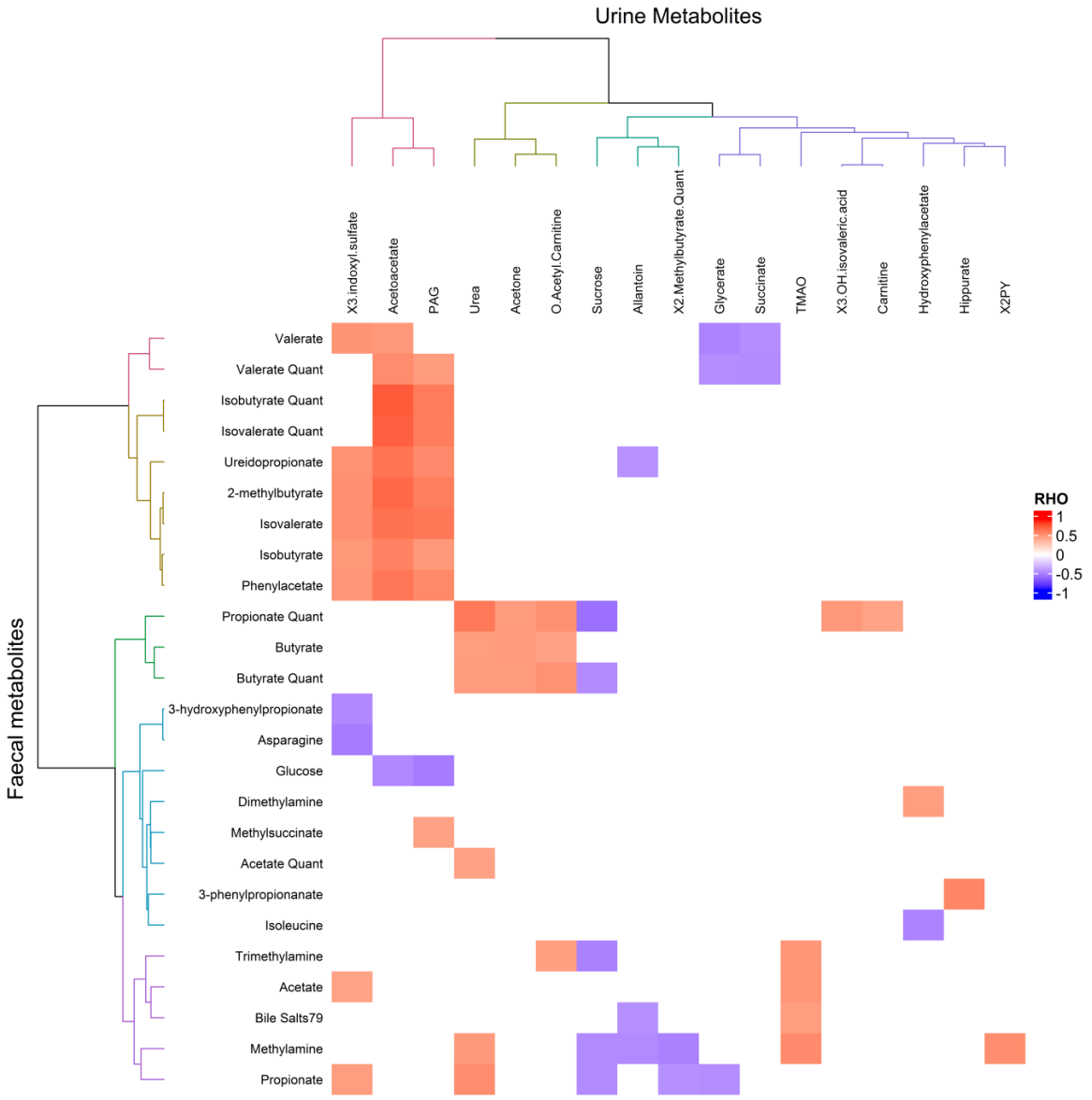


Figure 4.6 | Faecal – Urinary metabolic interactions. Demonstrating potential metabolic interactions of compounds derived specifically from host metabolism and products of the gut environment, significant Spearman's correlations (pFDR <0.05) between faecal and urinary metabolic datasets are shown, shaded according to strength of correlation coefficient (Rho). Correlations are clustered according to Euclidean distances.

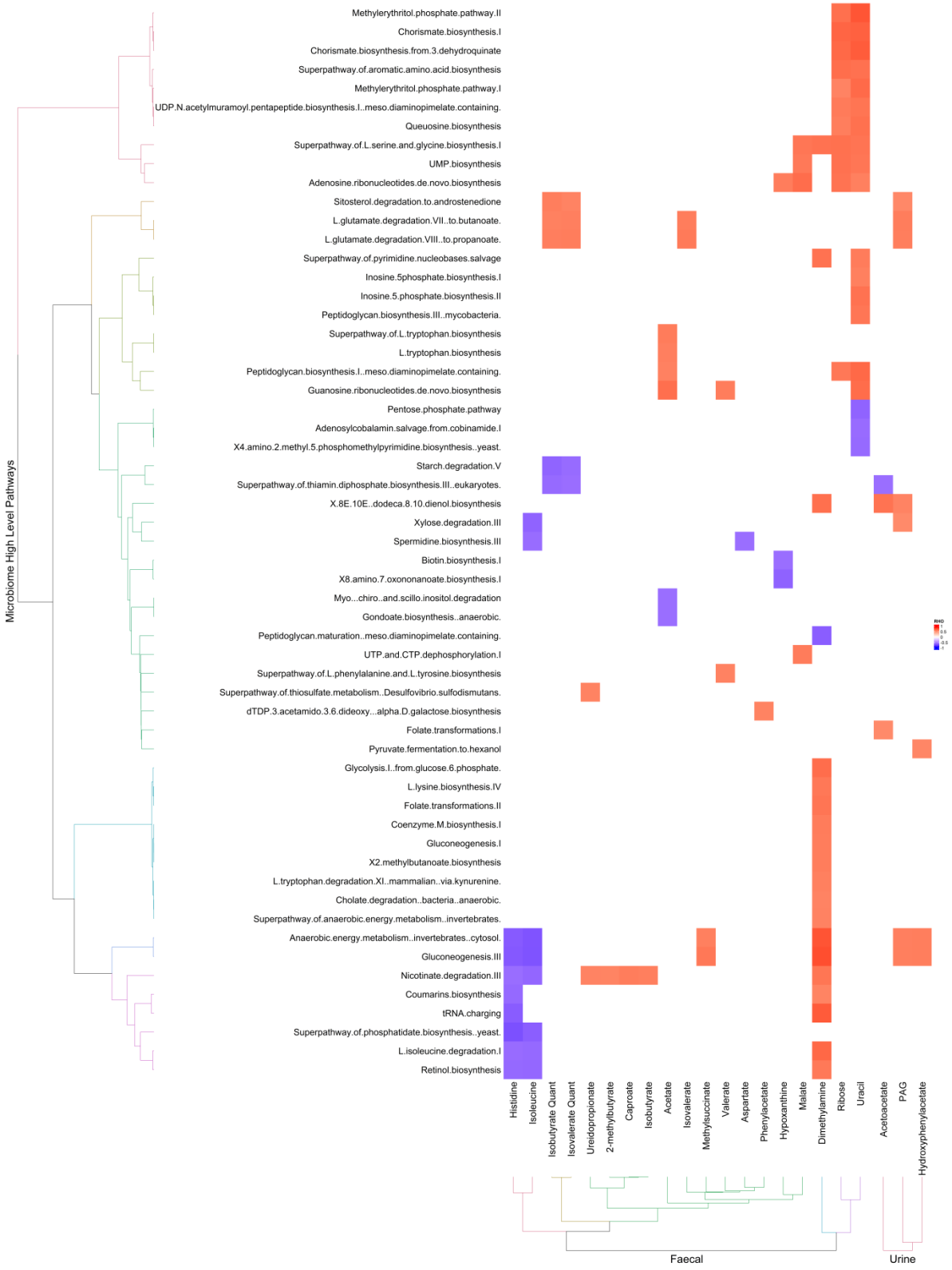


Figure 4.7 | Microbial metabolic superpathway–metabolite interactions.

Figure 4.7 continued | Microbial metabolic superpathway–metabolite interactions. Significant Spearman's correlations (pFDR <0.01) between microbial metabolic super-pathways and both faecal and urinary metabolic datasets are shown, presenting the microbial pathways detected with sequencing that may contribute to the generation of metabolomic features. The data is shaded according to strength of correlation coefficient (Rho), with positive values coloured red, and negative values coloured blue. Correlations are clustered according to Euclidean distances.

4.5 | Discussion & Biological Interpretation

Each of the models pertaining to diet healthiness, exercise or a combination of both displayed a unique metabolic, taxonomic, and functional microbial signature, with a significant proportion of identified metabolites discriminating between the various pairwise comparisons resulting from gut-host co-metabolism. Bacterial diversity was associated with a combination of increased diet healthiness and exercise and was correlated with a distinct array of microbially derived metabolites.

4.5.1 | Metabolic phenotype associated with diet and exercise

As expected, the effect of adherence to healthy dietary habits was associated with higher concentrations of markers of fruit and vegetable intake in the form of proline-betaine, formate and hippurate in addition to lower sugar levels [23, 39].

While investigating changes related to high physical activity, we found higher levels of O-acetyl carnitine, a metabolite associated with the intake of animal protein. Indeed, protein consumption was higher in the athlete group, and included whey protein supplementation. However, higher O-acetyl carnitine may also be due to increased fatty acid mobilisation and oxidation in this group [40]. Additionally, there were lower concentrations of urinary succinate in athletes, likely due to increased systemic utilisation through the TCA cycle. Controls assigned healthy diet scores were characterised by lower urinary levels of lactate and 2-hydroxybutyrate, markers of anaerobic respiration and oxidative stress, respectively. Oxidative stress raises the rate of hepatic glutathione synthesis, causing homocysteine to be metabolised through the transsulfuration pathway to replace depleted cysteine, with 2-hydroxybutyrate produced as a by-product [39].

4.5.2 | Evidence of altered gut microbial – host co-metabolism

Tryptophan Metabolism

We find lower levels of methylnicotinamide (NMND) in subjects assigned healthy eating scores (Model 1). In humans, the amino acid tryptophan is predominantly metabolised through the kynurenine pathway. The resulting nicotinamide is methylated by nicotinamide *N*-methyltransferase (NNMT), to form *N*-methylnicotinamide (NMND), which is then further metabolised to *N*-methyl-2-pyridone-5-carboxamide (2PY) [41]. Nicotinamide serves an important role in energy regulation. Indeed, NNMT knockout in mice, resulting in higher nicotinamide and reduced NMND and 2PY, is protective of diet induced obesity [42]. Thus, the lower levels of NMND seen here suggest increased energy expenditure in healthy eaters.

Conversely, in the exercise effect model we find higher levels of 2PY associated with physical activity (Model 2). In addition, there is increased gut microbial metabolism of tryptophan through the indole pathway leading to increased 3 indoxyl-sulfate (3IS). This is a similar metabolic signature to that seen in malnourished children [41]. In this case, however, increased 2PY is more readily explained through indoleamine 2,3-dioxygenase (IDO) mediated activation of the kynurenine pathway due to increased inflammation [19]. Increased 3IS is likely due to a shift towards proteolytic fermentation of increased colonic protein by gut bacteria due to higher protein intake in healthy eating athletes relative to healthy eating controls (median intake 240g/day vs. 104 g/day, $p<0.001$).

Phenylalanine Metabolism

Higher physical activity (Model 2) and the combined effect of healthy diet and exercise (Model 3) result in higher levels of urinary phenylacetylglutamine (PAG), produced through gut microbial conversion of phenylalanine to phenylacetate and subsequent conversion to PAG in the liver. We also see a resulting reduction in faecal phenylalanine and tyrosine. Consistent with these findings, PAG has previously been associated with a lean phenotype [43]. In our study PAG levels predominantly correlate with bacterial species from the *Actinobacteria* and *Proteobacteria* phyla as well as a number of bacterial pathways including increased gluconeogenesis, anaerobic energy metabolism and glutamate degradation to SCFAs. The ketone body acetoacetate, also raised with higher physical activity and the combined effect of healthy diet and exercise, can also be produced through ketogenic metabolism of phenylalanine.

Whilst PAG is positively associated with the group of healthy eating athletes (Model 2), the more minor microbial phenylalanine metabolites 3-hydroxyphenylacetate (HPA) and 3-(3-hydroxyphenyl)-3-hydroxypropionic acid (HPHPA) are inversely correlated, whereas HPA levels are higher in the urine of healthy eating controls (Model 1). The biological role of HPA and HPHPA is disputed. Some studies have associated HPA and HPHPA with neurological disorders [44]. Others however, have shown the contrary [45] and furthermore, HPA is thought to be a dietary marker of rutin intake, a flavonoid and antioxidant [46].

Benzoate Metabolism

It should also be noted that phenylalanine and other dietary aromatic compounds such as catechin can be metabolised to benzoate, which is further metabolised to hippurate [47]. Here we find higher hippurate in individuals with a healthy diet scores (Model 1). Higher hippurate levels have been correlated with a lean phenotype and lower blood

pressure as well as fruit intake [48, 49]. In this study, urinary hippurate levels correlate strongly with faecal 3-phenylpropionate, suggesting that the main source of hippurate production observed here is through degradation of catechin from the diet.

Choline Metabolism

We find higher levels of TMAO in both the healthy diet and exercise effect models in addition to the combined model. Circulating trimethylamine-*N*-Oxide (TMAO) is thought to predict cardiovascular disease (CVD), possibly through altered cholesterol metabolism and oxidative stress [4, 50]. However, the causative effect in humans is disputed and may be due to confounders such as kidney function and poor metabolic control [51]. Furthermore, TMAO is present in high concentrations in the tissues of cold water-dwelling fish where it acts as an antifreeze agent and is also found in high levels in the urine of Japanese populations whose diet contains a high portion of fish and who do not have high risk for CVD[48]. Moreover, TMAO was recently found to protect against impaired glucose tolerance and reduce endoplasmic reticulum stress [52].

As expected, we find urinary TMAO levels correlate with faecal trimethylamine and methylamine, although not dimethylamine. Further, TMAO correlates with urinary carnitine, acetyl-carnitine and urea levels, all markers of protein intake, which is consistent with the higher reported intake of protein in the healthy eating athletes represented in the exercise and combined models. The bacterial abundance data also show a dichotomy with respect to association between TMAO and DMA. For example, we identify a large number of bacterial species and pathways whose abundance correlates with faecal dimethylamine excretion but which do not appear to influence TMAO levels, whereas the heterofermentative lactic acid bacterium *Leuconostoc citreum* does correlate with higher TMAO levels. Intriguingly we also identify a number of lactococcal phage that correlate with TMAO. This

raises the possibility that changes to the gut environment, such as altered pH from changes to lactic acid bacteria, can alter levels of TMA / TMAO producing and metabolising bacteria leading to subsequent changes in host TMAO levels. Indeed, lactococcal phage levels also correlate with increased TMAO:DMA and TMA:DMA ratios, suggesting they may influence bacterial TMA dehydrogenase or TMAO aldolase activity.

4.5.3 | SCFA and Branched chain amino acid Metabolism

Faecal SCFAs are found in higher concentrations in subjects with increased physical activity (Model 2) and combined healthy eating and increased exercise (Model 3) relative to controls. SCFAs are produced through the bacterial fermentation of dietary fibre and complex carbohydrates. The most abundant SCFAs are acetate (C2), propionate (C3) and butyrate (C4). In addition, valerate and the branched SCFAs isobutyrate, isovalerate and 2-methyl-butyrate are produced by gut bacteria in smaller quantities [53]. Butyrate is the preferred source of energy for colonic epithelial cells [54], whereas propionate, along with remaining unused butyrate, is used predominantly by hepatocytes for gluconeogenesis [55], while acetate is mainly utilised by muscle cells to generate energy. Importantly, SCFAs have been associated with reduced appetite and weight loss through stimulating release of the anorectic gut hormones peptide YY (PYY) and glucagon like peptide-1 (GLP-1) [53, 56]. Interestingly, it has been reported that undigested proteins and amino acids in the colon may serve as an additional substrate for SCFA production [57]. Indeed, in this cohort we find that while faecal butyrate correlates strongly with fibre intake, propionate correlates most strongly with protein intake – increased in healthy eating athletes.

Alternatively, medium chain fatty acids (MCFA), branched SCFAs, and branched chain keto acids (BCKA) are correlated with higher bacterial diversity, seen after modelling the combined effects of healthy diet and exercise. The branched SCFAs isobutyrate,

isovalerate and 2-methylbutyrate are produced through the fermentation of BCAAs and have also been shown to modulate energy metabolism [58]. Here, isobutyrate and isovalerate are both correlated with bacterial amino acid (glutamate) degradation pathways to propionate and butyrate and show an inverse correlation to pathways encoding starch degradation. It is likely that bacteria here are using similar pathways to ferment BCAAs in place of glutamate. Furthermore, in faecal samples with higher bacterial diversity we detect higher branched chain keto acids (BCKA) 2-oxoisocaproate and 2-oxoisovalerate, produced through the catabolism of the BCAAs leucine and valine, respectively. This initial transamination reaction, catalysed by branched chain aminotransferase (BCAT), produces glutamate (also higher) and the respective BCKA. Subsequent leucine metabolism via 2-oxoisocaproate is ketogenic, while valine metabolism via 2-oxoisovalerate is glucogenic. This constellation of raised BCKAs, glutamate and branched-SCFAs suggests that the increased diversity observed here is associated with an adaptation towards increased degradation of protein and amino acids in the colon. Additionally, we see reduced glucose and amino acid levels (histidine, asparagine, isoleucine) with increased bacterial diversity. It is not clear whether these changes are a result of bacterial adaptation to a high protein / low sugar environment in the distal colon or whether they are driven by an increase in bacterial utilisation of amino acids. Intriguingly the reduction in amino acid levels, seen here with increased diversity, appears to indicate that enhanced protein degradation increases beyond increased intake. Corresponding changes are also seen in bacterial pathways related with increased diversity. Increased diversity correlates with higher; amino acid, BCAA and nicotinate degradation, co-factor and amino acid biosynthesis and gluconeogenesis pathways, along with reduced starch degradation, methionine and fatty acid biosynthesis pathways.

In addition, SCFA can also be utilised by bacteria for *de novo* amino acid biosynthesis [59]. Here, increased valerate, seen with increased exercise (Model 2 & 3) and bacterial diversity, is associated with bacterial phenylalanine and tyrosine biosynthesis pathways. Increased acetate, seen with increased exercise (Model 2 & 3), is associated with tryptophan biosynthesis pathways. Lastly, acetate is associated with guanosine ribonucleotide biosynthesis, suggesting that acetate has an important role in enabling the biosynthesis of aromatic compounds.

4.6 | Conclusion

The analysis of individuals' gut microbiome and metabolic profiles according to groupings based on an objective measure of their dietary habits, exercise levels and further analysis according to bacterial diversity, has elicited novel insights into this complex, multi-factorial paradigm. Each exercise/diet group appears to exhibit a distinct metabolic and functional microbial phenotype.

Model synthesis using an objective measure of dietary healthiness has enabled us to control for a number of closely related confounders, and more accurately ascertain the individual biological sequelae of diet healthiness and exercise, respectively. However, although reduced, a number of confounders in this paradigm still exist. Some changes exhibited in the exercise effect model may be due to increased meat and protein intake in this group. However, where possible this has been explained. Additionally, a subgroup of athletes consumed the probiotics *Lactobacillus casei* or *Lactobacillus acidophilus* regularly. However, their relative abundance in the faecal metagenomic data did not significantly correlate with any faecal or urinary metabolites. BMI has also been shown to be a major anthropometric factor influencing the gut microbiome [10, 60, 61]. While this present study begins to address the effects of diet and exercise as separate factors influencing the

microbiome, it does not circumvent the effects of BMI on the microbiome as the subjects included in the control models are still largely constrained to the *high* and *low* BMI groups described in chapter 3.

Importantly, increased microbial diversity has previously been associated with a healthy state [7-10]. Certainly, many microbial derived metabolites such as SCFA are known to be beneficial and furthermore, an increase in diversity is associated with a low glycaemic state. However, in this cohort we also see increased diversity relating to an increase towards proteolytic fermentation by gut bacteria. Protein metabolism within the gut has been associated with the production of toxic compounds such as ammonia, amines and sulphides [62, 63]. Thus, indicating that in this cohort the effects of increased diversity may include deleterious as well as beneficial effects on the host. This illustrates the importance of more in-depth analysis at a pathway and taxonomic level in future microbiome studies.

4.7 | References

[1] P. J. Turnbaugh, M. Hamady, T. Yatsunenko, B. L. Cantarel, A. Duncan, R. E. Ley, M. L. Sogin, W. J. Jones, B. A. Roe, J. P. Affourtit, M. Egholm, B. Henrissat, A. C. Heath, R. Knight, J. I. Gordon, A core gut microbiome in obese and lean twins, Nature 457 (7228) (2009) 480-4.

[2] P. J. Turnbaugh, R. E. Ley, M. A. Mahowald, V. Magrini, E. R. Mardis, J. I. Gordon, An obesity-associated gut microbiome with increased capacity for energy harvest, Nature 444 (7122) (2006) 1027-31.

[3] J. Qin, Y. Li, Z. Cai, S. Li, J. Zhu, F. Zhang, S. Liang, W. Zhang, Y. Guan, D. Shen, Y. Peng, D. Zhang, Z. Jie, W. Wu, Y. Qin, W. Xue, J. Li, L. Han, D. Lu, P. Wu, Y. Dai, X. Sun, Z. Li, A. Tang, S. Zhong, X. Li, W. Chen, R. Xu, M. Wang, Q. Feng, M. Gong, J. Yu, Y. Zhang, M. Zhang, T. Hansen, G. Sanchez, J. Raes, G. Falony, S. Okuda, M. Almeida, E. LeChatelier, P. Renault, N. Pons, J. M. Batto, Z. Zhang, H. Chen, R. Yang, W. Zheng, S. Li, H. Yang, J. Wang, S. D. Ehrlich, R. Nielsen, O. Pedersen, K. Kristiansen, J. Wang, A metagenome-wide association study of gut microbiota in type 2 diabetes, Nature 490 (7418) (2012) 55-60.

[4] Z. Wang, E. Klipfell, B. J. Bennett, R. Koeth, B. S. Levison, B. Dugar, A. E. Feldstein, E. B. Britt, X. Fu, Y. M. Chung, Y. Wu, P. Schauer, J. D. Smith, H. Allayee, W. H. Tang, J. A. DiDonato, A. J. Lusis, S. L. Hazen, Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease, Nature 472 (7341) (2011) 57-63.

[5] R. F. Schwabe, C. Jobin, The microbiome and cancer, Nat Rev Cancer 13 (11) (2013) 800-12.

[6] J. Kinross, R. Mirnezami, J. Alexander, R. Brown, A. Scott, D. Galea, K. Veselkov, R. Goldin, A. Darzi, J. Nicholson, J. R. Marchesi, A prospective analysis of mucosal microbiome-metabonome interactions in colorectal cancer using a combined MAS 1HNMR and metataxonomic strategy, Sci Rep 7 (1) (2017) 8979.

[7] J. L. Sonnenburg, F. Backhed, Diet-microbiota interactions as moderators of human metabolism, Nature 535 (7610) (2016) 56-64.

[8] E. Le Chatelier, T. Nielsen, J. Qin, E. Prifti, F. Hildebrand, G. Falony, M. Almeida, M. Arumugam, J. M. Batto, S. Kennedy, P. Leonard, J. Li, K. Burgdorf, N. Grarup, T. Jorgensen, I. Brandslund, H. B. Nielsen, A. S. Juncker, M. Bertalan, F. Levenez, N. Pons, S. Rasmussen, S. Sunagawa, J. Tap, S. Tims, E. G. Zoetendal, S. Brunak, K. Clement, J. Dore, M. Kleerebezem, K. Kristiansen, P. Renault, T. Sicheritz-Ponten, W. M. de Vos, J. D. Zucker, J. Raes, T. Hansen, H. I. T. c. Meta, P. Bork, J. Wang, S. D. Ehrlich, O. Pedersen, Richness of human gut microbiome correlates with metabolic markers, Nature 500 (7464) (2013) 541-6.

[9] C. Menni, M. A. Jackson, T. Pallister, C. J. Steves, T. D. Spector, A. M. Valdes, Gut microbiome diversity and high-fibre intake are related to lower long-term weight gain, Int J Obes (Lond) 41 (7) (2017) 1099-1105.

[10] A. Zhernakova, A. Kurilshikov, M. J. Bonder, E. F. Tigchelaar, M. Schirmer, T. Vatanen, Z. Mujagic, A. V. Vila, G. Falony, S. Vieira-Silva, J. Wang, F. Imhann, E. Brandsma, S. A. Jankipersadsing, M. Joossens, M. C. Cenit, P. Deelen, M. A. Swertz, s. LifeLines cohort, R. K. Weersma, E. J. Feskens, M. G. Netea, D. Gevers, D. Jonkers, L. Franke, Y. S. Aulchenko, C. Huttenhower, J. Raes, M. H. Hofker, R. J. Xavier, C. Wijmenga, J. Fu, Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity, Science 352 (6285) (2016) 565-9.

[11] L. A. David, C. F. Maurice, R. N. Carmody, D. B. Gootenberg, J. E. Button, B. E. Wolfe, A. V. Ling, A. S. Devlin, Y. Varma, M. A. Fischbach, S. B. Biddinger, R. J. Dutton, P. J. Turnbaugh, Diet rapidly and reproducibly alters the human gut microbiome, Nature 505 (7484) (2014) 559-63.

[12] A. Mika, W. Van Treuren, A. Gonzalez, J. J. Herrera, R. Knight, M. Fleshner, Exercise is More Effective at Altering Gut Microbial Composition and Producing Stable Changes in Lean Mass in Juvenile versus Adult Male F344 Rats, PLoS One 10 (5) (2015) e0125889.

[13] J. J. Choi, S. Y. Eum, E. Rampersaud, S. Daunert, M. T. Abreu, M. Toborek, Exercise attenuates PCB-induced changes in the mouse gut microbiome, Environ Health Perspect 121 (6) (2013) 725-30.

[14] M. I. Queipo-Ortuno, L. M. Seoane, M. Murri, M. Pardo, J. M. Gomez-Zumaquero, F. Cardona, F. Casanueva, F. J. Tinahones, Gut microbiota composition in male rat models under different nutritional status and physical activity and its association with serum leptin and ghrelin levels, PLoS One 8 (5) (2013) e65465.

[15] E. V. Lamoureux, S. A. Grandy, M. G. I. Langille, Moderate Exercise Has Limited but Distinguishable Effects on the Mouse Microbiome, mSystems 2 (4) (2017)

[16] M. Estaki, J. Pither, P. Baumeister, J. P. Little, S. K. Gill, S. Ghosh, Z. Ahmadi-Vand, K. R. Marsden, D. L. Gibson, Cardiorespiratory fitness as a predictor of intestinal microbial diversity and distinct metagenomic functions, Microbiome 4 (1) (2016) 42.

[17] C. Bressa, M. Bailen-Andrino, J. Perez-Santiago, R. Gonzalez-Soltero, M. Perez, M. G. Montalvo-Lominchar, J. L. Mate-Munoz, R. Dominguez, D. Moreno, M. Larrosa, Differences in gut microbiota profile between women with active lifestyle and sedentary women, PLoS One 12 (2) (2017) e0171352.

[18] W. Barton, N. C. Penney, O. Cronin, I. Garcia-Perez, M. G. Molloy, E. Holmes, F. Shanahan, P. D. Cotter, O. O'Sullivan, The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level, Gut (2017)

[19] S. F. Clarke, E. F. Murphy, O. O'Sullivan, A. J. Lucey, M. Humphreys, A. Hogan, P. Hayes, M. O'Reilly, I. B. Jeffery, R. Wood-Martin, D. M. Kerins, E. Quigley, R. P. Ross, P. W. O'Toole, M. G. Molloy, E. Falvey, F. Shanahan, P. D. Cotter, Exercise and associated dietary extremes impact on gut microbial diversity, Gut 63 (12) (2014) 1913-20.

[20] L. M. Petersen, E. J. Bautista, H. Nguyen, B. M. Hanson, L. Chen, S. H. Lek, E. Sodergren, G. M. Weinstock, Community characteristics of the gut microbiomes of competitive cyclists, Microbiome 5 (1) (2017) 98.

[21] K. Poslusna, J. Ruprich, J. H. de Vries, M. Jakubikova, P. van't Veer, Misreporting of energy and micronutrient intake estimated by food records and 24 hour recalls, control and adjustment methods in practice, Br J Nutr 101 Suppl 2 (2009) S73-85.

[22] K. L. Rennie, A. Coward, S. A. Jebb, Estimating under-reporting of energy intake in dietary surveys using an individualised method, Br J Nutr 97 (6) (2007) 1169-76.

[23] I. Garcia-Perez, J. M. Posma, R. Gibson, E. S. Chambers, T. H. Hansen, H. Vestergaard, T. Hansen, M. Beckmann, O. Pedersen, P. Elliott, J. Stamler, J. K. Nicholson, J. Draper, J. C. Mathers, E. Holmes, G. Frost, Objective assessment of dietary patterns by use of metabolic phenotyping: a randomised, controlled, crossover trial, Lancet Diabetes Endocrinol 5 (3) (2017) 184-195.

[24] N. J. Wareham, R. W. Jakes, K. L. Rennie, J. Mitchell, S. Hennings, N. E. Day, Validity and repeatability of the EPIC-Norfolk Physical Activity Questionnaire, Int J Epidemiol 31 (1) (2002) 168-74.

[25] S. Abubucker, N. Segata, J. Goll, A. M. Schubert, J. Izard, B. L. Cantarel, B. Rodriguez-Mueller, J. Zucker, M. Thiagarajan, B. Henrissat, O. White, S. T. Kelley, B. Methe, P. D. Schloss, D. Gevers, M. Mitreva, C. Huttenhower, Metabolic reconstruction for metagenomic data and its application to the human microbiome, PLoS Comput Biol 8 (6) (2012) e1002358.

[26] R. Caspi, T. Altman, R. Billington, K. Dreher, H. Foerster, C. A. Fulcher, T. A. Holland, I. M. Keseler, A. Kothari, A. Kubo, M. Krummenacker, M. Latendresse, L. A. Mueller, Q. Ong, S. Paley, P. Subhraveti, D. S. Weaver, D. Weerasinghe, P. Zhang, P. D. Karp, The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases, Nucleic Acids Res 42 (Database issue) (2014) D459-71.

[27] D. E. Wood, S. L. Salzberg, Kraken: ultrafast metagenomic sequence classification using exact alignments, Genome Biol 15 (3) (2014) R46.

[28] R Development Core Team, R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna, 2012), URL: http:// www. R-project. org (2015)

[29] J. Oksanen, F. G. Blanchet, R. Kindt, P. Legendre, P. R. Minchin, R. O'Hara, G. L. Simpson, P. Solymos, M. H. H. Stevens, H. Wagner, Multivariate analysis of ecological communities in R: vegan tutorial., R package version 2.3-1. (2013)

[30] H. Wickham, ggplot2: elegant graphics for data analysis, Springer Science & Business Media

[31] Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing, Journal of the Royal Statistical Society Series B-Methodological 57 (1) (1995) 289-300.

[32] A. C. Dona, B. Jimenez, H. Schafer, E. Humpfer, M. Spraul, M. R. Lewis, J. T. Pearce, E. Holmes, J. C. Lindon, J. K. Nicholson, Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping, Anal Chem 86 (19) (2014) 9887-94.

[33] R. Garcia-Villalba, J. A. Gimenez-Bastida, M. T. Garcia-Conesa, F. A. Tomas-Barberan, J. Carlos Espin, M. Larrosa, Alternative method for gas chromatography-mass spectrometry analysis of short-chain fatty acids in faecal samples, J Sep Sci 35 (15) (2012) 1906-13.

[34] K. A. Veselkov, L. K. Vingara, P. Masson, S. L. Robinette, E. Want, J. V. Li, R. H. Barton, C. Boursier-Neyret, B. Walther, T. M. Ebbels, I. Pelczer, E. Holmes, J. C. Lindon, J. K. Nicholson, Optimized preprocessing of ultra-performance liquid chromatography/mass spectrometry urinary metabolic profiles for improved information recovery, Anal Chem 83 (15) (2011) 5864-72.

[35] Z. Gu, R. Eils, M. Schlesner, Complex heatmaps reveal patterns and correlations in multidimensional genomic data, Bioinformatics 32 (18) (2016) 2847-9.

[36] J. M. Posma, I. Garcia-Perez, M. De Iorio, J. C. Lindon, P. Elliott, E. Holmes, T. M. Ebbels, J. K. Nicholson, Subset optimization by reference matching (STORM): an optimized statistical approach for recovery of metabolic biomarker structural information from 1H NMR spectra of biofluids, Anal Chem 84 (24) (2012) 10694-701.

[37] O. Cloarec, M. E. Dumas, A. Craig, R. H. Barton, J. Trygg, J. Hudson, C. Blancher, D. Gauguier, J. C. Lindon, E. Holmes, J. Nicholson, Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H NMR data sets, Anal Chem 77 (5) (2005) 1282-9.

[38] D. A. Cobb-Clark, S. C. Kassenboehmer, S. Schurer, Healthy habits: The connection between diet, exercise, and locus of control, Journal of Economic Behavior & Organization 98 (2014) 1-28.

[39] A. Pechlivanis, S. Kostidis, P. Saraslanidis, A. Petridou, G. Tsalis, V. Mougios, H. G. Gika, E. Mikros, G. A. Theodoridis, (1)H NMR-based metabonomic investigation of the effect of two different exercise sessions on the metabolic fingerprint of human urine, J Proteome Res 9 (12) (2010) 6405-16.

[40] D. C. Nieman, N. D. Gillitt, W. Sha, M. P. Meaney, C. John, K. L. Pappan, J. M. Kinchen, Metabolomics-Based Analysis of Banana and Pear Ingestion on Exercise Performance and Recovery, J Proteome Res 14 (12) (2015) 5367-77.

[41] J. Mayneris-Perxachs, A. A. Lima, R. L. Guerrant, A. M. Leite, A. F. Moura, N. L. Lima, A. M. Soares, A. Havt, S. R. Moore, R. Pinkerton, J. R. Swann, Urinary N-methylnicotinamide and beta-aminoisobutyric acid predict catch-up growth in undernourished Brazilian children, Sci Rep 6 (2016) 19780.

[42] D. Kraus, Q. Yang, D. Kong, A. S. Banks, L. Zhang, J. T. Rodgers, E. Pirinen, T. C. Pulinilkunnil, F. Gong, Y. C. Wang, Y. Cen, A. A. Sauve, J. M. Asara, O. D. Peroni, B. P. Monia, S. Bhanot, L. Alhonen, P. Puigserver, B. B. Kahn, Nicotinamide N-methyltransferase knockdown protects against diet-induced obesity, Nature 508 (7495) (2014) 258-62.

[43] P. Elliott, J. M. Posma, Q. Chan, I. Garcia-Perez, A. Wijeyesekera, M. Bictash, T. M. Ebbels, H. Ueshima, L. Zhao, L. van Horn, M. Daviglus, J. Stamler, E. Holmes, J. K. Nicholson, Urinary metabolic signatures of human adiposity, Sci Transl Med 7 (285) (2015) 285ra62.

[44] X. Xiong, D. Liu, Y. Wang, T. Zeng, Y. Peng, Urinary 3-(3-Hydroxyphenyl)-3-hydroxypropionic Acid, 3-Hydroxyphenylacetic Acid, and 3-Hydroxyhippuric Acid Are Elevated in Children with Autism Spectrum Disorders, Biomed Res Int 2016 (2016) 9485412.

[45] P. Emond, S. Mavel, N. Aidoud, L. Nadal-Desbarats, F. Montigny, F. Bonnet-Brilhault, C. Barthelemy, M. Merten, P. Sarda, F. Laumonnier, P. Vourc'h, H. Blasco, C. R. Andres, GC-MS-based urine metabolic profiling of autism spectrum disorders, Anal Bioanal Chem 405 (15) (2013) 5291-300.

[46] A. M. Aura, K. A. O'Leary, G. Williamson, M. Ojala, M. Bailey, R. Puupponen-Pimiä, A. M. Nuutila, K. M. Oksman-Caldentey, K. Poutanen, Quercetin Derivatives Are Deconjugated and Converted to Hydroxyphenylacetic Acids but Not Methylated by Human Fecal Flora in Vitro, Journal of Agricultural and Food Chemistry 50 (6) (2002) 1725-1730.

[47] H. J. Lees, J. R. Swann, I. D. Wilson, J. K. Nicholson, E. Holmes, Hippurate: the natural history of a mammalian-microbial cometabolite, J Proteome Res 12 (4) (2013) 1527-46.

[48] E. Holmes, R. L. Loo, J. Stamler, M. Bictash, I. K. Yap, Q. Chan, T. Ebbels, M. De Iorio, I. J. Brown, K. A. Veselkov, M. L. Daviglus, H. Kesteloot, H. Ueshima, L. Zhao, J. K. Nicholson, P. Elliott, Human metabolic phenotype diversity and its association with diet and blood pressure, Nature 453 (7193) (2008) 396-400.

[49] S. S. Heinzmann, C. A. Merrifield, S. Rezzi, S. Kochhar, J. C. Lindon, E. Holmes, J. K. Nicholson, Stability and robustness of human metabolic phenotypes in response to sequential food challenges, J Proteome Res 11 (2) (2012) 643-55.

[50] W. H. Tang, Z. Wang, B. S. Levison, R. A. Koeth, E. B. Britt, X. Fu, Y. Wu, S. L. Hazen, Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk, N Engl J Med 368 (17) (2013) 1575-84.

[51] D. M. Mueller, M. Allenspach, A. Othman, C. H. Saely, A. Muendlein, A. Vonbank, H. Drexel, A. von Eckardstein, Plasma levels of trimethylamine-N-oxide are confounded by impaired kidney function and poor metabolic control, Atherosclerosis 243 (2) (2015) 638-44.

[52] M. E. Dumas, A. R. Rothwell, L. Hoyles, T. Aranias, J. Chilloux, S. Calderari, E. M. Noll, N. Pean, C. L. Boulange, C. Blancher, R. H. Barton, Q. Gu, J. F. Fearnside, C. Deshayes, C. Hue, J. Scott, J. K. Nicholson, D. Gauguier, Microbial-Host Co-metabolites Are Prodromal Markers Predicting Phenotypic Heterogeneity in Behavior, Obesity, and Impaired Glucose Tolerance, Cell Rep 20 (1) (2017) 136-148.

[53] H. V. Lin, A. Frassetto, E. J. Kowalik, Jr., A. R. Nawrocki, M. M. Lu, J. R. Kosinski, J. A. Hubert, D. Szeto, X. Yao, G. Forrest, D. J. Marsh, Butyrate and propionate protect against diet-induced obesity and regulate gut hormones via free fatty acid receptor 3-independent mechanisms, PLoS One 7 (4) (2012) e35240.

[54] D. R. Donohoe, N. Garge, X. Zhang, W. Sun, T. M. O'Connell, M. K. Bunger, S. J. Bultman, The microbiome and butyrate regulate energy metabolism and autophagy in the mammalian colon, Cell Metab 13 (5) (2011) 517-26.

[55] C. Demigné, C. Yacoub, C. Rémésy, P. Fafournoux, Propionate and butyrate metabolism in rat or sheep hepatocytes, Biochimica et Biophysica Acta (BBA) - Lipids and Lipid Metabolism 875 (3) (1986) 535-542.

[56] E. S. Chambers, A. Viardot, A. Psichas, D. J. Morrison, K. G. Murphy, S. E. Zac-Varghese, K. MacDougall, T. Preston, C. Tedford, G. S. Finlayson, J. E. Blundell, J. D. Bell, E. L. Thomas, S. Mt-Isa, D. Ashby, G. R. Gibson, S. Kolida, W. S. Dhillo, S. R. Bloom, W. Morley, S. Clegg, G. Frost, Effects of targeted delivery of propionate to the human colon on appetite regulation, body weight maintenance and adiposity in overweight adults, Gut 64 (11) (2015) 1744-54.

[57] T. P. Bui, J. Ritari, S. Boeren, P. de Waard, C. M. Plugge, W. M. de Vos, Production of butyrate from lysine and the Amadori product fructoselysine by a human gut commensal, Nat Commun 6 (2015) 10062.

[58] E. Heimann, M. Nyman, A. K. Palbrink, K. Lindkvist-Petersson, E. Degerman, Branched short-chain fatty acids modulate glucose and lipid metabolism in primary adipocytes, Adipocyte 5 (4) (2016) 359-368.

[59] E. P. Neis, C. H. Dejong, S. S. Rensen, The role of microbial amino acid metabolism in host metabolism, Nutrients 7 (4) (2015) 2930-46.

[60] Y. Yun, H. N. Kim, S. E. Kim, S. G. Heo, Y. Chang, S. Ryu, H. Shin, H. L. Kim, Comparative analysis of gut microbiota associated with body mass index in a large Korean cohort, BMC Microbiol 17 (1) (2017) 151.

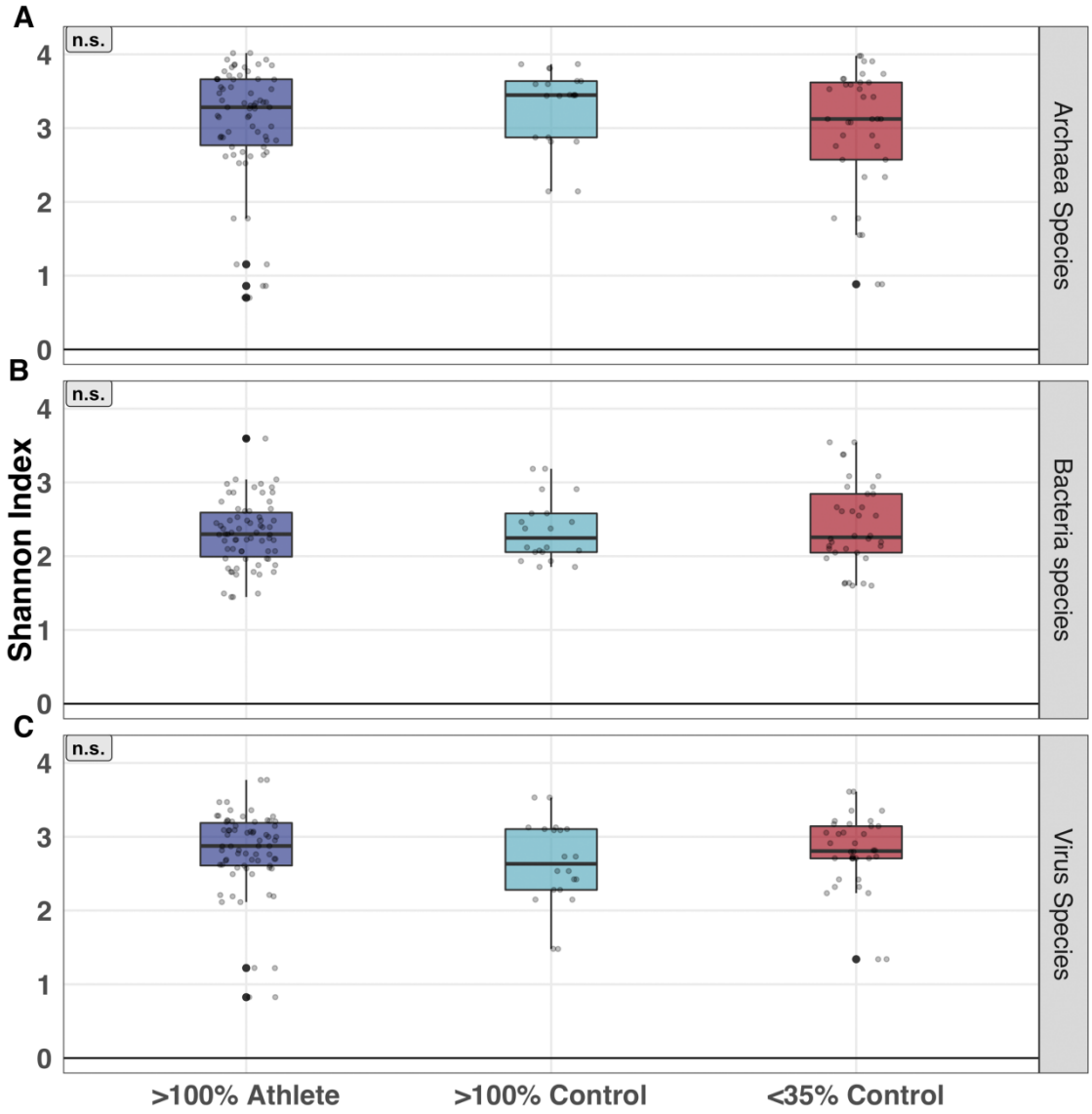
[61] D. Rothschild, O. Weissbrod, E. Barkan, A. Kurilshikov, T. Korem, D. Zeevi, P. I. Costea, A. Godneva, I. N. Kalka, N. Bar, S. Shilo, D. Lador, A. V. Vila, N. Zmora, M. Pevsner-Fischer,

D. Israeli, N. Kosower, G. Malka, B. C. Wolf, T. Avnit-Sagi, M. Lotan-Pompan, A. Weinberger, Z. Halpern, S. Carmi, J. Fu, C. Wijmenga, A. Zhernakova, E. Elinav, E. Segal, Environment dominates over host genetics in shaping human gut microbiota, Nature 555 (7695) (2018) 210-215.

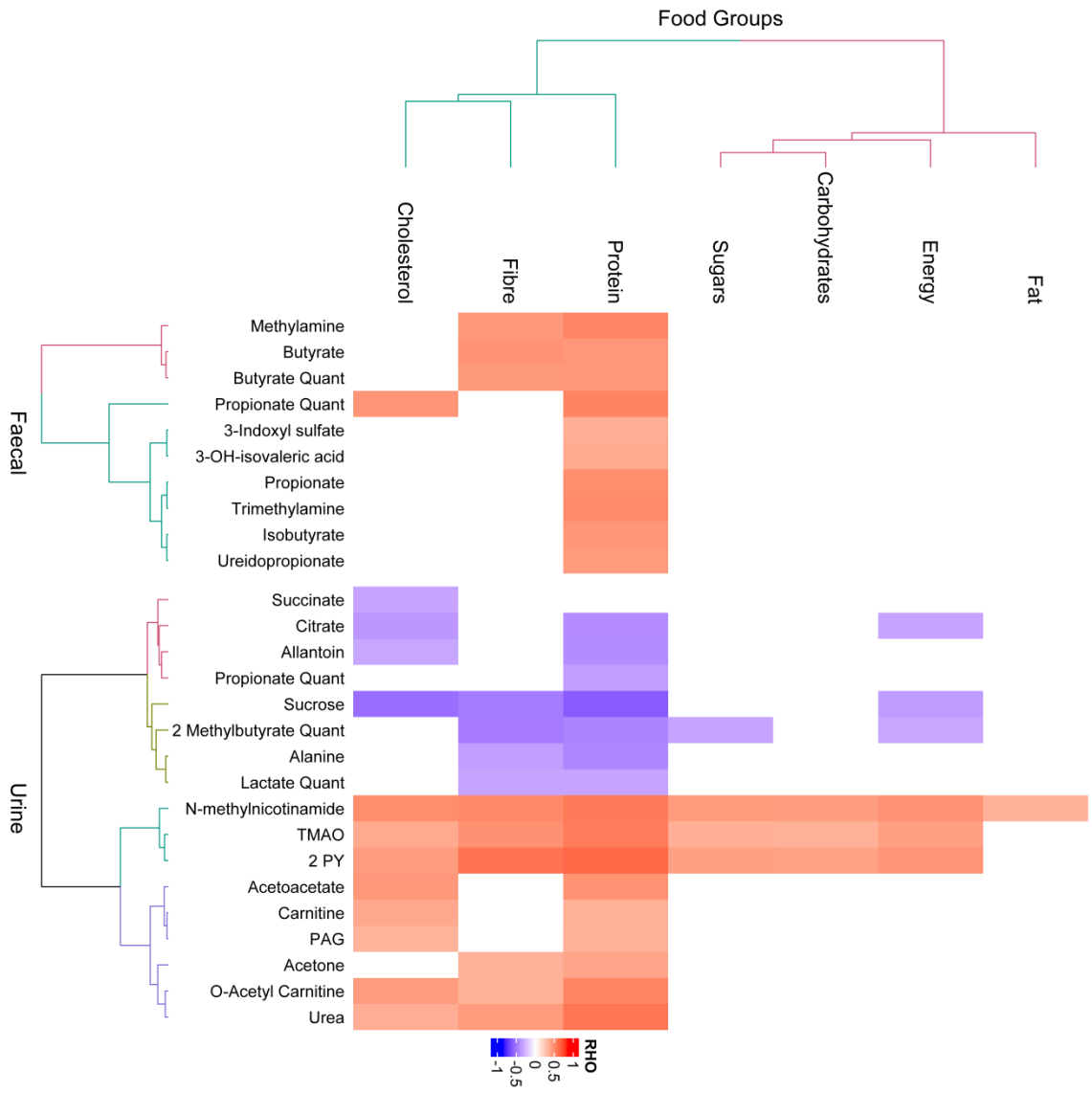
[62] G. T. Macfarlane, J. H. Cummings, C. Allison, Protein degradation by human intestinal bacteria, J Gen Microbiol 132 (6) (1986) 1647-56.

[63] R. Hughes, E. A. Magee, S. Bingham, Protein degradation in the large intestine: relevance to colorectal cancer, Curr Issues Intest Microbiol 1 (2) (2000) 51-8.

4.8 | Supplementary Content



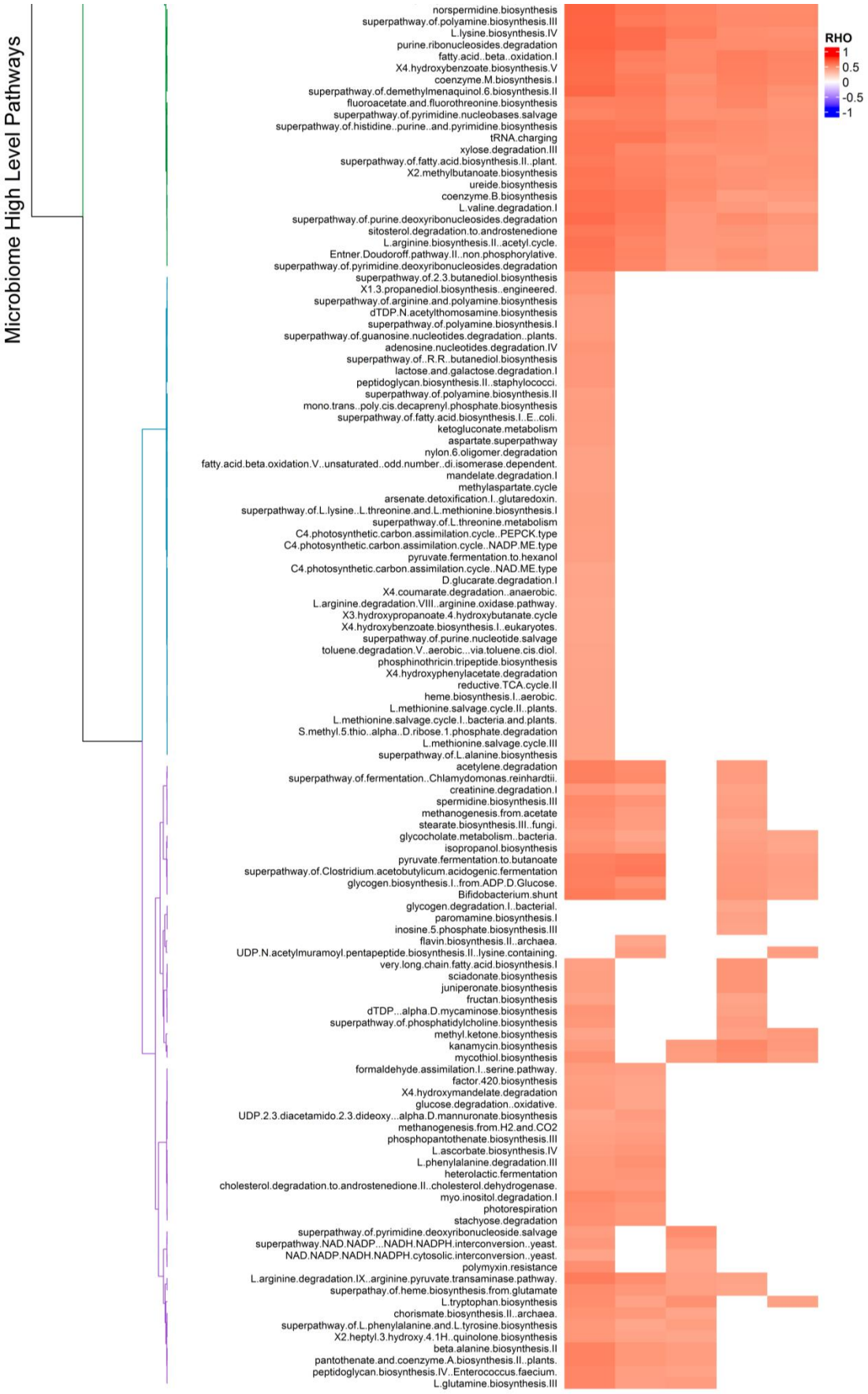
Supplementary Figure 4.1 | Taxonomic alpha diversity. Shannon H-index generated for species of Archaea (**A**), Bacteria (**B**), and viruses (**C**).



Supplementary Figure 4.2: Diet–Metabolite interactions. Significant Spearman's correlations (pFDR <0.05) between diet vs. faecal and urinary metabolic datasets are shown, shaded according to strength of correlation coefficient (Rho). Correlations are clustered according to Euclidean distances.

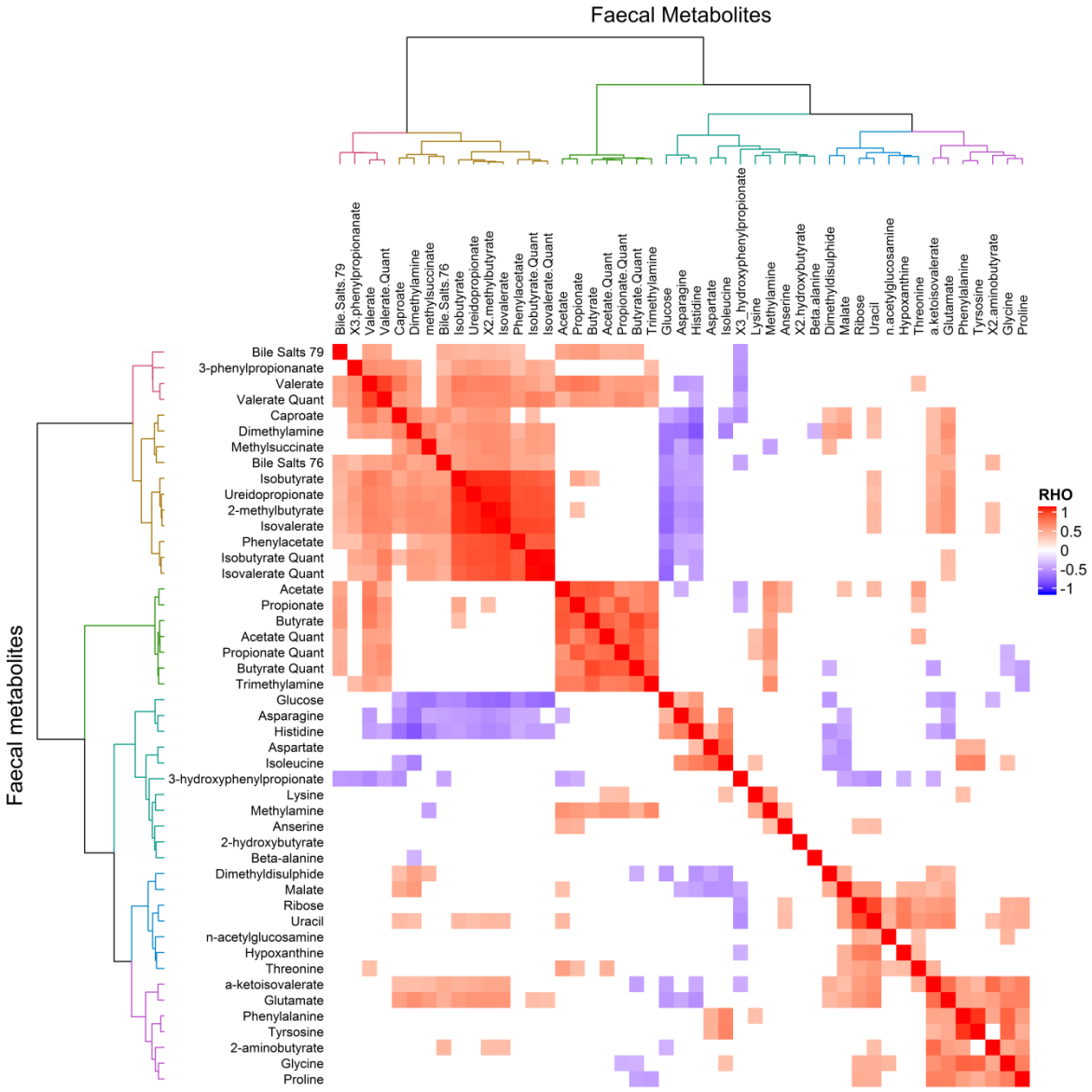
Supplementary Figure 4.3: Gut Species - Metabolite interactions. Significant Spearman's correlations (pFDR <0.01) between microbial species vs faecal and urinary metabolic datasets are shown, shaded according to strength of correlation coefficient (Rho). Correlations are clustered according to Euclidean distances (**Figure omitted due to size**).

Supplementary Figure 4.4 (Part A)

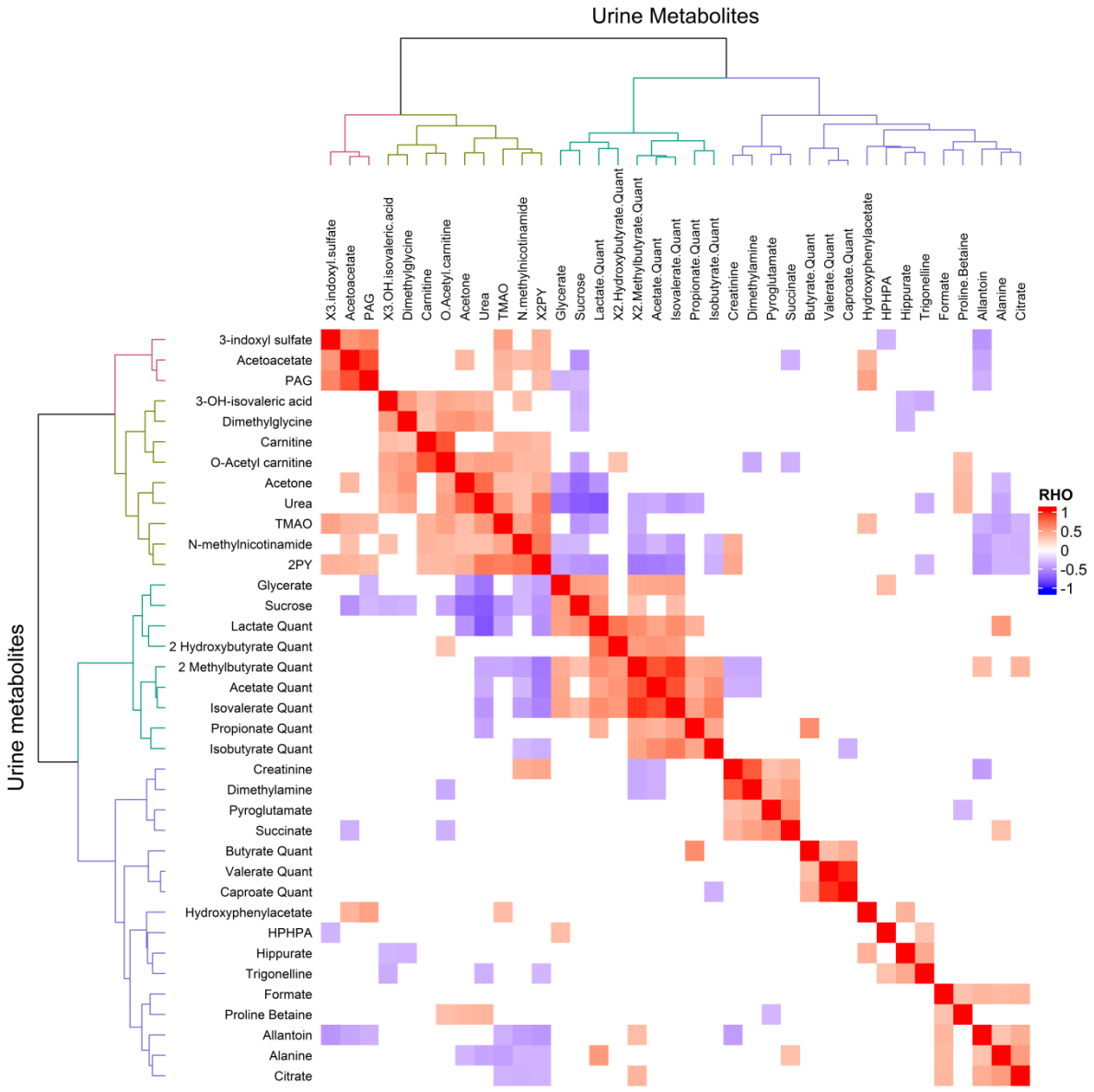


Supplementary Figure 4.4 (Part B)

Supplementary Figure 4.4: Metagenome Superpathway – Diversity Indices correlations. Significant Spearman's correlations (pFDR <0.01) between microbial metabolic super-pathways and diversity indices (Shannon, Simpson, PD, Chao1) are shown, shaded according to strength of correlation coefficient (Rho). Correlations are clustered according to Euclidean distances.



Supplementary Figure 4.5: Faecal – Faecal Metabolite correlations. Significant Spearman's correlations (pFDR <0.05) within the faecal metabolite dataset are shown, shaded according to strength of correlation coefficient (Rho). Correlations are clustered according to Euclidean distances.



Supplementary Figure 4.6: Urine – Urine Metabolite correlations. Significant Spearman's correlations (pFDR <0.05) within the urine metabolite dataset are shown, shaded according to strength of correlation coefficient (Rho). Correlations are clustered according to Euclidean distances.

Urinary Metabolites			
<35% controls Vs >100% control [Model 1 – Diet healthiness]			
Metabolite	**^{1}H chemical shift δ (multiplicity)[a]**	**Association**	**Source / Role**
Acetone	2.24 (s)	↑	Ketone
Citrate	2.55~ (d), 2.69~ (d)	↑	Fruit / TCA cycle
Creatinine	3.05 (s), 4.06 (s)	↑	Muscle mass
TMAO	3.27 (s)	↑	Choline metabolism, fish, meat
Proline Betaine	3.30 (s), 3.11 (s)	↑	Citrus fruits
3-Hydroxyphenylacetate	3.48(s), 6.79 (m), 6.81 (m), 6.86 (m), 7.26 (t)	↑	Gut microbial - Tyrosine metabolism
Hippurate	7.55 (t), 7.64 (t), 7.84 (d), 3.98 (d)	↑	Fruits & vegetables
Formate	8.46 (s)	↑	Fruits & vegetables
Lactate	Targeted GC-MS	↓	Anaerobic metabolism
2-hydroxybutyrate	Targeted GC-MS	↓	Glutathione synthesis
N-methylnicotinamide (NMND)	4.486 (s), 8.19 (t), 8.90 (d), 8.97 (d), 9.28 (s)	↓	Vit B3 / Endogenous Tryptophan metabolism
Pyroglutamate	2.04 (m), 2.41 (m), 2.51 (m), 4.18 (q)	↓	Glutamate metabolism
Sucrose	4.23 (d)	↓	Sugars
>100% controls Vs >100% athletes [Model 2 - Exercise]			
Acetoacetate	2.29 (s), 3.45 (s)	↑	Ketone
Phenylacetylglutamine (PAG)	2.11 (m), 2.27 (m), 3.67 (m), 4.19 (m), 7.36 (t), 7.43 (t)	↑	Gut microbial – Phenylalanine metabolism
Carnitine	3.23 (s), 2.45 (dd), 3.43 (m)	↑	Red meat
N-methylnicotinamide	4.486 (s), 8.19 (t), 8.90 (d), 8.97 (d), 9.28 (s)	↑	Niacin (B3)
3-Indoxyl sulfate *[UV]*	7.51(d), 7.71 (d), 7.21 (m), 7.28 (m)	↑	Gut microbial – Tryptophan metabolism
TMAO	3.27 (s)	↑	Choline metabolism, fish, meat
2 methyl 2 pyridone 5 carboxamide (2PY)	3.65 (d), 6.67 (d), 7.83 (dd), 8.34 (d)	↑	NAD degradation, Tryptophan metabolism
O-Acetyl Carnitine	3.19 (s), 2.15 (s)	↑	Red meat
2-methylbutyrate	Targeted GC-MS	↓	Branched SCFA
Succinate	2.41 (s)	↓	TCA cycle
Citrate *[UV]*	2.55~ (d), 2.69~ (d)	↓	Fruit / TCA cycle
3-Hydroxyphenylacetate	3.48(s), 6.79 (m), 6.81 (m), 6.86 (m), 7.26 (t)	↓	Gut microbial / Tyrosine metabolism
Allantoin	5.39 (s)	↓	Gut microbial – oxidative stress
3-(3-hydroxyphenyl)-3-hydroxypropionic acid (HPHPA)	2.62 (dd), 2.71 (dd)*, 5.02 (dd), 6.85 (ddd), 6.91 (t), 6.98 (m), 7.30 (t)	↓	Gut microbial - dysbiosis
<35% controls Vs >100% athletes [Model 3 – Diet healthiness & Exercise]			
Acetoacetate	2.29 (s), 3.45 (s)	↑	Ketone
Phenylacetylglutamine (PAG)	2.11 (m), 2.27 (m), 3.67 (m), 4.19 (m), 7.36 (t), 7.43 (t)	↑	Gut microbial – Phenylalanine metabolism
Carnitine	3.23 (s), 2.45 (dd), 3.43 (m)	↑	Red meat
TMAO	3.27 (s)	↑	Choline metabolism, fish, meat
Proline Betaine	3.30 (s), 3.11 (s)	↑	Citrus fruits
3-Indoxyl sulfate	7.51(d), 7.71 (d), 7.21 (m), 7.28 (m)	↑	Gut microbial

Urinary Metabolites			
Hippurate	7.55 (t), 7.64 (t), 7.84 (d), 3.98 (d)	↑	Fruits & vegetables
2 methyl 2 pyridone 5 carboxamide (2PY)	3.65 (d), 6.67 (d), 7.83 (dd), 8.34 (d)	↑	NAD degradation, Tryptophan metabolism
O-Acetyl Carnitine	3.19 (s), 2.15 (s)	↑	Red meat
Acetate	Targeted GC-MS	↓	SCFA
Propionate	Targeted GC-MS	↓	SCFA
2-methylbutyrate	Targeted GC-MS	↓	Branched SCFA
Isovalerate	Targeted GC-MS	↓	Branched SCFA
Lactate	Targeted GC-MS	↓	Anaerobic metabolism
Alanine	1.49 (d), 3.79 (q)	↓	Amino acid
Pyroglutamate	2.04 (m), 2.41 (m), 2.51 (m), 4.18 (q)	↓	Glutamate metabolism
Sucrose	4.23 (d)	↓	Sugars
Allantoin	5.39 (s)	↓	Gut microbial – oxidative stress

Supplementary Table 4.1: Urinary metabolites discriminating between model classes. Tables show urinary metabolites increasing or decreasing with increased diet healthiness, exercise and combined diet healthiness and exercise in models 1, 2 and 3 respectively. [a]Multiplicity key is as follows: s=singlet, d=doublet, t=triplet, q=quartet, dd=doublet of doublets, m=multiplet. ^{1}H shifts marked by an asterisk (*) were not confirmed experimentally due to the absence of unambiguous resonance values and are taken from the Human Metabolome Database (HMDB).

Faecal Metabolites			
<35% controls Vs >100% control [Model 1 – Diet healthiness]			
Metabolite	**^{1}H chemical shift δ (multiplicity)[a]**	**Association**	**Source / Role**
SMCSO metabolite 1	2.78 (s)	↑	Cruciferous vegetables
SMCSO metabolite 2	2.81 (s)	↑	Cruciferous vegetables
S-methyl-L-cysteine-sulfoxide (SMCSO)	2.83 (s)	↑	Cruciferous vegetables
3-hydroxyphenylpropionate	2.47 (t), 2.84 (t), 6.75 (m), 6.81 (t), 6.87 (m), 7.25 (t)	↑	Gut microbial / Caffeine
Bile Salts 73	0.73 (s)	↓	Bile acids
2-aminobutyrate	0.98 (t)	↓	Glutathione synthesis
Methylsuccinate (citramalate)	1.09 (d), 2.13 (dd)*, 2.53 (dd), 2.63 (m)*	↓	Microbial, malate analogue
2-hydroxybutyrate	1.20 (d)	↓	Glutathione synthesis
Acetate	1.92 (s)	↓	SCFA
n-acetylglucosamine	2.06 (s), 3.47-3.85 (m), 5.21 (d)	↓	Amino sugar, bacterial cell wall constituent
Beta-alanine	2.56 (t), 3.18 (t)	↓	Beta amino acid Uracil metabolism
Ribose	3.52 (dd), 4.11 (t), 4.93 (d), 5.25 (d) +…	↓	Monosaccharide, RNA constituent
Hypoxanthine	8.21 (d)	↓	Purine derivative
>100% controls Vs >100% athletes [Model 2 - Exercise]			
Bile salts 79	0.79 (s)	↑	Bile acids
Valerate	0.89 (t), 1.30 (m), 1.54 (m), 2.18 (t)	↑	SCFA
Butyrate	0.90 (t), 1.56 (m), 2.16 (t) + Targeted GC-MS	↑	SCFA
Propionate	1.06 (t), 2.19 (q) + Targeted GC-MS	↑	SCFA
Threonine	1.33 (d), 3.59 (d), 4.26 (dqt)	↑	Amino acid
Acetate	1.92 (s) + Targeted GC-MS	↑	SCFA
Methylamine	2.60 (s)	↑	Choline metabolism
Trimethylamine	2.88 (s)	↑	Choline metabolism
Uracil	5.81 (d), 7.55 (d)	↑	RNA nucleobase
Anserine	7.17 (s), 8.35 (s), 3.65 (s)	↑	Dipeptide
3-hydroxyphenylpropionate	2.47 (t), 2.84 (t), 6.75 (m), 6.81 (t), 6.87 (m), 7.25 (t)	↓	Gut microbial / Caffeine
Dimethylamine	2.73 (s)	↓	Choline metabolism
Glycine	3.56 (s)	↓	Amino acid
Phenylalanine	3.20 (dd), 3.29 (dd), 3.99 (dd), 7.34 (d), 7.38 (m), 7.43 (m)	↓	Amino acid
Tyrosine	6.91 (d), 7.21 (d)	↓	Amino acid
<35% controls Vs >100% athletes [Model 3 – Diet healthiness & Exercise]			

Faecal Metabolites			
Acetate	1.92 (s)	↑	SCFA
Anserine	7.17 (s), 8.35 (s), 3.65 (s)	↑	Dipeptide
Bile salts 79	0.79 (s)	↑	Bile acids
Butyrate	0.90 (t), 1.56 (m), 2.16 (t) + Targeted GC-MS	↑	SCFA
Methylamine	2.60 (s)	↑	Choline metabolism
Propionate	1.06 (t), 2.19 (q) + Targeted GC-MS	↑	SCFA
Trimethylamine	2.88 (s)	↑	Choline metabolism
Valerate	0.89 (t), 1.30 (m), 1.54 (m), 2.18 (t) + Targeted GC-MS	↑	SCFA
2-hydroxybutyrate	1.20 (d)	↓	Glutathione synthesis
Proline	2.01 (m), 2.35 (m), 4.14 (dd)	↓	Amino acid
n-acetylglucosamine	2.06 (s), 3.47-3.85 (m), 5.21 (d)	↓	Amino sugar, bacterial cell wall constituent
Glycine	3.56 (s)	↓	Amino acid
Bacterial Diversity (Shannon / Simpson / PD) [Model 4 - Diversity]			
Bile Salts 76	0.76 (s)	↑	Bile acids
2-methylbutyrate	0.86 (t), 1.05 (d), 1.39 (m), 1.5 (m), 2.21 (m)	↑	Branched SCFA
Caproate	0.86 (t), 1.27 (m), 1.53 (m), 2.17 (t)	↑	MCFA
Valerate	0.89 (t), 1.30 (m), 1.54 (m), 2.18 (t)	↑	SCFA
Isovalerate	0.91 (d), 1.96 (m), 2.06 (d)	↑	Branched SCFA
α-ketoisocaproate (2-oxoisocaproate)	0.94 (d), 2.10 (m), 2.62 (d)	↑	BCAA degradation
Isobutyrate	1.07 (d), 2.35 (m)	↑	Branched SCFA
α-ketoisovalerate (2-oxoisovalerate)	1.13 (d), 3.02 (m)	↑	BCAA degradation
Glutamate	2.05 (m), 2.12 (m), 2.35 (m), 3.75 (dd)	↑	Amino acid
Malate	2.36 (dd), 2.67 (dd), 4.32 (dd)	↑	TCA cycle
Ureidopropionate	2.39 (t), 3.29 (q)*	↑	Uracil metabolism
3-phenylpropionate (Hydrocinnamate)	2.49 (t), 2.90 (t), 7.30 (m)	↑	Phenylalanine metabolism
Dimethylamine	2.73 (s)	↑	Choline metabolism
2-methylproline	2.75 (s), [3.60-61 (s)]	↑	Arginine & Proline metabolism
Phenylacetate	3.54 (s), 7.31 (m), 7.39 (m)	↑	Phenylalanine metabolism
Isoleucine	1.02 (d), 3.68 (d)	↓	BCAA
Asparagine	2.95 (dd), 2.87 (dd), 4.01 (dd)	↓	Amino acid
Glucose	3.25 (dd), 3.42 (m), 3.49 (m), 3.54 (dd), 3.74 (m), 3.84 (m), 3.91 (dd), 4.65 (d), 5.24 (d)	↓	Sugars
Uridine	3.80 (dd)*, 3.91 (dd)*, 4.12 (m)*, 4.22 (dd)*, 4.34 (dd)*, 5.91 (d), 5.92 (d),	↓	Nucleoside

Faecal Metabolites			
	7.88 (d)		
Histidine	7.11 (s), 7.93 (s)	↓	Amino acid
Nicotinate	7.53 (dd)*, 8.26 (m), 8.62 (dd), 8.94 (d)	↓	Vit B3

Supplementary Table 4.2: Faecal metabolites discriminating between model classes. Tables show faecal metabolites increasing or decreasing with increased diet healthiness, exercise, combined diet healthiness and exercise, and gut bacterial diversity in models 1, 2, 3 and 4 respectively. [a]Multiplicity key is as follows: s=singlet, d=doublet, t=triplet, q=quartet, dd=doublet of doublets, m=multiplet. ^{1}H shifts marked by an asterisk (*) were not confirmed experimentally due to the absence of unambiguous resonance values and are taken from the Human Metabolome Database (HMDB)

Chapter 5

COMPUTATIONAL ELUCIDATION OF MICROBIAL DYNAMICS IN THE GUT MICROBIOME OF ELITE ATHLETES

5.1 | Abstract

Strong influences have been illustrated for the gut microbiome's role in numerous aspects of human health. Previously we demonstrated that the gut microbiome of high-performance athletes is distinct from less active healthy controls, both in terms of taxonomic composition and functional potential. The robust differences in the athlete microbiome are intriguing, however the high-intensity physical activity of these individuals is paired with dietary habits that sharply deviate from what is typical in the general population. Accordingly, separating the influences of the athletes' diet from their levels of exercise is an important challenge to overcome. Here we continue the analysis of this cohort through the use of an advanced computational modelling approach based on Flux Balance Analysis (FBA), providing an opportunity to build upon prior observations of the athlete microbiome. Simultaneously, the FBA based analysis, due to its capacity to modify nutritional resources (i.e. diet) in simulations of gut microbiota metabolism, is demonstrated as a relevant, and indeed, valuable approach in generating novel insights from metagenomic studies.

In addition to the author, the original material subject of this chapter was contributed to accordingly:

Guidance on modelling approach: Federico Baldini, Dr. Almut Heinkin, & Professor. Ines Thiele

Metabolomic processing and data analysis: Dr. Nicholas C. Penney, Dr. Isabel Garcia Perez, & Professor. Elaine Holmes

Presented material has not been published at the time of this thesis' submission.

5.2 | Introduction

Physical exercise has been shown to confer a broad range health benefits, specifically affecting systemic and end-organ inflammation regulation, efficiency of carbohydrate metabolism, and trophic effects at the level of the central nervous system [1-7]. Accordingly, physical activity has been implemented as an effective treatment and preventative strategy for many chronic conditions in which the gut microbiome has been implicated [8-10]. Sedentary lifestyle is alternatively a major morbidity factor in Western society, contributing to numerous *diseases of affluence*, such as obesity, diabetes, and cardiovascular disease [11-14]. The gut microbiome is shown with mounting evidence to have an integral role in these diseases [15-23].

Described in previous chapters, the gut microbiome of high-performance athletes has high taxonomic diversity and a metabolomic profile associated with health [24]. This evidence, resulting from combined shotgun metagenomic sequencing and metabolomic analysis, was put forth after initial 16s rRNA amplicon sequencing of the same cohort of elite international-level rugby players and controls matched for age and BMI range (>28 kg/m^2 and <25 kg/m^2) [25]. In this present study, we continue the analysis of this cohort through the use of an advanced computational modelling approach. Here, Flux Balance Analysis (FBA) provides an opportunity to build upon prior observations of the athlete microbiome. In particular, our previous analysis of the athletes highlighted the challenge of separating the influences of exercise and associated diet on the microbiome. The FBA method addresses this, while simultaneously demonstrating the relevance of modelling approaches in generating novel insights from metagenomic studies.

5.3 | Study objective

Having previously demonstrated with a multi-omics led interrogation that elite athletes present a gut microbiome with metabolic function primed for health and greater phylogenetic diversity, it is clear that these athletes are suitable for further investigation. The primary objective of this chapter is to further profile the gut microbiome of high-performance athletes with computational based elaboration via flux balance analysis (FBA) and the AGORA pipeline. The modelling approach permits manipulation of dietary constraints, subsequently allowing for the separation of influences exerted by diet and exercise. Additionally, this chapter aims to support evidence for the use of modelling approaches in gaining novel insights from previously generated metagenomic data, providing firm basis for subsequent experimentation.

5.4 | Methods

5.4.1 | Study population

Elite professional male athletes (n = 40) and healthy controls (n = 46) matched for age and gender were enrolled in 2011 as described in section 3.4.1 the initial study [25]. Approval for this study was granted by the Cork Clinical Research Ethics Committee.

5.4.2 | Acquisition of clinical, exercise and dietary data

Self-reported dietary intake information was accommodated by a research nutritionist within the parameters of a food frequency questionnaire (FFQ) in conjunction with a photographic food atlas as per sections 2.3.1 and 3.4.2, as well as the initial investigation [25]. Fasting blood samples of all participants and physical activity levels of controls were collected, as detailed in sections 3.4.2 and 2.3.2 respectively. Creatine kinase levels were used as a proxy for level of physical activity across all groups.

5.4.3 | Preparation of Metagenomic libraries

DNA derived from faecal samples was extracted and purified using the QIAmp DNA Stool Mini Kit (cat. no. 51504) prior to storage at -80°C. DNA libraries were prepared with the Nextera XT DNA Library Kit (cat. no. FC-131-1096) prior to processing on the Illumina HiSeq 2500 sequencing platform (see section 2.5 for further detail).

5.4.4 | Metagenomic statistical and bioinformatic analysis

Presented more elaborately in section 2.6; delivered raw FASTQ sequence files were quality checked as follows: contaminating sequences of human origin were first removed through the NCBI Best Match Tagger (BMTagger). Poor quality and duplicate read removal, as well as trimming was implemented using a combination of SAM and Picard tools. Refined reads were functionally profiled by the Human Microbiome Project (HMP) Unified Metabolic Analysis Network (HUMAnN2 V.0.5.0) pipeline [26]. Microbial pathway data was statistically analysed in the R software environment (V.3.2.2) (for further details see sections 2.6 and 2.8.3) [27]. All presented p values were corrected for multiple comparisons using the Benjamini-Hochberg False Discovery Rate (pFDR) method [28].

5.4.5 | Microbiome modelling

Taxonomic profiles previously generated with the Metagenomic Phylogenetic Analysis (MetaPhlAn2) software [29] were manually filtered against the Assembly of Gut Organisms through Reconstruction and Analysis (AGORA) database of genome-scale metabolic reconstructions of gut microbiota [30]. The resulting compositional profiles consisted of relative abundances for taxa that had been annotated in the AGORA database. As nutritional input to the metabolic models, that is, the resources utilised by AGORA taxa in simulations of microbial metabolic exchange, dietary information from multiple sources, was used to create the following nutrition constructions: personalised; directly derived from the participants' FFQ data, standard; based on EU averages derived from a large-scale dietary survey, and rich; an enrichment of all nutritional constituents to higher than average levels. Standard diet construction was based on the European average as delineated by a 2012 Austrian survey [31]. Modelling was performed with the Microbiome Modelling Toolbox (preprint) [32], an extension of the constraint-based reconstruction and analysis (COBRA) MATLAB software suite [arXiv:1710.04038]. Described briefly; microbial profiles were generated by matching taxa to AGORA constructions and combined with dietary micronutrients to produce microbe-nutrient reactions, finally providing flux of uptake and secretion of metabolites.

5.4.6 | Metabolic profiling

Urine and faecal samples were prepared for metabonomic analysis as previously described [33, 34]. Refer to section 2.7 and 2.8 for further detail.

5.5 | Results

5.5.1 | Community dynamics captured by AGORA models

Of the 495 species detected from the original taxonomic profiling of both athletes and controls, 309 were present in the AGORA database. Measurements of α-diversity for all species and those represented in AGORA showed similar diversity levels for participants in all of the groups (Figure 5.1 A & B). Variance in α-diversity (percent Δ) between the original and AGORA taxonomic profiles demonstrated minimal reduction of diversity, with similar alteration between controls and athletes (Figure 5.1 C). Presented with non-metric multidimensional scaling (NMDS), β-diversity of all detected species and AGORA models illustrated similar clustering of the participants (Figure 5.1 D & E).

The 19,300 total metabolic pathways generated by metagenomic sequencing of faecal samples were separated into general pathways (e.g. Biotin biosynthesis) and species-specific pathways (e.g. Biotin biosynthesis in *Bacteroides vulgatus*). Comparison of these metabolic pathway models to FBA reactions generated from the personalised diet of participants revealed similar group clustering with NMDS (Figure 5.2 A, C, & D). Measurements of α-diversity were produced for general pathways, species-specific pathways, and FBA reactions (Figure 5.2 B, D, & F). Median values for the separate experimental groups illustrated similar ranking of median values for species-specific pathways and reactions, with the low BMI control group presenting the highest median values (Figure 5.2 D & E).

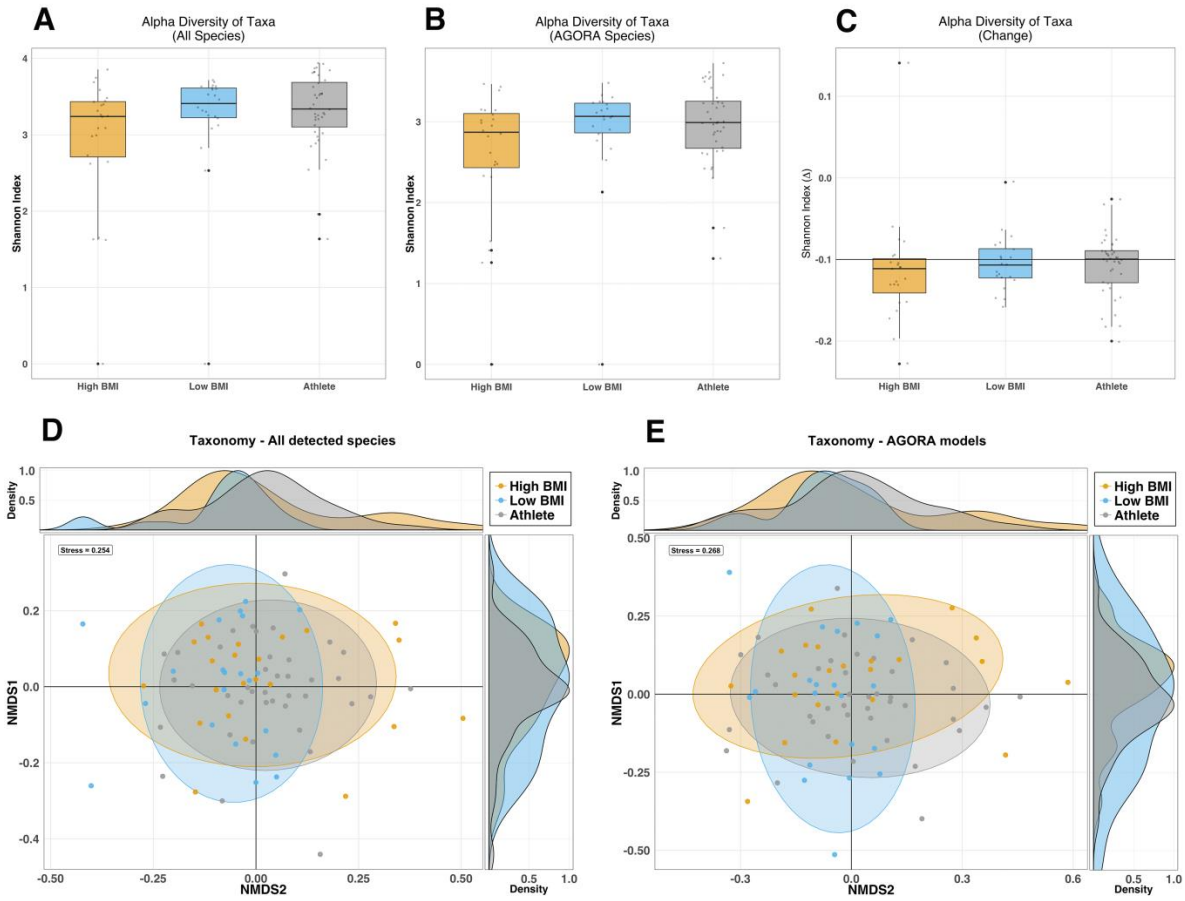


Figure 5.1 | Comparison of complete taxonomic profiles with selected AGORA models. Diversity comparisons show retention of group dynamics originally described with the complete taxonomic profile with the species represented in the AGORA models. (**A - C**) α-diversity presented as Shannon H-index for the original taxonomic profile of all species (**A**), only species represented by AGORA models (**B**), and the difference between them (percent Δ, **C**), demonstrated high fidelity of original α-diversity dynamics (i.e. with all detected species) for all groups with the AGORA models. β-diversity of taxonomic profiles of all species (**D**) and just those in the AGORA models (**E**), presented with non-metric multidimensional scaling (NMDS). NMDS shows similar clustering of participants in both profiles. Density plots along the axes of all panels display the concentrations of data points, and were generated from kernel density estimates and scaled to a maximum value of 1.

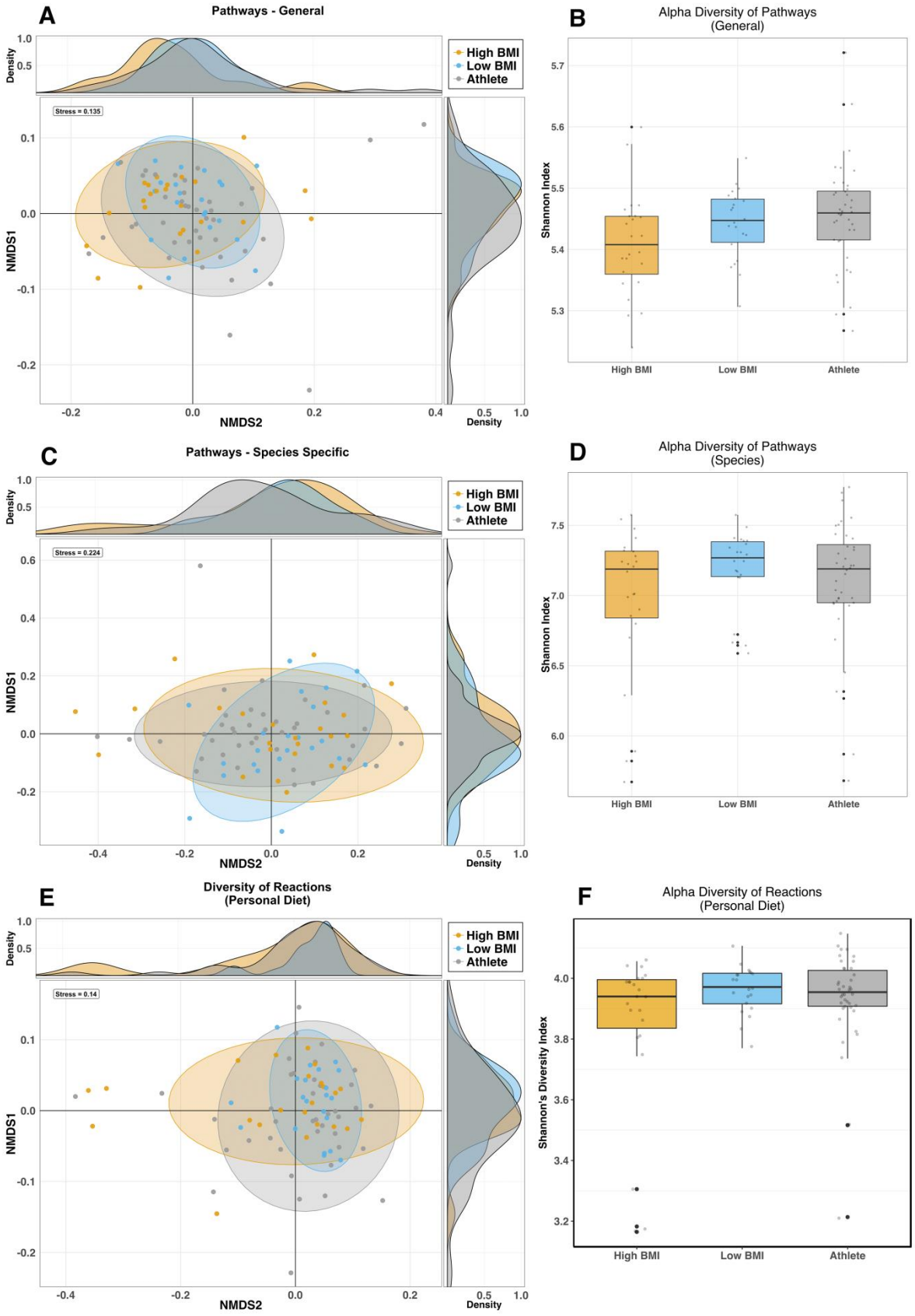


Figure 5.2 | Comparison between metabolic pathways and FBA reactions.

Figure 5.2 continued | Comparison of diversity between metabolic pathways and FBA reactions. Diversity of model reactions generated from personalised diet recapitulate elements of metabolic pathway diversity. (**A, C, & E**) β-diversity presented with NMDS for general pathways (**A**), pathways that are specific to taxa (**C**), and exchange reactions yielded from the modelling approach (**E**), show variations in clustering. Species-specific pathways project the greatest inter-participant variation, while ordination of general pathways and exchange reactions show similar clustering of participants within their respective groups. α-diversity presented as Shannon H-index for general pathways (**B**), those that are specific to taxa (**D**), and exchange reactions (**F**), similarly demonstrates that species-specific pathways have the highest levels of diversity.

5.5.2 | Influence of dietary nutrition construction on reactions

Diversity of reactions from AGORA species in response to the different nutrition constructions, illustrates a high degree of similarity between personalised and standard diets, while the rich diet eliminates separation between participants (Figure 5.3 A). Similarly, α-diversity measurements of reactions produced from the diets show near identical values for each group under the rich diet while the other two diets capture a shared variation between the experimental groups (Figure 5.3 B-D).

5.5.3 | Reaction variation between athletes and controls

Utilising the personalised diet reconstruction, a total of 134 reactions were shown to be significantly varied between the groups (pFDR < 0.05, Supplementary Table 5.1). Group comparisons demonstrated that the highest number of significantly varied reactions was between low BMI controls and athletes, with 94 reactions (Figure 5.4 A). Intriguingly, the reaction for trimethylamine N-oxide (TMAO) was significantly varied between the groups. The reaction's variance was significant between athletes and low BMI controls and between the two control groups (pFDR < 0.001), but not between the athletes and high BMI group. This suggests similar levels of the compound between athletes and high BMI controls within the simulation. Group separation according to reactions was further illustrated with Principal Component Analysis (PCA) supplemented with a Correspondence Analysis (CA) and k-Nearest Neighbour (k-NN) (Figure 5.4 B & C). This semi-supervised learning method demonstrated that clusters of reactions separated together within specific groups. No significant correlations were identified between reactions and previously generated faecal metabolites (Spearman's rank-order, data not shown).

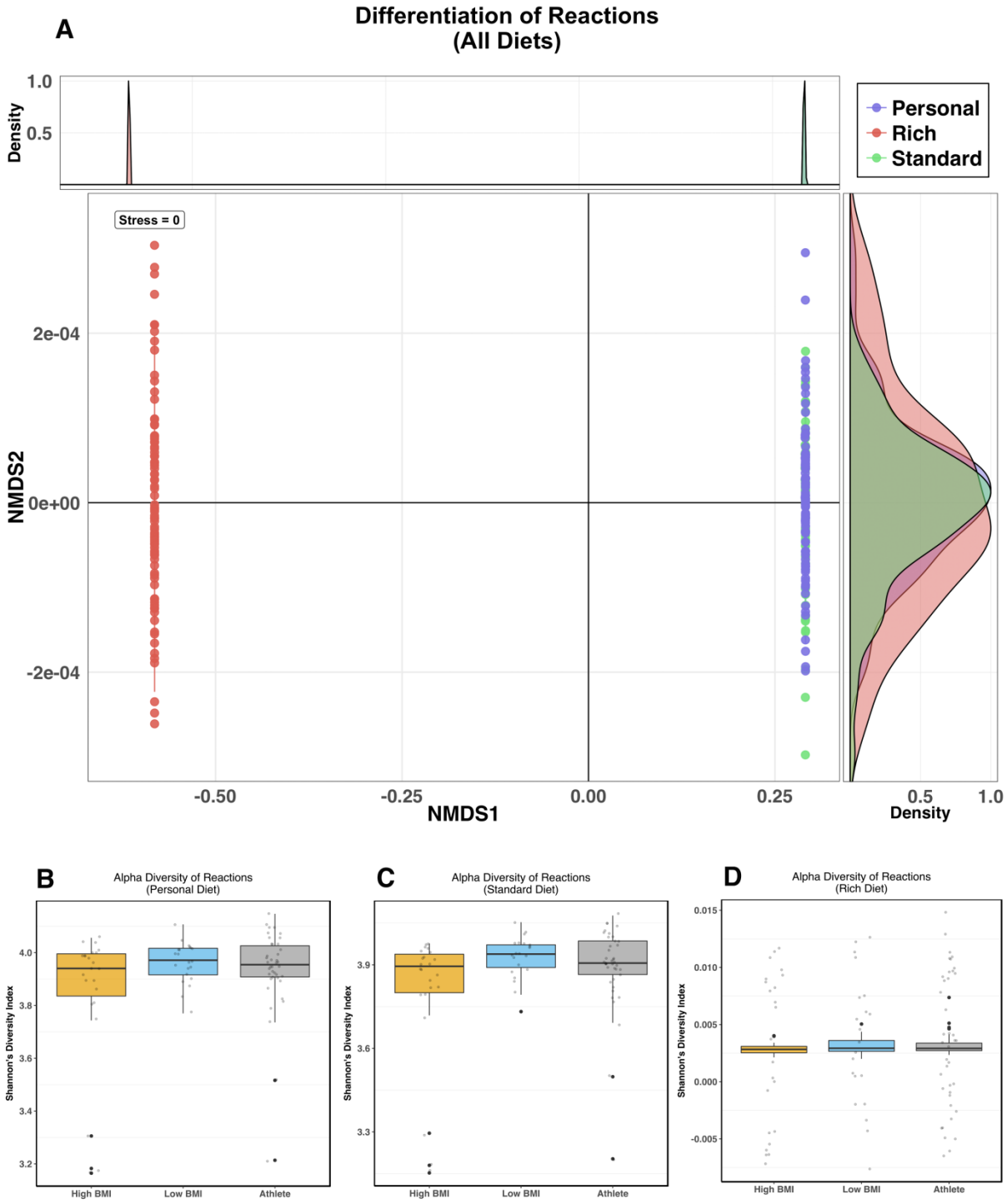


Figure 5.3 | Influence of nutritional inputs on reactions.

Figure 5.3 continued | Influence of nutritional inputs on reactions. Variations in exchange reactions, as described in terms of diversity measurements, results from the different input diets used in the metabolic simulations. (**A**) β-diversity of reactions resulting from personalised, standard, and rich diets are plotted with NMDS, showing both an extreme divergence between reactions from the rich and the remaining diets, and high congruence of personal and standard diet reactions. (**B - D**) α-diversity presented as Shannon H-index for reactions produced from personal (**B**), standard (**C**), and rich diets (**D**), presents highly similar values for each group with the personal and standard diets, while minimal inter-group variability is present in the rich diet. Density plots along the axes of all panels display the concentrations of data points, and were generated from kernel density estimates and scaled to a maximum value of 1.

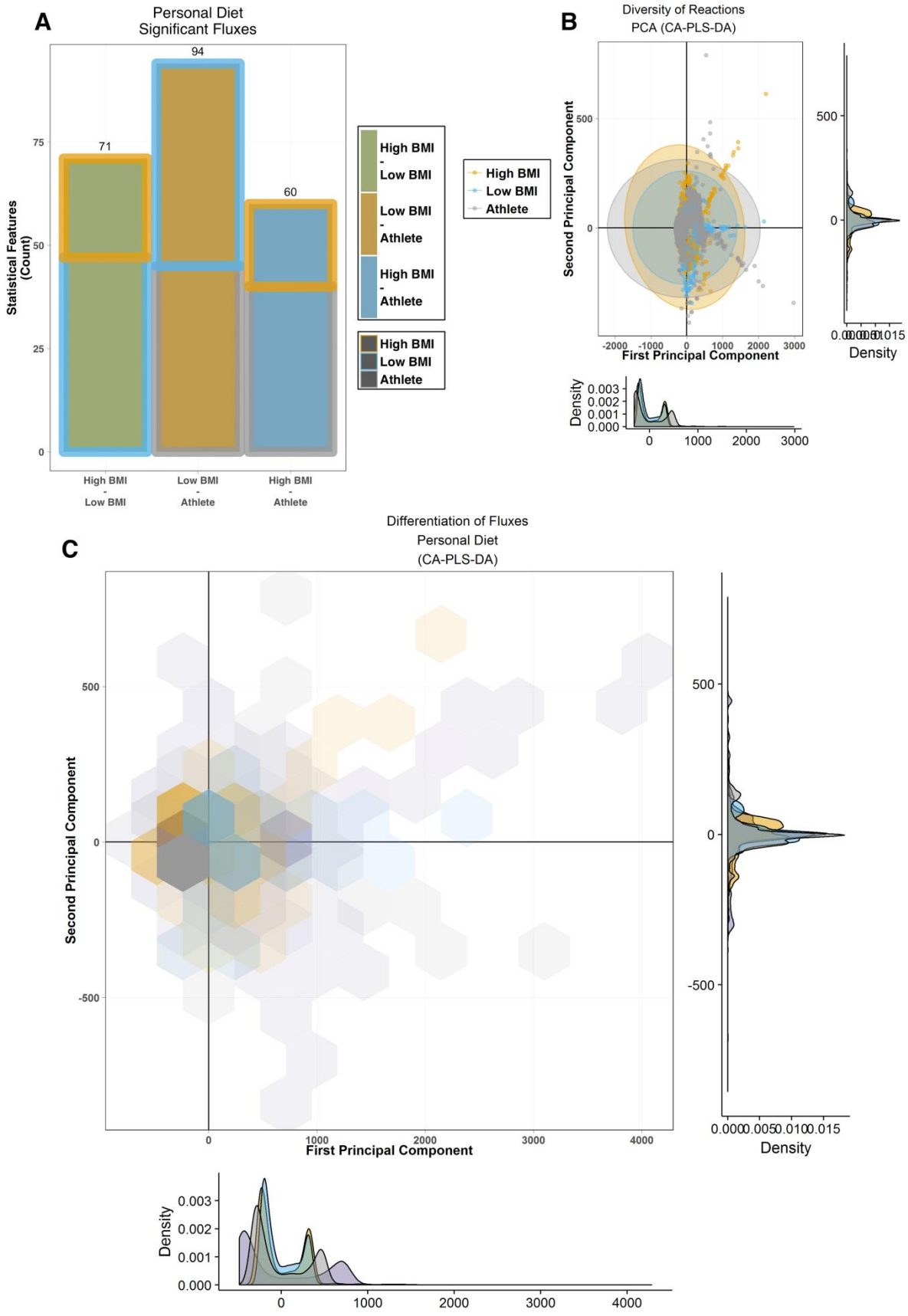


Figure 5.4 | Reaction variation between athletes and controls.

Figure 5.4 continued | Reaction variation between athletes and controls**.** Variations in exchange reactions illustrate differences between participant groups. (**A**) Number of significantly varied reactions between pairs of groups illustrates that the greatest variation (i.e. the most number of significant differences) occurs between athletes and low BMI controls (See Table 5.1 for complete list of significantly altered reactions). Within each test pair (e.g. high BMI – low BMI), coloured borders correspond to the number of reactions that were higher in the respective group. (**B**) The Principal Component Analysis (PCA) procedure, Partial Least Squares-Discriminant Analysis (PLS-DA) with Correspondence Analysis (CA) illustrates group separation according to reactions. (**C**) With the inclusion of combined reactions for all participants, group specificity of reaction clustering is further demonstrated with hexagonal binning of plotted points. Density plots along the axes of all panels display the concentrations of data points. Statistical testing performed with Wilcoxon signed-rank test. Significance determined with a cut-off of pFDR < 0.05(Benjimini-Hochberg).

5.6 | Discussion

To continue elucidation of variations in the athlete gut microbiome, this present chapter applied Flux Balance Analysis (FBA) based modelling approaches to a collection of previously generated metagenomic sequences from the gut microbiome of high-performance athletes and controls. Previous investigation of the elite athlete microbiome illustrated an increase of microbial diversity in athletes when compared to controls [24]. Furthermore, the athletes were shown to have increased relative abundance of metabolic pathways that have potential for contributing to health. Indeed, profiling of the athlete gut environment with metabolomic analysis revealed an enrichment of SCFAs, compounds with association to health and lean body composition [35-37]. Trimethylamine N-oxide (TMAO), a metabolite that is an end product of dietary protein degradation, was also present at high levels in the athlete gut environment. It has been shown by others that elevated concentrations of TMAO are associated with atherosclerosis and cardiovascular disease (CVD) [15-17, 22, 38]. This association with disease is disputed however, as some populations with low occurrence of CVD also maintain high levels of excreted TMAO [39]. In terms of the athlete population, it was not initially clear if elevated TMAO was the result of the combination of their unique dietary habits and microbiome, or a by-product of either factor alone.

This present study has importantly illustrated that parameters of metagenomic sequencing based investigations of the gut microbiome are preserved in the FBA modelling. Particularly, the initial input of AGORA models recapitulates the trends in diversity observed in the complete profile of taxonomy (Figure. 5.1). Minimal reductions in α-diversity measured as percent change of Shannon H-index were consistent across athlete and control groups, illustrating equivalent representation with the AGORA models.

Similarly, β-diversity for all taxa and the AGORA subset demonstrated little variation in clustering of the participants. Examination of the separate diversity dynamics for metabolic pathways and reactions, demonstrated subject clustering that was most similar between pathways specific to species and the reactions (Figure 5.2). Such was also the case when looking at the α-diversity of pathways and reactions. A comparison of different nutritional inputs for the models showed nearly identical diversity of reactions generated from personalised and standard diets, while the rich diet uniformly affected subjects in the athlete and control groups (Figure 5.3). This highlights the accuracy of the EU standardised diet, and also the potential for using the rich diet to identify dynamics of the metabolic models for the individual participants that are not directly linked to diet (i.e. those that arise from unique microbial composition). Lastly, examination of the reactions generated demonstrates group variability that has been identified with previous investigation of the samples. Of particular note, statistical analysis of the reactions between athletes and controls revealed that the exchange reaction pertaining to TMAO was significantly varied between the athletes and low BMI control group, and between the two control groups. These results infer that the high BMI controls and athletes have similar potential for TMAO generation, which further implies that the concentrations of TMAO originally detected with metabolomic profiling, may indeed be the result of dietary intake and not influenced significantly by exercise. This consideration requires more experimental elaboration to draw conclusions, however it is warranted with these results.

5.7 | References

[1] A. Harkin, Muscling in on depression, N Engl J Med 371 (24) (2014) 2333-4.

[2] F. B. Benatti, B. K. Pedersen, Exercise as an anti-inflammatory therapy for rheumatic diseases-myokine regulation, Nat Rev Rheumatol 11 (2) (2015) 86-97.

[3] J. A. Hawley, A. Krook, Metabolism: One step forward for exercise, Nat Rev Endocrinol 12 (1) (2016) 7-8.

[4] L. Hoffman-Goetz, N. Pervaiz, N. Packer, J. Guan, Freewheel training decreases pro- and increases anti-inflammatory cytokine expression in mouse intestinal lymphocytes, Brain Behav Immun 24 (7) (2010) 1105-15.

[5] W. Barton, F. Shanahan, P. D. Cotter, O. O'Sullivan, The metabolic role of the microbiota, Clinical Liver Disease 5 (4) (2015) 91-93.

[6] K. L. Szuhany, M. Bugatti, M. W. Otto, A meta-analytic review of the effects of exercise on brain-derived neurotrophic factor, J Psychiatr Res 60 (2015) 56-64.

[7] S. M. Ryan, Y. M. Nolan, Neuroinflammation negatively affects adult hippocampal neurogenesis and cognition: can exercise compensate?, Neurosci Biobehav Rev 61 (2016) 121-31.

[8] E. Johannesson, M. Simren, H. Strid, A. Bajor, R. Sadik, Physical activity improves symptoms in irritable bowel syndrome: a randomized controlled trial, Am J Gastroenterol 106 (5) (2011) 915-22.

[9] T. E. Robsahm, B. Aagnes, A. Hjartaker, H. Langseth, F. I. Bray, I. K. Larsen, Body mass index, physical activity, and colorectal cancer by anatomical subsites: a systematic review and meta-analysis of cohort studies, Eur J Cancer Prev 22 (6) (2013) 492-505.

[10] L. Schwingshackl, B. Missbach, S. Dias, J. Konig, G. Hoffmann, Impact of different training modalities on glycaemic control and blood lipids in patients with type 2 diabetes: a systematic review and network meta-analysis, Diabetologia 57 (9) (2014) 1789-97.

[11] A. Biswas, P. I. Oh, G. E. Faulkner, R. R. Bajaj, M. A. Silver, M. S. Mitchell, D. A. Alter, Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis, Ann Intern Med 162 (2) (2015) 123-32.

[12] R. V. Same, D. I. Feldman, N. Shah, S. S. Martin, M. Al Rifai, M. J. Blaha, G. Graham, H. M. Ahmed, Relationship Between Sedentary Behavior and Cardiovascular Risk, Curr Cardiol Rep 18 (1) (2016) 6.

[13] E. G. Wilmot, C. L. Edwardson, F. A. Achana, M. J. Davies, T. Gorely, L. J. Gray, K. Khunti, T. Yates, S. J. Biddle, Sedentary time in adults and the association with diabetes, cardiovascular disease and death: systematic review and meta-analysis, Diabetologia 55 (11) (2012) 2895-905.

[14] Y. C. Chen, Y. K. Tu, K. C. Huang, P. C. Chen, D. C. Chu, Y. L. Lee, Pathway from central obesity to childhood asthma. Physical fitness and sedentary time are leading factors, Am J Respir Crit Care Med 189 (10) (2014) 1194-203.

[15] R. A. Koeth, Z. Wang, B. S. Levison, J. A. Buffa, E. Org, B. T. Sheehy, E. B. Britt, X. Fu, Y. Wu, L. Li, J. D. Smith, J. A. DiDonato, J. Chen, H. Li, G. D. Wu, J. D. Lewis, M. Warrier, J. M. Brown, R. M. Krauss, W. H. Tang, F. D. Bushman, A. J. Lusis, S. L. Hazen, Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis, Nat Med 19 (5) (2013) 576-85.

[16] W. H. W. Tang, Z. E. Wang, B. S. Levison, R. A. Koeth, E. B. Britt, X. M. Fu, Y. P. Wu, S. L. Hazen, Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk, New England Journal of Medicine 368 (17) (2013) 1575-1584.

[17] W. H. W. Tang, S. L. Hazen, Microbiome, trimethylamine N-oxide, and cardiometabolic disease, Translational Research (2016)

[18] A. Woting, N. Pfeiffer, G. Loh, S. Klaus, M. Blaut, Clostridium ramosum promotes high-fat diet-induced obesity in gnotobiotic mouse models, MBio 5 (5) (2014) e01530-14.

[19] K. M. Utzschneider, M. Kratz, C. J. Damman, M. Hullar, Mechanisms linking the gut microbiome and glucose metabolism., The Journal of clinical endocrinology and metabolism 101 (4) (2016) jc20154251.

[20] P. J. Turnbaugh, R. E. Ley, M. A. Mahowald, V. Magrini, E. R. Mardis, J. I. Gordon, An obesity-associated gut microbiome with increased capacity for energy harvest, Nature 444 (7122) (2006) 1027-31.

[21] N. C. Williams, M. A. Johnson, D. E. Shaw, I. Spendlove, J. Vulevic, G. R. Sharpe, K. A. Hunter, A prebiotic galactooligosaccharide mixture reduces severity of hyperpnoea-induced bronchoconstriction and markers of airway inflammation, Br J Nutr 116 (5) (2016) 798-804.

[22] Z. Wang, E. Klipfell, B. J. Bennett, R. Koeth, B. S. Levison, B. Dugar, A. E. Feldstein, E. B. Britt, X. Fu, Y. M. Chung, Y. Wu, P. Schauer, J. D. Smith, H. Allayee, W. H. Tang, J. A. DiDonato, A. J. Lusis, S. L. Hazen, Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease, Nature 472 (7341) (2011) 57-63.

[23] A. Zhernakova, A. Kurilshikov, M. J. Bonder, E. F. Tigchelaar, M. Schirmer, T. Vatanen, Z. Mujagic, A. V. Vila, G. Falony, S. Vieira-Silva, J. Wang, F. Imhann, E. Brandsma, S. A. Jankipersadsing, M. Joossens, M. C. Cenit, P. Deelen, M. A. Swertz, s. LifeLines cohort, R. K. Weersma, E. J. Feskens, M. G. Netea, D. Gevers, D. Jonkers, L. Franke, Y. S. Aulchenko, C. Huttenhower, J. Raes, M. H. Hofker, R. J. Xavier, C. Wijmenga, J. Fu, Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity, Science 352 (6285) (2016) 565-9.

[24] W. Barton, N. C. Penney, O. Cronin, I. Garcia-Perez, M. G. Molloy, E. Holmes, F. Shanahan, P. D. Cotter, O. O'Sullivan, The microbiome of professional athletes differs from that of

more sedentary subjects in composition and particularly at the functional metabolic level, Gut (2017)

[25] S. F. Clarke, E. F. Murphy, O. O'Sullivan, A. J. Lucey, M. Humphreys, A. Hogan, P. Hayes, M. O'Reilly, I. B. Jeffery, R. Wood-Martin, D. M. Kerins, E. Quigley, R. P. Ross, P. W. O'Toole, M. G. Molloy, E. Falvey, F. Shanahan, P. D. Cotter, Exercise and associated dietary extremes impact on gut microbial diversity, Gut 63 (12) (2014) 1913-20.

[26] S. Abubucker, N. Segata, J. Goll, A. M. Schubert, J. Izard, B. L. Cantarel, B. Rodriguez-Mueller, J. Zucker, M. Thiagarajan, B. Henrissat, O. White, S. T. Kelley, B. Methe, P. D. Schloss, D. Gevers, M. Mitreva, C. Huttenhower, Metabolic reconstruction for metagenomic data and its application to the human microbiome, PLoS Comput Biol 8 (6) (2012) e1002358.

[27] R Development Core Team, R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna, 2012), URL: http:// www. R-project. org (2015)

[28] Y. Benjamini, Y. Hochberg, Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing, Journal of the Royal Statistical Society Series B-Methodological 57 (1) (1995) 289-300.

[29] N. Segata, L. Waldron, A. Ballarini, V. Narasimhan, O. Jousson, C. Huttenhower, Metagenomic microbial community profiling using unique clade-specific marker genes, Nat Methods 9 (8) (2012) 811-4.

[30] S. Magnusdottir, A. Heinken, L. Kutt, D. A. Ravcheev, E. Bauer, A. Noronha, K. Greenhalgh, C. Jager, J. Baginska, P. Wilmes, R. M. Fleming, I. Thiele, Generation of genome-scale metabolic reconstructions for 773 members of the human gut microbiota, Nat Biotechnol 35 (1) (2017) 81-89.

[31] H. V. Elmadfa I, Wagner K, Oesterreichischer Ernaehrungsbericht, (2012)

[32] F. Baldini, A. Heinken, L. Heirendt, S. Magnusdottir, R. M. T. Fleming, I. Thiele, The Microbiome Modeling Toolbox: from microbial interactions to personalized microbial communities, (2018)

[33] A. C. Dona, B. Jimenez, H. Schafer, E. Humpfer, M. Spraul, M. R. Lewis, J. T. Pearce, E. Holmes, J. C. Lindon, J. K. Nicholson, Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping, Anal Chem 86 (19) (2014) 9887-94.

[34] R. Garcia-Villalba, J. A. Gimenez-Bastida, M. T. Garcia-Conesa, F. A. Tomas-Barberan, J. Carlos Espin, M. Larrosa, Alternative method for gas chromatography-mass spectrometry analysis of short-chain fatty acids in faecal samples, J Sep Sci 35 (15) (2012) 1906-13.

[35] A. Koh, F. De Vadder, P. Kovatcheva-Datchary, F. Backhed, From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites, Cell 165 (6) (2016) 1332-1345.

[36] V. K. Ridaura, J. J. Faith, F. E. Rey, J. Cheng, A. E. Duncan, A. L. Kau, N. W. Griffin, V. Lombard, B. Henrissat, J. R. Bain, M. J. Muehlbauer, O. Ilkayeva, C. F. Semenkovich, K. Funai, D. K. Hayashi, B. J. Lyle, M. C. Martini, L. K. Ursell, J. C. Clemente, W. Van Treuren, W. A. Walters, R. Knight, C. B. Newgard, A. C. Heath, J. I. Gordon, Gut microbiota from twins discordant for obesity modulate metabolism in mice, Science 341 (6150) (2013) 1241214.

[37] H. M. Hamer, D. M. Jonkers, A. Bast, S. A. Vanhoutvin, M. A. Fischer, A. Kodde, F. J. Troost, K. Venema, R. J. Brummer, Butyrate modulates oxidative stress in the colonic mucosa of healthy humans, Clin Nutr 28 (1) (2009) 88-93.

[38] B. J. Bennett, T. Q. de Aguiar Vallim, Z. Wang, D. M. Shih, Y. Meng, J. Gregory, H. Allayee, R. Lee, M. Graham, R. Crooke, P. A. Edwards, S. L. Hazen, A. J. Lusis, Trimethylamine-N-oxide, a metabolite associated with atherosclerosis, exhibits complex genetic and dietary regulation, Cell Metab 17 (1) (2013) 49-60.

[39] E. Holmes, R. L. Loo, J. Stamler, M. Bictash, I. K. Yap, Q. Chan, T. Ebbels, M. De Iorio, I. J. Brown, K. A. Veselkov, M. L. Daviglus, H. Kesteloot, H. Ueshima, L. Zhao, J. K. Nicholson, P. Elliott, Human metabolic phenotype diversity and its association with diet and blood pressure, Nature 453 (7193) (2008) 396-400.

5.8 | Supplementary Content

Reaction	Adjusted *P* value		
	High BMI--Low BMI	Low BMI--Athlete	High BMI--Athlete
'EX_pro_L[fe]'	0.00E+00	4.00E-04	0.00E+00
'EX_pyr[fe]'	0.00E+00	3.40E-03	0.00E+00
'EX_fum[fe]'	0.00E+00	6.32E-02	0.00E+00
'EX_glyc[fe]'	0.00E+00	4.93E-01	0.00E+00
'EX_meoh[fe]'	0.00E+00	6.24E-01	0.00E+00
'EX_pydxn[fe]'	2.00E-04	2.34E-01	0.00E+00
'EX_csn[fe]'	2.00E-04	2.86E-01	0.00E+00
'EX_gln_L[fe]'	2.00E-04	4.82E-01	0.00E+00
'EX_2hyoxplac[fe]'	2.00E-04	5.07E-01	0.00E+00
'EX_mnl[fe]'	2.00E-04	6.33E-01	0.00E+00
'EX_ile_L[fe]'	1.00E-03	2.43E-01	0.00E+00
'EX_glcn[fe]'	1.00E-03	3.00E-01	0.00E+00
'EX_lys_L[fe]'	4.40E-03	0.00E+00	0.00E+00
'EX_ura[fe]'	1.25E-01	4.00E-04	0.00E+00
'EX_tyr_L[fe]'	1.45E-01	9.60E-03	0.00E+00
'EX_asp_L[fe]'	3.51E-01	3.42E-02	0.00E+00
'EX_ribflv[fe]'	4.34E-01	0.00E+00	0.00E+00
'EX_amp[fe]'	4.53E-01	0.00E+00	0.00E+00
'EX_thm[fe]'	4.67E-01	2.22E-01	0.00E+00
'EX_asn_L[fe]'	5.27E-01	0.00E+00	0.00E+00
'EX_Lcystin[fe]'	5.67E-01	3.80E-03	0.00E+00
'EX_actn_R[fe]'	6.34E-01	0.00E+00	0.00E+00
'EX_cys_L[fe]'	6.96E-01	0.00E+00	0.00E+00
'EX_diact[fe]'	7.14E-01	0.00E+00	0.00E+00
'EX_etoh[fe]'	7.23E-01	2.80E-03	0.00E+00
'EX_ppa[fe]'	8.08E-01	0.00E+00	0.00E+00
'EX_succ[fe]'	8.52E-01	0.00E+00	0.00E+00
'EX_spmd[fe]'	1.81E-01	6.84E-01	2.00E-04
'EX_acald[fe]'	4.81E-01	2.34E-01	2.00E-04
'EX_ala_D[fe]'	5.79E-01	1.52E-01	2.00E-04
'EX_cit[fe]'	6.86E-01	4.00E-04	2.00E-04
'EX_rbflvrd[fe]'	8.23E-01	4.00E-04	2.00E-04
'EX_h2o[fe]'	4.40E-03	6.06E-01	4.00E-04
'EX_but[fe]'	5.82E-02	4.12E-01	4.00E-04
'EX_lac_L[fe]'	5.09E-01	8.60E-02	4.00E-04
'EX_strch2[fe]'	8.45E-01	2.34E-01	4.00E-04
'EX_ttdca[fe]'	8.66E-01	1.54E-01	6.00E-04
'EX_acmana[fe]'	1.45E-01	5.14E-01	1.00E-03
'EX_h[fe]'	2.00E-04	2.34E-01	1.20E-03
'EX_ppi[fe]'	1.80E-03	0.00E+00	1.20E-03

Reaction	Adjusted *P* value		
	High BMI--Low BMI	Low BMI--Athlete	High BMI--Athlete
'EX_fol[fe]'	4.40E-03	4.24E-01	1.20E-03
'EX_hxan[fe]'	1.65E-01	0.00E+00	1.20E-03
'EX_gly[fe]'	4.95E-01	5.83E-01	1.20E-03
'EX_pnto_R[fe]'	5.23E-01	3.75E-01	1.20E-03
'EX_pheme[fe]'	0.00E+00	4.00E-04	1.80E-03
'EX_2dmmq8[fe]'	6.08E-01	4.00E-04	1.80E-03
'EX_mqn7[fe]'	1.60E-03	6.42E-01	2.20E-03
'EX_glc_D[fe]'	7.46E-01	2.06E-01	2.20E-03
'EX_co2[fe]'	2.12E-02	5.54E-01	2.60E-03
'EX_malt[fe]'	7.60E-01	1.01E-01	2.60E-03
'EX_rib_D[fe]'	8.30E-01	1.04E-02	7.00E-03
'EX_glu_L[fe]'	1.82E-01	3.60E-01	8.40E-03
'EX_glyc3p[fe]'	7.46E-01	2.74E-01	8.40E-03
'EX_xtsn[fe]'	8.73E-01	2.34E-01	8.40E-03
'EX_12dgr180[fe]'	2.20E-03	0.00E+00	2.24E-02
'EX_malttr[fe]'	0.00E+00	4.00E-04	2.84E-02
'EX_sheme[fe]'	1.60E-03	2.59E-01	3.16E-02
'EX_urea[fe]'	5.42E-01	4.00E-04	3.45E-02
'EX_mqn8[fe]'	0.00E+00	3.40E-03	3.70E-02
'EX_q8[fe]'	0.00E+00	0.00E+00	4.56E-02
'EX_ade[fe]'	2.42E-01	0.00E+00	6.45E-02
'EX_dgsn[fe]'	2.20E-03	3.55E-01	9.52E-02
'EX_chor[fe]'	2.00E-04	0.00E+00	1.03E-01
'EX_dpcoa[fe]'	0.00E+00	2.10E-02	1.23E-01
'EX_ser_L[fe]'	2.22E-01	0.00E+00	1.24E-01
'EX_akg[fe]'	2.00E-04	4.00E-04	1.45E-01
'EX_ca2[fe]'	2.00E-04	5.80E-03	1.45E-01
'EX_cl[fe]'	2.00E-04	5.80E-03	1.62E-01
'EX_cobalt2[fe]'	2.00E-04	5.80E-03	1.62E-01
'EX_cu2[fe]'	2.00E-04	5.80E-03	1.67E-01
'EX_gcald[fe]'	0.00E+00	0.00E+00	1.83E-01
'EX_glyclt[fe]'	0.00E+00	0.00E+00	1.89E-01
'EX_k[fe]'	2.00E-04	5.80E-03	2.13E-01
'EX_ptrc[fe]'	2.00E-04	4.00E-04	2.34E-01
'EX_mg2[fe]'	2.00E-04	5.80E-03	2.34E-01
'EX_mn2[fe]'	2.00E-04	5.80E-03	2.34E-01
'EX_so4[fe]'	2.00E-04	5.80E-03	2.34E-01
'EX_zn2[fe]'	2.00E-04	5.80E-03	2.34E-01
'EX_dtmp[fe]'	1.80E-03	2.92E-02	2.34E-01
'EX_trypta[fe]'	2.20E-03	2.80E-01	2.34E-01
'EX_fe2[fe]'	2.12E-02	3.42E-02	2.34E-01
'EX_arg_L[fe]'	2.80E-01	0.00E+00	2.34E-01

	Adjusted *P* value		
Reaction	High BMI--Low BMI	Low BMI--Athlete	High BMI--Athlete
'EX_oaa[fe]'	4.31E-01	4.20E-03	2.34E-01
'EX_isobut[fe]'	4.34E-01	0.00E+00	2.34E-01
'EX_xyl_D[fe]'	5.55E-01	5.40E-03	2.34E-01
'EX_his_L[fe]'	7.46E-01	3.42E-02	2.34E-01
'EX_phe_L[fe]'	7.92E-01	4.26E-02	2.34E-01
'EX_pydx[fe]'	8.15E-01	3.80E-03	2.34E-01
'EX_met_L[fe]'	2.00E-04	4.00E-04	2.42E-01
'EX_4hbz[fe]'	1.00E-03	3.80E-03	2.61E-01
'EX_2obut[fe]'	4.40E-03	4.20E-03	2.80E-01
'EX_dad_2[fe]'	2.98E-01	0.00E+00	2.83E-01
'EX_gthox[fe]'	3.16E-01	6.01E-01	2.95E-01
'EX_adn[fe]'	2.61E-01	0.00E+00	3.13E-01
'EX_val_L[fe]'	5.27E-01	4.00E-04	3.51E-01
'EX_ispre[fe]'	7.89E-02	4.20E-03	3.60E-01
'EX_15dap[fe]'	1.45E-01	2.27E-02	3.76E-01
'EX_26dap_M[fe]'	2.00E-04	0.00E+00	3.92E-01
'EX_hista[fe]'	2.20E-03	8.25E-02	4.30E-01
'EX_nac[fe]'	7.76E-01	3.14E-01	4.44E-01
'EX_tym[fe]'	2.20E-03	1.24E-01	4.59E-01
'EX_ncam[fe]'	5.82E-02	4.40E-03	4.73E-01
'EX_nh4[fe]'	1.60E-03	4.00E-04	4.85E-01
'EX_tma[fe]'	2.00E-04	4.00E-04	4.96E-01
'EX_acgam[fe]'	2.00E-04	0.00E+00	5.34E-01
'EX_ins[fe]'	5.56E-02	0.00E+00	5.47E-01
'EX_thymd[fe]'	8.00E-03	4.00E-04	5.59E-01
'EX_4abut[fe]'	1.60E-03	0.00E+00	5.74E-01
'EX_oxa[fe]'	2.20E-03	4.00E-04	5.85E-01
'EX_ala_L[fe]'	2.20E-03	0.00E+00	6.51E-01
'EX_cgly[fe]'	2.00E-04	0.00E+00	6.92E-01
'EX_cytd[fe]'	2.42E-02	0.00E+00	7.30E-01
'EX_dcyt[fe]'	7.16E-02	0.00E+00	7.38E-01
'EX_fe3[fe]'	3.67E-01	4.20E-02	7.43E-01
'EX_fru[fe]'	0.00E+00	0.00E+00	7.52E-01
'EX_gsn[fe]'	2.02E-01	0.00E+00	7.78E-01
'EX_gua[fe]'	4.40E-03	0.00E+00	7.86E-01
'EX_h2[fe]'	1.01E-01	0.00E+00	7.95E-01
'EX_h2s[fe]'	2.00E-04	0.00E+00	8.03E-01
'EX_hdca[fe]'	3.94E-02	2.34E-01	8.11E-01
'EX_indole[fe]'	3.84E-01	1.04E-02	8.19E-01
'EX_isocapr[fe]'	1.00E-03	0.00E+00	8.27E-01
'EX_isoval[fe]'	1.00E-03	0.00E+00	8.40E-01
'EX_lac_D[fe]'	4.40E-03	0.00E+00	8.40E-01

Reaction	Adjusted *P* value		
	High BMI--Low BMI	Low BMI--Athlete	High BMI--Athlete
'EX_leu_L[fe]'	0.00E+00	0.00E+00	8.44E-01
'EX_mal_L[fe]'	0.00E+00	0.00E+00	8.52E-01
'EX_ocdca[fe]'	4.00E-01	4.00E-04	8.73E-01
'EX_orn[fe]'	0.00E+00	0.00E+00	8.88E-01
'EX_thr_L[fe]'	9.24E-02	0.00E+00	9.01E-01
'EX_tmao[fe]'	1.60E-03	4.00E-04	9.08E-01
'EX_trp_L[fe]'	4.16E-01	4.20E-03	9.15E-01
'EX_uri[fe]'	2.86E-02	4.00E-04	9.21E-01
'EX_xan[fe]'	1.81E-01	3.60E-03	9.28E-01

Supplementary Table 5.1 | Significantly altered reactions. Exchange reactions generated from personalised diet that are shown to have statistically significant variation between groups.

Chapter 6

A Prospective Metagenomic and Metabolomic Analysis of the Impact of Exercise and/or Whey Protein Supplementation on the Gut Microbiome of Sedentary Adults

6.1 | Abstract

Many components of modern living exert influence on the resident intestinal microbiota of humans with resultant impact on host health. For example, exercise-associated changes in gut microbial diversity, composition, and functional profiles have been described in cross-sectional studies of habitual athletes. However, this relationship is compounded by changes in diet that coincide with exercise such as dietary and supplementary protein consumption. To determine whether increasing physical activity and/or increased protein intake modulates gut microbial composition and function, we prospectively challenged healthy but sedentary adults with a short-term exercise regime, with and without concurrent daily whey protein consumption. Metagenomic and metabolomic-based assessments demonstrated modest changes in gut microbial composition and function following increases in physical activity. Significant changes in the diversity of the gut virome were evident in participants receiving daily whey protein supplementation. Results indicate that improved body composition with exercise is not dependent on major changes in gut microbial diversity.

In addition to the author, the original material subject of this chapter was contributed to accordingly:

Clinical work: Dr. Owen Cronin,

Metabolomic processing and data analysis: Dr. Nicholas C. Penney, Dr. Isabel Garcia Perez, & Professor Elaine Holmes

Presented material has been published at the time of this thesis' submission.

A Prospective Metagenomic and Metabolomic Analysis of the Impact of Exercise and/or Whey Protein Supplementation on the Gut Microbiome of Sedentary Adults.

Owen Cronin,[1,2*^] Wiley Barton,[1,2,3^] Peter Skuse,[3] Nicholas C. Penney,[4,5] Isabel Garcia-Perez,[4] Eileen F. Murphy,[6] Trevor Woods,[7] Helena Nugent,[2] Aine Fanning,[1] Silvia Melgar,[1] Eanna C. Falvey,[2,8] Elaine Holmes,[4] Paul D. Cotter,[2,3] Orla O'Sullivan,[2,3] Michael G. Molloy,[1,2] Fergus Shanahan.[1,2]

[1] APC Microbiome Ireland, National University of Ireland, Cork, T12 YT20, Ireland.

[2] Department of Medicine, National University of Ireland, Cork, T12 DC4A, Ireland.

[3] Teagasc Food Research Centre, Moorepark, Fermoy, P61 C996, Cork, Ireland.

[4] Section of Biomolecular Medicine, Division of Computational Systems Medicine, Department of Surgery and Cancer, Imperial College London, SW7 2AZ, United Kingdom.

[5] Division of Surgery, Department of Surgery and Cancer, Imperial College London, SW7 2AZ, United Kingdom.

[6] Alimentary Health Ltd., Cork Airport Business Park, Cork, T12 N84F, Ireland.

[7] Human Performance Laboratory, Department of Sport and Physical Activity, National University of Ireland, Cork, T12 N1FK, Ireland.

[8] Department of Sports Medicine, Sports Surgery Clinic, Northwood, Santry, Dublin 9, D09 C523, Ireland.

6.2 | Introduction

Most of the elements of human lifestyle and environment influence the composition or function of the gut microbiome [1, 2]. Indeed, the microbiome has been viewed as a transducer of nutrient and other environmental signals for the host [3]. Therefore, several investigators have begun to explore whether a sedentary lifestyle, or more specifically, exercise and fitness, is associated with changes in the gut microbiota [4-7]. This has been assessed in cross-sectional studies of habitual exercisers [8-10] and professional athletes [10-12] in addition to experimental models.

In elite athletes, distinct compositional and functional microbial characteristics are evident in the gut including increased α-diversity, enhanced microbial production of short-chain fatty acids, and greater metabolic capacity [11, 12]. These microbial features positively correlate with the athletes' levels of physical activity, in addition to the quantity of dietary protein consumed. In many professional sporting disciplines, as well as amateur sport, intentional protein supplementation (e.g. whey protein) provides a sizeable proportion of athletes' daily protein intake [12].

Evidence from animal studies highlights the potential for taxonomic manipulation of colonic microbiota following exercise interventions, both with and without concurrent dietary alterations [6, 13, 14]. Previously, we have proposed several mechanisms by which exercise and resultant fitness may directly influence the gut microbiota including effects on gastro-intestinal transit time [15], a known driver of gut microbial diversity [16, 17]. It appears that physical activity initiated in the juvenile period of development demonstrates a greater potential for fostering a preferential microbiota than exercise commenced in adulthood [6, 18].

However, in humans the relationship between exercise and alterations in the microbiome is compounded by changes in dietary consumption that often accompany physical activity e.g. increased protein supplement intake.

Building on previous work [11, 12], the present study sought to interrogate correlations between the gut microbiome and levels of physical activity and protein consumption. To do so, using a combination of next generation shotgun sequencing and metabolomic analysis, we prospectively examined the impact of exercise, with and without whey protein supplementation, on the adult human gut microbiome. We report that 8-weeks of combined aerobic and resistance training lead to modest alterations in the composition and activity of the gut microbiome of sedentary individuals. Participants consuming daily whey protein did experience a marked increase in the diversity of their gut virome following 8-weeks of oral supplementation.

6.3 | Materials and Methods

6.3.1 | Experimental models and subject details

A description of the human study model used here is outlined in Results under "Study overview" and in figure 6.1. Male and female volunteers were enrolled. The Cork Clinical Research Ethics Committee (CREC) approved the study before it commenced. Recruitment and assignment of interventions are outlined in detail below.

6.3.2 | Study recruitment and safe participation

Male and female participants, aged between 18 and 40 years (inclusive) were recruited via online, e-mail and poster advertisement of the study details. This information was circulated to the study institutions (University College Cork and Cork University Hospital) and local businesses in Cork City, Ireland. Participants were informed that free gymnasium

membership would be supplied for the study period. Interested individuals contacted researchers via the study telephone line and were screened initially for inclusion criteria (see Table 6.1 in the supplemental material). Baseline levels of physical activity were assessed using the International Physical Activity Questionnaire short form [19]. If appropriate, a subsequent screening visit at the study site was arranged for further assessment of the exclusion criteria. Safe participation in the exercise program was ensured by medical screening of all participants using an adapted version of the safe participation questionnaire of the American College of Sports Medicine [20].

6.3.3 | Intervention group allocation

Eligible volunteers were randomized into 2 intervention groups, namely, an exercise-only group (E group) and an exercise plus daily whey protein supplementation group (EP group) (Figure 6.1). A separate parallel group consuming whey protein supplementation alone (P group) was included in the study as a control. Participants in the P group were instructed to maintain their usual levels of light physical activity. To encourage recruitment to the control group, volunteers were offered an exercise program at a later date but were not followed extensively during that period. All participants were observed and their responses measured for 8 weeks (n = 30 for each group).

6.3.4 | Combined exercise intervention

Combined aerobic and resistance training was performed at the Mardyke Arena gymnasium at University College Cork, Ireland. All exercise sessions took place at this venue. Volunteers in the P group were asked to maintain usual levels of physical inactivity for the 8-week period. Participants in the E and EP groups were instructed to adhere to the assigned exercise program and to avoid additional (moderate to vigorous) physical activity

outside that prescribed. Details of the structured exercise intervention are presented in section 2.4. Compliance and withdrawal from the study

Compliance with the prescribed exercise program was monitored remotely by the investigators using a FitLinxx activity monitoring system (Activelinxx, Shelton, CT), and is further outlined in section 2.2. This activity monitoring system was used to record the quantity of aerobic and resistance training performed by participants (supplementary table 6.4).

Participants noted to have not complied with the exercise regime for more than 7 consecutive days were withdrawn from the study. Individuals requiring antibiotics during the intervention period were also withdrawn from the study, as were participants not complying with whey protein intake requirement in the EP and P groups.

6.3.5 | Measurement visits

Measurement visits took place at 2 sites: Cork University Hospital and the Mardyke Arena, University College Cork. Baseline measurement was conducted within the 4 days prior to the commencement of the intervention period and once more after the 8-week intervention. Participants were asked to refrain from the use of alcohol and medication and moderate to vigorous physical activity for at least 24 h prior to measurement. To minimize potential effects of diurnal variation, measurement visits took place between 7:00 a.m. and 10:30 a.m. Initially, participants attended the Department of Medicine research facility at Cork University Hospital and sat restfully in a quiet environment. Participants proceeded to participate in measurement of clinical variables, e.g., recording of weight, blood pressure, and heart rate, before undergoing phlebotomy by a trained nurse using universal precautions. Approximately 16 ml of venous blood was withdrawn. Plasma and serum samples were transported immediately to the clinic laboratories at the Mercy University

Hospital, Cork. Standardized laboratory techniques were employed for the measurement of haematology and biochemistry indices. Following phlebotomy, individuals underwent a total body dual-energy X-ray absorptiometry (DEXA) scan to assess body composition (refer to section 2.3.3 for further detail). When possible, volunteers were asked to provide fresh urine and faecal samples, which were transported at room temperature to the Teagasc Food Research Centre, Moorepark, Fermoy, Co. Cork, where DNA extraction took place in accordance with the procedures outlined in section 2.5.1. Following completion of the body composition assessment, participants proceeded to the indoor track at the Mardyke Arena gymnasium to undergo a submaximal cardiorespiratory fitness assessment as described in section 2.3.2.

6.3.6 | Inflammatory cytokine measurement

Blood samples (4 ml) from participants were collected in serum separator clot activator blood collection tubes (Greiner Bio-One, Stonehouse, United Kingdom; reference no. 454071). Serum C reactive protein (CRP) was measured. Additionally, resting levels of proinflammatory cytokines interleukin-6 (IL-6), IL-8, IL-10, tumour necrosis factor alpha (TNF-α), and gamma interferon were measured using a mesoscale discovery (MSD) platform. Further description of blood sample preparation and immunoassay are presented in section 2.3.4. Samples from all 3 intervention groups were dispersed across each MSD plate.

6.3.7 | Dietary data collection

Dietary data were collected by means of a 146-item food frequency questionnaire (FFQ) as described in section 2.3.1 and outlined previously [12]. Participants were asked to record their usual pattern of dietary intake over the previous 8 weeks. Completed FFQs were

coded and dietary data were visualized with correspondence analysis using the ade4 package [21] in the R programing environment (V.3.3.2).

6.3.8 | DNA extraction and metagenomic sequencing of faecal microbiome and whey protein supplement

DNA was extracted from the donated fresh faecal samples received at the Teagasc Moorepark research facility and prepared for shotgun metagenomic sequencing as detailed in section 2.5. A total of 8 equimolar library pools of samples were made prior to shipping of the pools on dry ice for sequencing on an Illumina Hiseq 2500 (chemistry V.4.0) sequencing platform (Beckman Coulter, Inc.; Genomics Inc., Danvers, MA). High-throughput sequencing was performed using the high-output run mode for 2 × 125-bp paired-end reads with the addition of a PhiX library (1%) to estimate sequence quality. A sample of the whey protein used in the study and a sample of an oat-based nutritional supplement used as a control were both processed in a manner identical to that used with the faecal samples for the extraction of microbial DNA and preparation of metagenomic libraries. Sequencing of the supplement libraries was performed using an Illumina MiSeq (chemistry V.3.0) platform in high-output run mode for 2 × 300-bp paired-end reads (Teagasc sequencing facility).

6.3.9 | Bioinformatic processing of microbial metagenomic sequencing

Metagenomic FASTQ sequence files were processed according to section 2.6. Models of microbial metabolic pathways produced by HUMAnN2 were derived from the MetaCyc database [22] and were the basis for analyses performed on microbial metabolic profiling. Taxonomic profiling was facilitated by use of the Kraken taxonomy assignment software tool (V.0.10.6) [23].

6.3.10 | Metabolomic sample preparation

Samples of urine and faeces were stored at −80°C prior to analysis. Metabolomic processing proceeded as described in sections 2.7 and 2.8.4.

6.3.11 | Quantification and statistical analysis

Statistical analysis was carried out using the Statistical Package for the Social Sciences V.23 (SPSS, Inc., Chicago, Illinois) and the R statistical programing environment (V.3.3.2). Due to the predominance of non-normally distributed data, nonparametric analyses were performed to compare baseline clinical and demographic variables between groups. Similarly, nonparametric statistical tests were employed in the analysis of microbiome and metabolomics data. Clinical data are presented as medians and interquartile ranges (IQR), unless stated otherwise. Between-group differences in baseline, follow-up, and postintervention changes (Δ) in clinical and demographic data were compared using the Kruskal-Wallis test. For significantly different results, a Mann-Whitney *U* test was performed to determine the groups between which this difference applied. Where stated, the Wilcoxon signed-rank test was used to compare baseline and postintervention values within intervention groups. A type I error rate of ≤0.05 was considered significant in all cases. Correction of *P* values relating to microbiome and metabolomic analysis was performed using the Benjamini-Hochberg false-discovery rate (FDR) [24] in the base *stats* package in R. Statistical assessment of dissimilarity matrices (Bray-Curtis) derived from microbial data was facilitated with the adonis2 function in the vegan R package (V.2.4-3) [25]. Identification of statistically relevant taxonomic features was performed with the analysis of composition of microbiomes (ANCOM) test as implemented in the R package of the same name (V.1.1-3) [26]. Detection of underlying features of metabolic pathways was performed with unsupervised cross-validated partial-least-squares–discriminant analysis (PLS-DA)

and the KODAMA algorithm from the R package of the same name (V.1.4) [27]. Measurements of alpha diversity and calculations of relative abundances were also performed with the vegan package. Relative-abundance data were generated separately for identified species within each phylogenetic domain (e.g., Bacteria).

For metabolomic analysis, the resulting ^{1}H-NMR and LC-MS data sets were imported into MatLab to conduct multivariate statistical analysis. Data were centred and scaled to account for the repeated-measures design and then modelled using partial-least-squares–discriminant analysis (PLS-DA) with Monte Carlo cross-validation (MCCV) [28]. The fit and predictability of the models obtained were determined and expressed as R2 and Q2 values, respectively.

6.3.12 | Data and software availability

The microbial DNA sequences have been deposited in the European Nucleotide Database (ENA) database under ID code PRJEB20054.

6.3.13 | Acknowledgements

We thank the staff of the Mardyke Arena and Mercy University Hospital, Cork, Ireland, for their assistance in conducting this study, in addition to the study volunteers for their time and efforts.

This research was funded by Science Foundation Ireland (SFI) in the form of a centre grant (APC microbiome Ireland grant SFI/12/RC/2273) and by the Irish Centre for Arthritis Research and Education (I CARE). Research in the Cotter laboratory is funded by SFI through the PI award "Obesibiotics" (11/PI/1137). O.O. and W.B. are funded by SFI through a Starting Investigator research grant award (13/SIRG/2160). N.C.P. is funded by the Diabetes Research & Wellness Foundation through the Sutherland-Earl Clinical Research Fellowship 2015. We thank the Imperial-National Institute for Health Research (NIHR)

Clinical Phenome Centre for support. The centre is supported by the NIHR Imperial Biomedical Research Centre based at Imperial College London Healthcare National Health Service (NHS) Trust and Imperial College London. The views expressed are ours and not necessarily those of the NHS, the NIHR, or the Department of Health. Carbery Foods Ltd., Ballineen, Co. Cork, donated whey protein supplements for use in this study.

F.S. is a founder shareholder in Atlantia Food Clinical Trials, Tucana Health Ltd., and Alimentary Health Ltd. He is director of the APC Microbiome Ireland, a research centre that is funded in part by Science Foundation Ireland (APC/SFI/12/RC/2273) and which is/has recently been in receipt of research grants from Abbvie, Alimentary Health, Cremo, Danone, Janssen, Friesland Campina, General Mills, Kerry, MeadJohnson, Nutricia, 4D Pharma, Second Genome, and Sigmoid Pharma.

6.4 | Results

6.4.1 | Study overview

Following local advertisement, healthy Irish male and female Caucasian volunteers (n=90) aged 18 to 40 years and with a body mass index (BMI) of between 22 and 35kg/m^2 (predominantly overweight or obese) were recruited between January and August 2014 (Figure 6.1). The study was conducted in accordance with the Declaration of Helsinki, and, prior to commencement, ethical approval was granted by the Cork Clinical Research Ethics Committee (CREC). All volunteers provided written informed consent. To prospectively measure the effect of *de novo* exercise training on gut microbiota, subjects were required to be physically inactive for at least 3 months prior to study entry (i.e. not engaged in regular structured or unstructured exercise beyond light physical activities of daily life). All participants were screened for specific exclusion criteria, including regular medication use and history of cardiovascular disease (CVD), diabetes mellitus, or autoimmune disorders (see Table 6.1 in the supplementary material). Volunteers who had received oral antibiotics or bowel preparations or had suffered gastroenteritis 1 month prior to study enrolment were excluded.

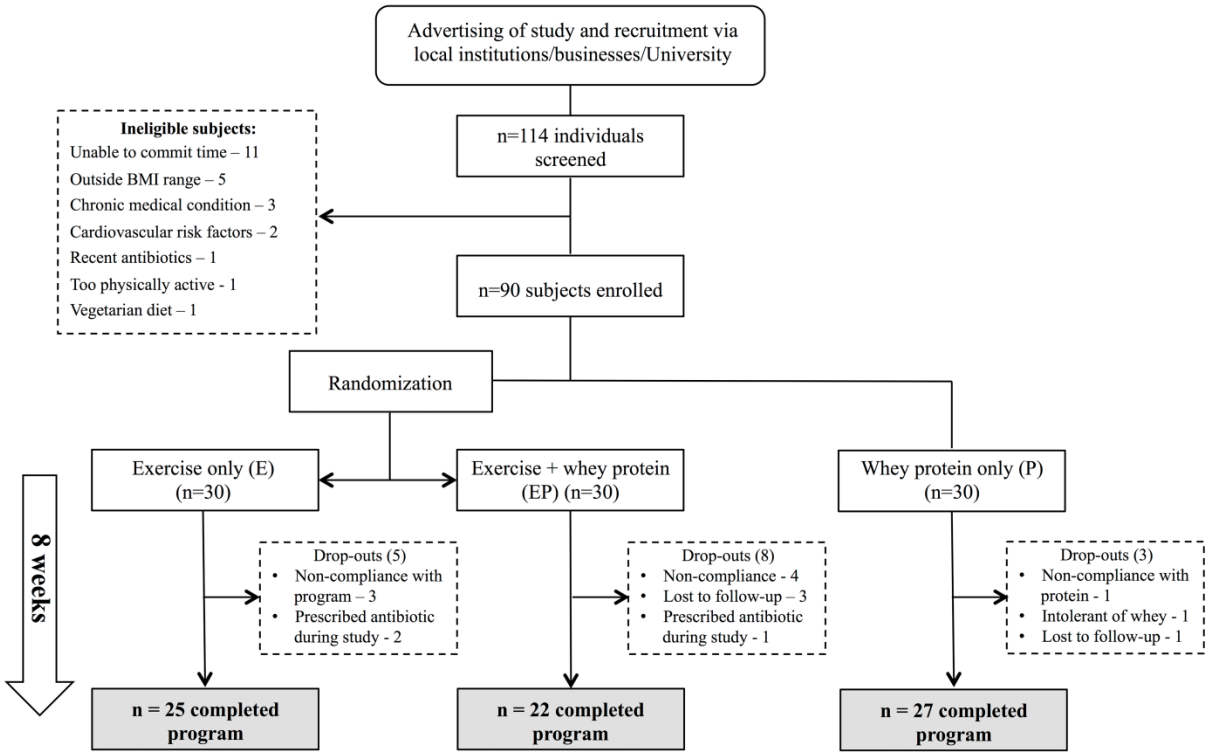


Figure 6.1 | Study design. The figure presents details of study recruitment and allocation of participants to intervention groups as follows: exercise-only group (E), exercise and protein supplementation group (EP), and whey protein supplementation-only group (P). Reasons for volunteer dropout and completion numbers are also outlined. See also supplementary table 6.1.

Eligible volunteers were randomised into 2 intervention groups: an exercise-only group (E group) and an exercise plus daily whey protein supplementation group (EP group) (Figure 6.1). A separate parallel group consuming whey protein supplementation but not participating in exercise programs (P group) was included in the study as a control. To encourage recruitment to this control group, volunteers were offered an exercise program at a later date, but their participation in the program after the conclusion of the study was not followed extensively. All participants were observed and measured for 8 weeks (n = 30 for each group). The exercise-only group (E) participated in an 8-week mixed aerobic and resistance exercise training program. The exercise plus whey protein supplementation group (EP) followed the same exercise program, in addition to consuming the once-daily whey protein supplement. All volunteers were asked to maintain usual *ad libitum* dietary intake during the intervention period and to refrain from taking additional vitamin, dietary, or herbal supplements.

Participants in the E and EP groups were required to train 3 times per week for 8 weeks. The exercise program consisted of combined aerobic and resistance training. Aerobic exercise was standardized, progressive, and similar in energy expenditure to a "couch-to-5km-running" program. The intensity of aerobic exercise was moderate, being graded between 5 and 7 of 10 on the modified Borg rating of perceived exertion (RPE) scale [29]. Resistance training consisted of 7 machine-based resistance exercises. Starting weights were calculated at induction at 70% of the individual's one-repetition maximum value (1RM). Subjects were required to perform a minimum 3 sets of 8 repetitions. Resistance training was progressive with aims of increasing resistance weight by 15% to 20% over the 8-week period.

To ensure a uniform and consistent increase in daily protein consumption, subjects in the P and EP arms of the study were required to take a daily 30-g protein supplement containing 24 g of whey protein (donated as an unrestricted grant by Carbery Group, Ballineen, Co. Cork, Ireland). The supplement comprised a blend of whey protein concentrate, isolate, and hydrolysed whey protein concentrate (see supplementary table 6.2 for full nutritional details). Subjects' compliance to daily whey protein supplementation was encouraged using daily text message reminders. Volunteers were required to return empty whey protein sachets to the study site fortnightly before the issuing of further supplement. Subjects with a compliance rate of less than 90% were excluded from the study. Baseline measurements were not significantly different between the three study groups (Table 6.1).

Patient characteristic	Values			*P* value
	Exercise (E) only (n = 25)	Exercise + protein (EP) (n = 22)	Protein only (P) (n = 27)	
Age (yrs)	35 (28, 38)	32 (28, 35)	34 (28, 36)	0.528
No. (%) of females[b]	n = 14 (56)	n = 12 (55)	n = 11 (41)	0.48
Height (cm)	172 (165, 181)	169 (166, 183)	172 (163, 178)	0.67
Weight (kg)	78.8 (70.1, 94.5)	82.3 (69, 98.9)	76.4 (69.8, 87)	0.67
BMI (kg/m^2)	27.9 (25.1, 29.2)	27.5 (25.7, 30)	27 (24.9, 28.7)	0.761
Resting heart rate (bpm)	72 (65, 81)	68 (61, 79)	74 (66, 78)	0.36
Systolic BP (mm Hg)	128 (117, 134)	125 (121, 136)	125 (118, 130)	0.706
Diastolic BP (mm Hg)	78 (74, 89)	76 (72, 84)	79 (75, 84)	0.543
Waist/hip ratio	0.85 (0.83, 0.89)	0.84 (0.8, 0.93)	0.83 (0.78, 0.88)	0.365
Body fat (%)	32.8 (29, 38.7)	34.7 (29, 37.2)	34.5 (29.3, 39.4)	0.659
Fat mass (kg)	26.3 (22.6, 30.6)	26 (23, 33.1)	26.8 (20.7, 32.9)	0.96
Fat mass (trunk) (kg)	14.1 (10.8, 16.8)	14.1 (11.2, 17.6)	13.7 (9.4, 17.1)	0.878
Lean tissue mass (kg)	52.4 (40.7, 61.4)	51.3 (41.5, 61.5)	47.2 (42.9, 53.3)	0.44
Weekly PA (METS)	462 (298, 1,139)	564 (413, 844)	657 (424, 1,145)	0.599
Weekly PA (kCal)	761 (381, 1,618)	748 (525, 1,127)	762 (512, 1,773)	0.767
Sitting time (h per wk)	56 (40, 61)	62 (47, 76)	51 (33, 62)	0.114
Motorized transport (h per wk)	5 (3.25, 8.3)	3.5 (2, 6)	4.1 (0.8, 7)	0.27

[a]Values represent medians (interquartile ranges) except where otherwise indicated. *P* values represent results of Kruskal-Wallis tests or chi-square tests. BMI, body mass index; IPAQ, International Physical Activity Questionnaire; METS, metabolic equivalents.

[b]Data indicate chi-square test results.

Table 6.1 | Baseline demographic and anthropometric characteristics of the study participants with comparisons between the 3 intervention groups.

6.4.2 | 8 weeks of aerobic and resistance training improves body composition and cardio-respiratory fitness profiles in sedentary subjects

A total of 74 of the 90 participants enrolled in the study completed the 8-week study period (reasons for dropping out are detailed in Figure 6.1). At entry, the intervention groups shared similar clinical and anthropometric characteristics. Following the intervention period, both E and EP group participants demonstrated significant and similar improvements in predicted maximal aerobic capacity (VO_{2max}) (Figure 6.2A). Furthermore, resting heart rate was significantly reduced following the intervention period in both of the exercising groups (E and EP) compared with the protein-only group ($P = 0.005$) (supplementary table 6.3). Compliance with the prescribed exercise program was high, with a median of 21 sessions (87.5%) performed in both the E and EP groups. The types and levels of exercise training undertaken in both groups were similar, with no statistically significant differences in the aerobic- and resistance-training workloads recorded (supplementary table 6.4).

Participants were predominantly overweight, with body fat percentages above 30%. There were no significant differences in the participants' baseline levels of physical activity when assessed using the International Physical Activity Questionnaire [30]. All baseline values are expressed as medians and interquartile ranges (IQR). In contrast to the protein-only group, the exercise-only group and the exercise plus protein supplementation group experienced significant decreases in percentage body fat, total fat mass, and trunk fat mass during the intervention period (Figure 6.2B and C), in addition to an increase in total lean tissue mass (Figure 6.2D) (all $P < 0.001$; see also supplementary table 6.3).

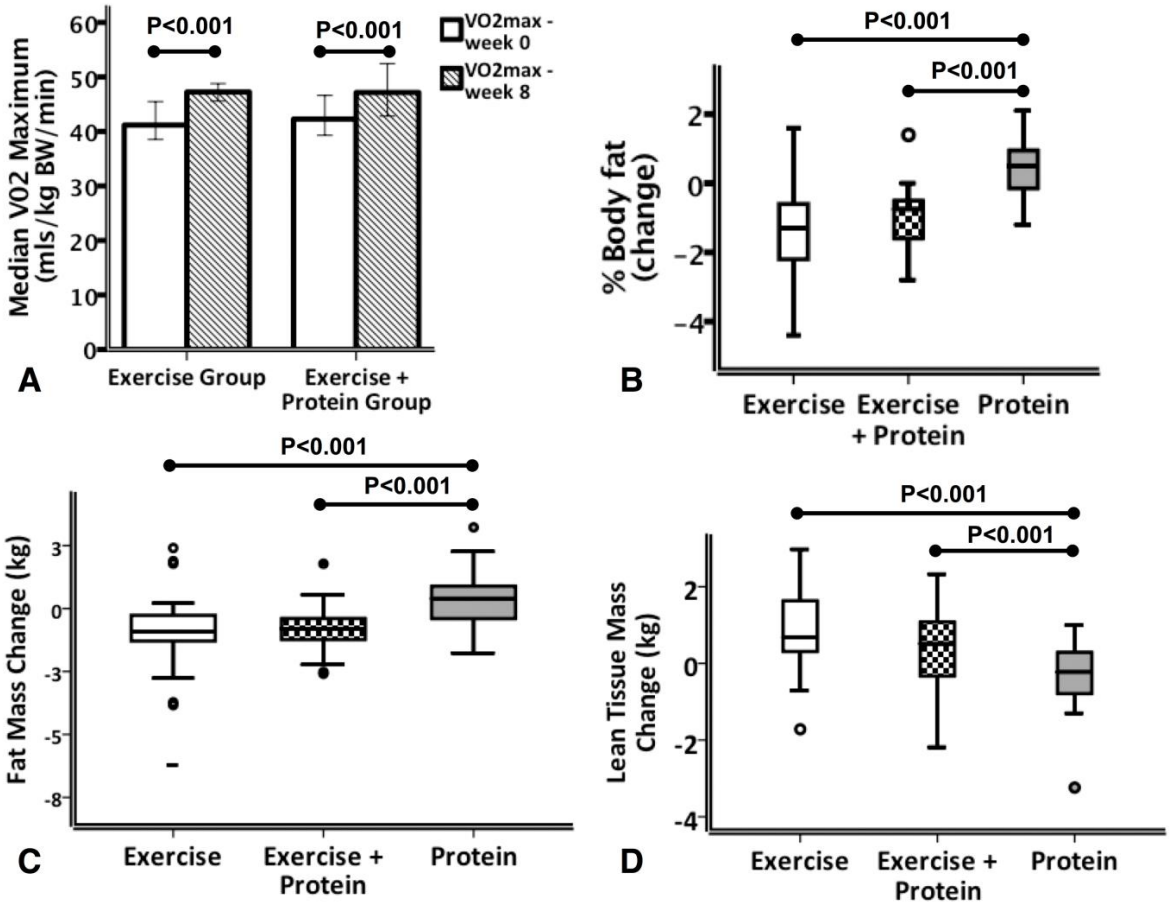


Figure 6.2 | Alterations in cardiorespiratory fitness and body composition following exercise interventions, protein interventions, and combined interventions. (**A**) Peak aerobic capacity (VO_{2max}) per kilogram of body weight as predicted using the Rockport 1-mile walk test was higher in both the E and EP groups following the intervention period, indicating improved levels of cardiorespiratory fitness. Within-group comparisons were tested using the Wilcoxon signed-rank test ($P < 0.001$). (**B**) Changes in percentages of body fat following the intervention period as measured using DEXA. Percent body fat reduction was significantly greater in the exercise-only group and in the exercise plus protein supplementation group compared to the protein-only group. (**C**) Absolute changes in body fat mass (in kilograms) following the intervention period demonstrated a significantly greater reduction in both the exercise and exercise plus protein supplementation groups. (**D**) Absolute change in lean tissue mass (kg), measured using a three-compartment model, indicating significantly greater lean mass accretion in the E and EP groups than in the P group. Error bars represent 95% confidence intervals. See also supplementary tables 6.3 to 6.5

Compliance with daily whey protein supplementation in the EP and P groups was high with only one participant excluded due to poor adherence to whey protein supplementation. Whey protein supplementation aside, dietary frequency patterns did not deviate from the volunteers' usual intake at study entry (supplementary figure 6.1). Addition of the 30 g daily whey protein supplement did not favour the EP over E group with respect to body composition improvement; however, the study was not designed to test this hypothesis. No clinically relevant differences in resting-state serum proinflammatory markers were evident following any of the interventions (supplementary figure 6.2; see also supplementary table 6.5).

6.4.3 | Metagenomic assessment of microbiota after exercise and/or dietary adjustment

Postintervention alterations (percent Δ) in gut microbial α-diversity did not identify significant modulation in taxonomic composition or metabolic pathways for any of the intervention groups compared to baseline (Figure 6.3A to D). A trend of median increase in bacterial diversity was observed for the E and EP groups (Figure 6.3B). These findings of moderate alterations of α-diversity were consistent across pairwise comparisons of the groups, with a few notable exceptions. Increased α-diversity of Archaea species in the P group following intervention was observed, as was a moderate enhancement of archaeal diversity in the P group compared to the EP group ($P < 0.05$ and $P < 0.01$, respectively; Figure 6.3E). After the intervention period, bacterial diversity was greater in the EP group than in the P group ($P < 0.05$; Figure 6.3F), while the diversity of virus species was lower in EP group than in the E group ($P < 0.05$; Figure 6.3G).

Principal-coordinate analysis (PCoA) was used to present separation of measures from the taxonomic composition and metabolic pathway models (Figure 6.3H to O). Prior to

intervention, all 3 groups demonstrated similarity in measures of microbial metabolic pathways and taxonomic β-diversity (Figure 6.3H to K). A significant separation between the intervention groups was detected in the Bray-Curtis-derived dissimilarity matrices generated from participants postintervention for metabolic pathways (P = 0.054; Figure 6.3L), the entirety of detected species (P < 0.001; Figure 6.3M), and species of bacteria (P < 0.05) and viruses (P < 0.001) (Figure 6.3N and O, respectively). Archaea species did not differentiate with intervention (data not presented).

Pairwise analysis of taxonomy compared according to high-level phylogeny (Archaea, Bacteria, and viruses) demonstrated significant alterations of detected virus species in both the EP and P groups following the intervention period that were absent from the exercise-only group (P < 0.001; Figure 6.4). There were no further significant separations for Archaea or Bacteria species or for metabolic pathways (supplementary figure 6.3A to I). An unsupervised partial-least-squares–discriminant analysis (PLS-DA) approach was used to identify underlying features of the metabolic pathways before and after the intervention period (supplementary figure 6.3J and K). Pathways associated with *Prevotella copri* were shown to cluster with the E group prior to intervention. Following intervention, this cluster was still present but, in addition, separate clusters of *P. copri*- and *Bacteroides vulgatus*-associated pathways were apparent within the EP and P groups, respectively. Forty-eight species were detected as being differentially abundant within the three groups (false-discovery rate [FDR] = 0.05). The majority of identified taxa were virus species, predominantly *Lactococcus* phage, within the P and EP groups. No Archaea were found to have significantly varied in abundance with treatment in any of the groups (supplementary table 6.6).

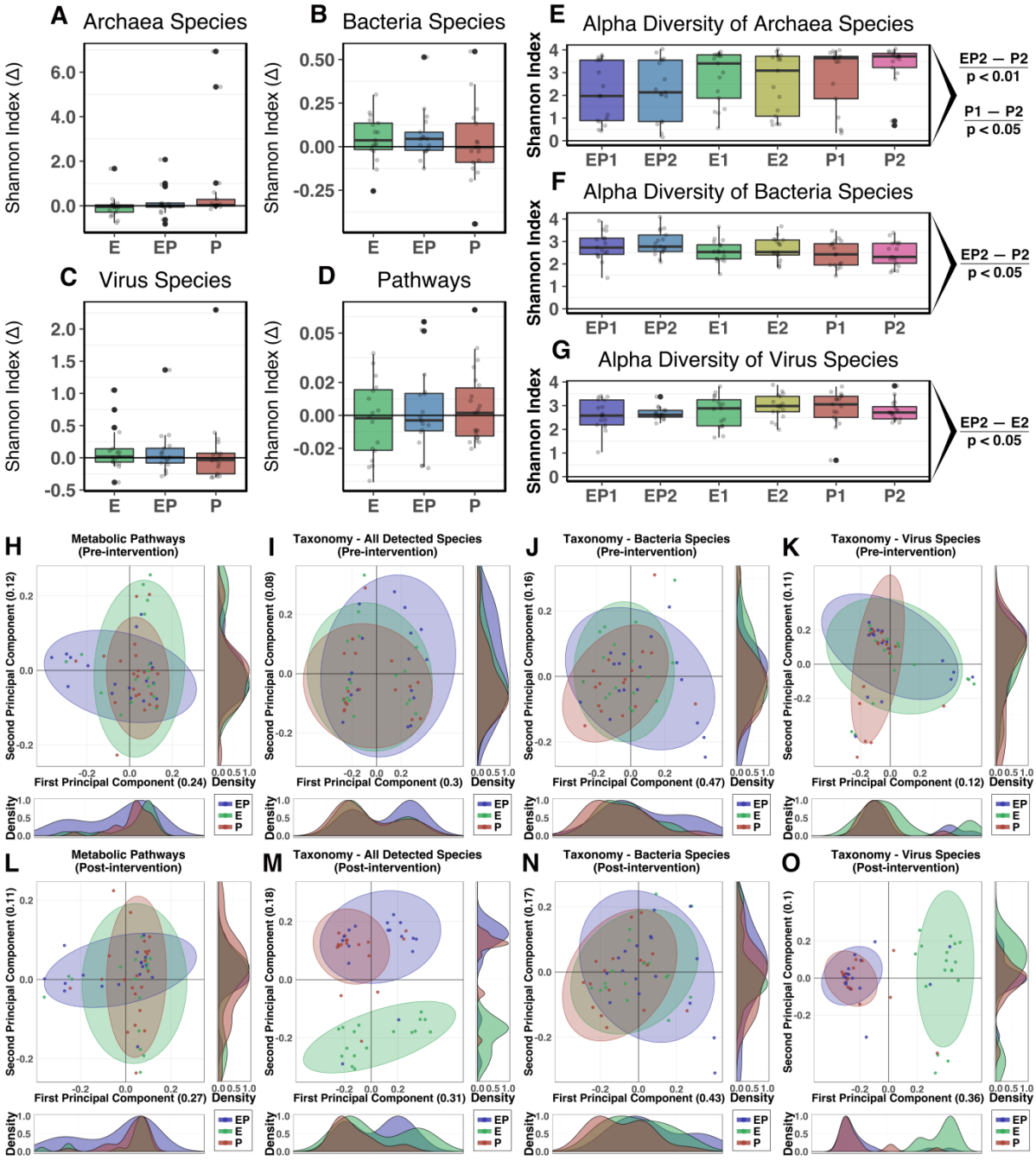


Figure 6.3 | Intervention effects on taxonomic and functional pathway diversity of the intestinal microbiome.

Figure 6.3 Continued

(**A** to **D**) Percent change (Δ) of Shannon α-diversity H-index values following intervention. No significant variations were presented for taxonomic measurements (**A** to **C**) or metabolic pathways (**D**). (**E** to **G**) Pairwise statistical assessment of taxonomy α-diversity demonstrates equal data with respect to the presence of taxonomy between groups at baseline. EP1, combined exercise and protein supplementation group, week 0; EP2, combined exercise and protein supplementation group, week 8; E1, exercise-only group, week 0; E2, exercise-only group, week 8; P1, protein-only group, week 0; P2, protein-only group, week 8. (**E**) The diversity of Archaea was significantly altered after intervention within the P group ($P < 0.05$) and, similarly, was greater in the P group (P2) than in the EP group (EP2) ($P < 0.01$). (**F**) Postintervention bacterial diversity was greater in the EP group (EP2) in testing against the P group (P2) ($P < 0.05$). (**G**) Similar levels of virus diversity were presented in the protein supplementation groups (EP and P) following the intervention, with significantly lower diversity in the EP group than in the E group ($P < 0.05$). (**H** to **O**) Principal-coordinate analysis (PCoA) of relative abundance profiles for taxonomic and metabolic pathway constructions of the three groups demonstrates the influence of the interventions on the diversity of microbial populations. (**H** to **K**) Prior to intervention, group profiles of taxonomic and metabolic pathway diversity were not significantly differentiated. (**L** to **O**) Following intervention, a significant separation was identified between the groups for measures of (**L**) metabolic pathways ($P = 0.054$), (**M**) all detected species unsegregated by phylogeny ($P < 0.001$), (**N**) bacteria ($P < 0.05$), and (**O**) virus species ($P < 0.001$). Specific separations in diversity per intervention group are outlined further in Figure 6.4 for virus species and supplementary figure 6.3 for all other comparisons. Statistical assessment of PCoA dissimilarity matrices was performed with the Adonis2 permutational multivariate analysis of variance (PERMANOVA) test. (**H** to **O**) Density plots were derived from kernel density estimates and scaled to a maximum estimated value of 1 and display concentrations of plotted data along the corresponding plot axis. *P* values were calculated for α-diversity comparisons using the Wilcoxon signed-rank test.

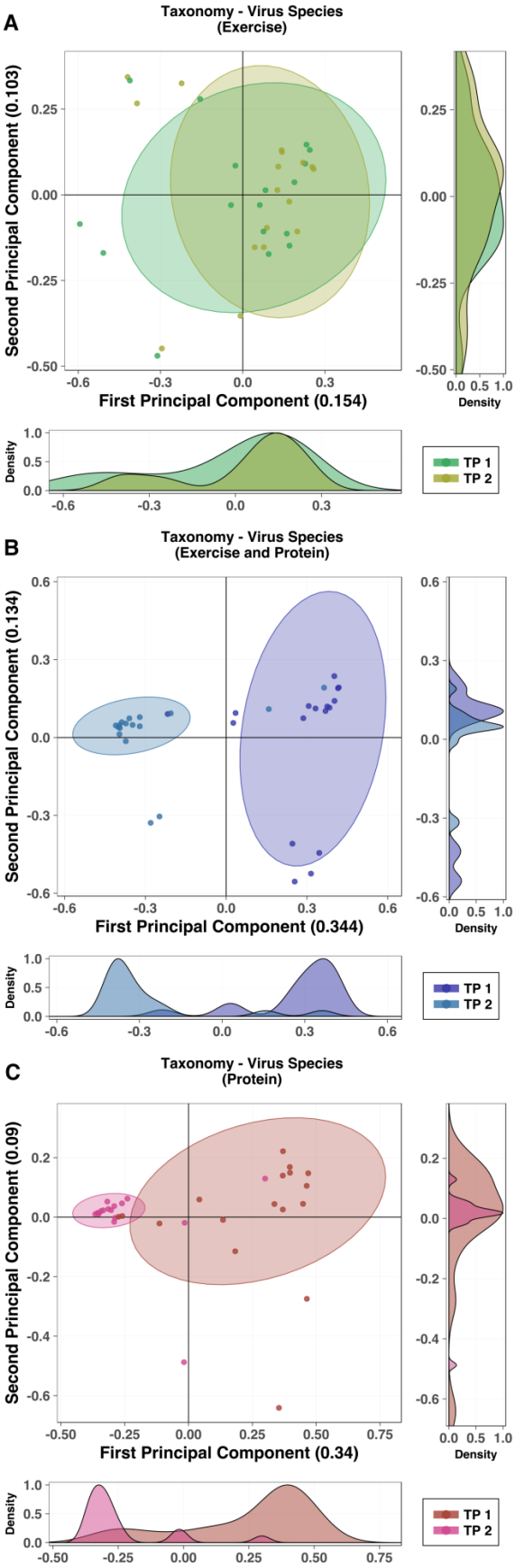


Figure 6.4 | Pairwise analysis of detected virus taxonomy prior to and following intervention.

Figure 6.4 Continued

(**A** to **C**) PCoA of virus species for each group, comparing virus profiles before and after the intervention period (time point 1 [TP1] [week 0] and time point 2 [TP2] [week 8], respectively). (**A**) The exercise-only group had virus diversity that was not significantly altered by intervention. (**B** and **C**) Diversity of viruses was significantly affected during the intervention period for both groups receiving protein supplementation ($P < 0.001$). The exercise plus protein supplementation group (**B**) and the protein-only group (**C**) demonstrated reduced variability of diversity following intervention. Results of pairwise analysis of additional taxonomic and metabolic pathway profiles are presented in supplementary figure 6.3. Statistical assessment of PCoA dissimilarity matrices was performed with the Adonis2 PERMANOVA test. (**A** to **C**) Density plots were derived from kernel density estimates and scaled to a maximum estimated value of 1 and display concentrations of plotted data along the corresponding plot axis.

Of the total 23,019 general (e.g., coenzyme A biosynthesis) and taxonomically specific (e.g., coenzyme A biosynthesis in *Akkermansia muciniphila*) metabolic pathways included in the metagenomic construction of the models, 619 were identified as having significantly differed among the three intervention groups at either the pre-treatment or posttreatment time point ($P < 0.05$). Significantly altered pathways were organized according to a metabolic pathway hierarchy defined by the MetaCyc database and were structured as a heat map of low-level categories of classification (e.g., nucleotide biosynthesis) (supplementary figure 6.4). Scaled group means of pathway relative abundances demonstrated modest alterations of microbial metabolic potential. A complete list of the categorized pathways can be found in the supplemental material (supplementary table 6.6). Further assessment of pathways differing among all groups was performed both within each group (before and after treatment) and between the separate groups. No significant variation within groups was evident following *P* value correction for multiple testing.

Untargeted metabolomic analysis of participant faecal-water and urinary samples revealed no significant separations either within each group pre- and postintervention or between groups at each time point with analysis of the full spectrum of metabolites. Subsequent targeted metabolomic quantification, guided by previous findings [31-33], revealed significant changes following intervention in the amount of glutamate (faecal water) and *trans*-aconitate (urine) in the protein-only group ($P < 0.01$ and $P < 0.05$, respectively; supplementary table 6.6, section 6.7 and Appendix B for metabolite quantifications). Comparisons of differences (percent Δ) in metabolite quantifications between all groups demonstrated significant variation in the levels of phenylacetylglycine (PAG) and trimethylamine N-oxide (TMAO) ($P < 0.01$ and $P < 0.05$, respectively) in urine, as well as of glutamate ($P < 0.05$) in faecal water, within all groups (supplementary table 6.6).

Such differences were also present in the paired comparisons. Levels of both PAG and TMAO were significantly reduced (Δ = −0.196 and −0.518, respectively) in the E group following the intervention period in comparison to the P group ($P < 0.05$, supplementary table 6.6).

6.4.4 | Characterization of whey protein supplement microbial content

Metagenomic sequencing of the whey protein supplement and of a non-dairy-based dietary supplement control revealed a taxonomic profile in the former that was characterized by high proportions of bacteriophage associated with lactic acid bacteria. Notably, these phage were also enriched in participants in receipt of the whey supplement (supplementary table 6.6). The taxonomic composition of the whey protein and control demonstrated highly divergent microbial contents of the supplements, including taxa detected in participants.

6.5 | Discussion

To accurately and consistently increase daily protein intake, we selected a whey protein supplement. Whey protein, a widely used commercial supplement in elite sport and amateur fitness milieu, is known for its muscle accretion effects [34], in addition to its positive influence on energy metabolism [35-37] and, more recently, on appetite control [38]. In addition, its use facilitated analysis of the effect of a widely available exercise adjunct on the diversity, composition, and activity of microbial populations in the gut. Somewhat unexpectedly, individuals in the whey protein supplementation-only group (P) experienced a significant alteration in the β-diversity of the gut virome (Figure 6.4C). Furthermore, this change was mirrored in the combined exercise and protein supplementation (EP) group (Figure 6.4B), suggesting a robust effect of whey protein on the taxonomic richness of the gut virome. To explore this dynamic, a sample of the whey supplement and a sample of a non-dairy-based dietary supplement were sequenced for microbial content. Intriguingly, all

bacteriophage and two of the four bacterial species that were significantly altered in the groups receiving whey protein were present in high relative abundance within the whey protein supplement but not the control supplement. Further in-depth experimentation is required to determine whether virus particles from whey protein conclusively transmit to the human gut from consumption and, if so, whether they remain biologically active. However, the overlap in the taxonomic compositions of the whey supplement and participants' gut microbiome provides a convincing explanation for the source of virome changes observed.

While this examination did not identify a significant impact of short-term combined aerobic and resistance exercise on the diversity of bacterial or archaeal constituents of the gut microbiome, subtle compositional and functional changes were detected in this analysis (Figure 6.3 and 6.4; see also supplementary figures 6.3 and 6.4). Although the results were not statistically significant, the groups engaged in exercise demonstrated less change in archaeal diversity than the protein-only group after the intervention period. In the case of the exercise-only group, a reduction in Archaea diversity was observed, suggesting that exercise acts against intestinal Archaea. More-extensive investigation is necessary to resolve this issue, but in view of a putative role for Archaea in intestinal disorders and as modulators of TMAO concentrations, such an inquiry is justified [39]. Changes in Bacteria diversity were similarly below the threshold of statistical significance; however, the median differences between groups indicated that those undertaking exercise had increases in bacterial diversity that were absent from the intervention group excluded from exercise. Curiously, the diversity of bacterial species was elevated in the EP group after intervention but the diversity of virus species was uniformly lower. The inverse relationship of these measures is counterintuitive, given the predominance of bacteriophage in the detected

viruses. However, the influx of such bacteriophage may explain the overall reduction of virus diversity within the group. Furthermore, this increase in the levels of bacteriophage may have been insufficient to profoundly influence the overall diversity of the Bacteria due to their selective targeting of only a few bacterial species.

The absence of substantial modulation of the diversity of microbial populations in the gut following the 8-week exercise intervention mirrors recent findings in mice [40]. To date, most of the work in humans has focused on elite or professional athletes [10-12] and as a result has explored the relationship between established physical 'fitness' and the gut microbiota. Few prospective studies have examined the effect of exercise on the gut microbiota of physically inactive human volunteers [41]. The current study is the largest to have done so. It should be acknowledged that the unperturbed adult intestinal microbiome is resilient [42] and may not be subject to significant alteration following an 8-week intervention period. It is likely that the diverse, metabolically favourable intestinal microbiome evident in the elite athlete is the cumulative manifestation of many years of optimized nutrition and of high degrees of physical condition throughout youth and adolescence and during adult participation in professional sports [43]. Initial examination of the acute effects of extreme and prolonged endurance exercise, such as in trained military regiments, suggests that prolonged physical stress negatively impacts intestinal permeability and gut microbiota composition [44]. However, the results of the present study indicate that exercise at moderate intensity does not exert a deleterious effect on gut microbial composition or function in the untrained subject. Furthermore, the results of this study signify that exercise-induced improvements in cardiorespiratory fitness and body composition are not dependent on substantial alteration of the diversity of microbial populations in the gut. Whether the limited changes in microbiome composition and

function detected in this study contributed to the witnessed improvements in body composition and fitness profiles remains unknown.

An intriguing exception to the otherwise minimal differences in metabolomic modification is represented by the controversial metabolite TMAO. Associations between TMAO and cardiovascular disease (CVD) have framed the metabolite as a disease factor; however, high levels have also been observed in populations with low CVD risk [45-48]. Elevated levels of TMAO have previously been found in elite athletes [11], and while the presence of TMAO may or may not have deleterious health implications, we demonstrate a potential modulatory effect of exercise on urinary TMAO levels. Participants in the exercise-only group showed levels of urinary TMAO that were reduced below baseline with intervention, while the groups receiving whey protein had increased levels of the metabolite, with the combined-treatment group demonstrating lower levels than the protein-only group. The TMAO precursor phosphatidylcholine comprised less than 0.1% of the constituents of the whey protein supplement used in this study, suggesting a possible direct effect of whey protein and/or exercise on TMAO production. While the data represent a promising paradigm, further work will be necessary to determine the specific mechanisms involved and to rule out unintended dietary influence or influences of host biology (e.g., altered absorption of TMAO with exercise). Additionally, known microbial producers of TMAO [49] were absent from the taxonomic profiling. PAG concentrations were similarly reduced in the exercise-only group, although the metabolite has previously been associated with lean body composition and has been found to be present in increased concentrations in athletes. It has also recently been shown to decrease in urine in thoroughbred racehorses following exercise [50].

Likewise, the data reflecting increased abundance of *P. copri*-associated pathways detected in the EP group postintervention supports the work of others which suggested an active role for *Prevotella* species in host metabolic [51, 52] and immune health [53]. Studies have linked *Prevotella* with inflammatory and metabolic disorders, including rheumatoid arthritis [54, 55], ankylosing spondylitis [56], and type 2 diabetes mellitus [57]. Conversely, and consistent with our findings, increased physical activity has been associated with increases in *Prevotella*-related metabolic pathways in the gut microbiome [10].

It is pertinent to acknowledge the difficulty in controlling all potential confounders of gut microbial composition and activity in this investigation (e.g., diet, wide BMI range). This study attempted to control for the potential impact of dietary variation by instruction of volunteers to maintain their usual *ad libitum* dietary intake. Food frequency questionnaire (FFQ) dietary analysis indicated stability in the volunteers' dietary patterns; however, FFQ assessment is subject to its limitations, including recall bias [58].

In conclusion, this prospective examination demonstrated that short-to-medium-term combined exercise in healthy, physically inactive adults does not induce drastic alterations in the diversity of microbial populations in the gut. We highlight an interaction between whey protein intake and the β-diversity of the adult gut virome which requires further exploration. Furthermore, the functional activity of the gut microbiota does not appear to be extensively manipulated by short-term, moderate-intensity exercise and/or whey protein supplementation, although some changes, including alteration of levels of urinary TMAO and PAG excretion, were evident. The alterations in the diversity, composition, and metabolomic profiles of microbial gut populations that we and others have observed in habitual exercisers and professional athletes may represent late responses to exercise or fitness.

6.6 | References

[1] J. Debelius, S. J. Song, Y. Vazquez-Baeza, Z. Z. Xu, A. Gonzalez, R. Knight, Tiny microbes, enormous impacts: what matters in gut microbiome studies?, Genome Biol 17 (1) (2016) 217.

[2] J. L. Sonnenburg, F. Backhed, Diet-microbiota interactions as moderators of human metabolism, Nature 535 (7610) (2016) 56-64.

[3] F. Shanahan, D. van Sinderen, P. W. O'Toole, C. Stanton, Feeding the microbiota: transducer of nutrient signals for the host, Gut (2017)

[4] S. Bermon, B. Petriz, A. Kajeniene, J. Prestes, L. Castell, O. L. Franco, The microbiota: an exercise immunology perspective, Exerc Immunol Rev 21 (2015) 70-9.

[5] B. Cerda, M. Perez, J. D. Perez-Santiago, J. F. Tornero-Aguilera, R. Gonzalez-Soltero, M. Larrosa, Gut Microbiota Modification: Another Piece in the Puzzle of the Benefits of Physical Exercise in Health?, Front Physiol 7 (2016) 51.

[6] A. Mika, W. Van Treuren, A. Gonzalez, J. J. Herrera, R. Knight, M. Fleshner, Exercise is More Effective at Altering Gut Microbial Composition and Producing Stable Changes in Lean Mass in Juvenile versus Adult Male F344 Rats, PLoS One 10 (5) (2015) e0125889.

[7] O. Cronin, M. G. Molloy, F. Shanahan, Exercise, fitness, and the gut, Curr Opin Gastroenterol 32 (2) (2016) 67-73.

[8] M. Estaki, J. Pither, P. Baumeister, J. P. Little, S. K. Gill, S. Ghosh, Z. Ahmadi-Vand, K. R. Marsden, D. L. Gibson, Cardiorespiratory fitness as a predictor of intestinal microbial diversity and distinct metagenomic functions, Microbiome 4 (1) (2016) 42.

[9] C. Bressa, M. Bailen-Andrino, J. Perez-Santiago, R. Gonzalez-Soltero, M. Perez, M. G. Montalvo-Lominchar, J. L. Mate-Munoz, R. Dominguez, D. Moreno, M. Larrosa, Differences in gut microbiota profile between women with active lifestyle and sedentary women, PLoS One 12 (2) (2017) e0171352.

[10] L. M. Petersen, E. J. Bautista, H. Nguyen, B. M. Hanson, L. Chen, S. H. Lek, E. Sodergren, G. M. Weinstock, Community characteristics of the gut microbiomes of competitive cyclists, Microbiome 5 (1) (2017) 98.

[11] W. Barton, N. C. Penney, O. Cronin, I. Garcia-Perez, M. G. Molloy, E. Holmes, F. Shanahan, P. D. Cotter, O. O'Sullivan, The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level, Gut. Published online 30th March, 2017.

[12] S. F. Clarke, E. F. Murphy, O. O'Sullivan, A. J. Lucey, M. Humphreys, A. Hogan, P. Hayes, M. O'Reilly, I. B. Jeffery, R. Wood-Martin, D. M. Kerins, E. Quigley, R. P. Ross, P. W. O'Toole, M. G. Molloy, E. Falvey, F. Shanahan, P. D. Cotter, Exercise and associated dietary extremes impact on gut microbial diversity, Gut 63 (12) (2014) 1913-20.

[13] C. C. Evans, K. J. LePard, J. W. Kwak, M. C. Stancukas, S. Laskowski, J. Dougherty, L. Moulton, A. Glawe, Y. Wang, V. Leone, D. A. Antonopoulos, D. Smith, E. B. Chang, M. J.

Ciancio, Exercise prevents weight gain and alters the gut microbiota in a mouse model of high fat diet-induced obesity, PLoS One 9 (3) (2014) e92193.

[14] M. I. Queipo-Ortuno, L. M. Seoane, M. Murri, M. Pardo, J. M. Gomez-Zumaquero, F. Cardona, F. Casanueva, F. J. Tinahones, Gut microbiota composition in male rat models under different nutritional status and physical activity and its association with serum leptin and ghrelin levels, PLoS One 8 (5) (2013) e65465.

[15] O. O'Sullivan, O. Cronin, S. F. Clarke, E. F. Murphy, M. G. Molloy, F. Shanahan, P. D. Cotter, Exercise and the microbiota, Gut Microbes 6 (2) (2015) 131-6.

[16] D. Vandeputte, G. Falony, S. Vieira-Silva, R. Y. Tito, M. Joossens, J. Raes, Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates, Gut 65 (1) (2016) 57-62.

[17] B. K. Song, K. O. Cho, Y. Jo, J. W. Oh, Y. S. Kim, Colon transit time according to physical activity level in adults, J Neurogastroenterol Motil 18 (1) (2012) 64-9.

[18] A. Mika, M. Fleshner, Early-life exercise may promote lasting brain and metabolic health through gut bacterial metabolites, Immunol Cell Biol 94 (2) (2016) 151-7.

[19] C. L. Craig, A. L. Marshall, M. Sjostrom, A. E. Bauman, M. L. Booth, B. E. Ainsworth, M. Pratt, U. Ekelund, A. Yngve, J. F. Sallis, P. Oja, International physical activity questionnaire: 12-country reliability and validity, Med Sci Sports Exerc 35 (8) (2003) 1381-95.

[20] L. S. Pescatello, ACSM's guidelines for exercise testing and prescripton, Wolters Kluwer/Lippincott Williams and Wilkins (Health)

[21] S. Dray, A.-B. Dufour, The ade4 package: implementing the duality diagram for ecologists, Journal of statistical software 22 (4) (2007) 1-20.

[22] R. Caspi, H. Foerster, C. A. Fulcher, P. Kaipa, M. Krummenacker, M. Latendresse, S. Paley, S. Y. Rhee, A. G. Shearer, C. Tissier, The MetaCyc Database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases, Nucleic acids research 36 (suppl 1) (2008) D623-D631.

[23] D. E. Wood, S. L. Salzberg, Kraken: ultrafast metagenomic sequence classification using exact alignments, Genome Biol 15 (3) (2014) R46.

[24] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, Journal of the royal statistical society. Series B (Methodological) (1995) 289-300.

[25] J. Oksanen, F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. O'Hara, G. L. Simpson, P. Solymos, M. Henry, H. Stevens, E. Szoecs, H. Wagner, Vegan: Community Ecology Package. R package version 2.4-3,

[26] S. Mandal, W. Van Treuren, R. A. White, M. Eggesbo, R. Knight, S. D. Peddada, Analysis of composition of microbiomes: a novel method for studying microbial composition, Microb Ecol Health Dis 26 (2015) 27663.

[27] S. Cacciatore, L. Tenori, C. Luchinat, P. R. Bennett, D. A. MacIntyre, KODAMA: an R package for knowledge discovery and data mining, Bioinformatics 33 (4) (2017) 621-623.

[28] I. Garcia-Perez, J. M. Posma, R. Gibson, E. S. Chambers, T. H. Hansen, H. Vestergaard, T. Hansen, M. Beckmann, O. Pedersen, P. Elliott, J. Stamler, J. K. Nicholson, J. Draper, J. C. Mathers, E. Holmes, G. Frost, Objective assessment of dietary patterns by use of metabolic phenotyping: a randomised, controlled, crossover trial, Lancet Diabetes Endocrinol 5 (3) (2017) 184-195.

[29] G. Borg, Borg's perceived exertion and pain scales, Human kinetics

[30] Y. Kim, I. Park, M. Kang, Convergent validity of the international physical activity questionnaire (IPAQ): meta-analysis, Public Health Nutr 16 (3) (2013) 440-52.

[31] W. Barton, N. C. Penney, O. Cronin, I. Garcia-Perez, M. G. Molloy, E. Holmes, F. Shanahan, P. D. Cotter, O. O'Sullivan, The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level, Gut (2017)

[32] H. H. Lillefosse, M. R. Clausen, C. C. Yde, D. B. Ditlev, X. Zhang, Z. Y. Du, H. C. Bertram, L. Madsen, K. Kristiansen, B. Liaset, Urinary loss of tricarboxylic acid cycle intermediates as revealed by metabolomics studies: an underlying mechanism to reduce lipid accretion by whey protein ingestion?, J Proteome Res 13 (5) (2014) 2560-70.

[33] B. D. Piccolo, K. B. Comerford, S. E. Karakas, T. A. Knotts, O. Fiehn, S. H. Adams, Whey protein supplementation does not alter plasma branched-chained amino acid profiles but results in unique metabolomics patterns in obese women enrolled in an 8-week weight loss trial, J Nutr 145 (4) (2015) 691-700.

[34] R. W. Morton, C. McGlory, S. M. Phillips, Nutritional interventions to augment resistance training-induced skeletal muscle hypertrophy, Front Physiol 6 (2015) 245.

[35] L. McAllan, D. Keane, H. Schellekens, H. M. Roche, R. Korpela, J. F. Cryan, K. N. Nilaweera, Whey protein isolate counteracts the effects of a high-fat diet on energy intake and hypothalamic and adipose tissue expression of energy balance-related genes, Br J Nutr 110 (11) (2013) 2114-26.

[36] G. T. Sousa, F. S. Lira, J. C. Rosa, E. P. de Oliveira, L. M. Oyama, R. V. Santos, G. D. Pimentel, Dietary whey protein lessens several risk factors for metabolic diseases: a review, Lipids Health Dis 11 (2012) 67.

[37] R. E. Brimelow, N. P. West, L. T. Williams, A. W. Cripps, A. J. Cox, A role for whey-derived lactoferrin and immunoglobulins in the attenuation of obesity-related inflammation and disease, Crit Rev Food Sci Nutr 57 (8) (2017) 1593-1602.

[38] R. A. Reimer, H. J. Willis, J. M. Tunnicliffe, H. Park, K. L. Madsen, A. Soto-Vaca, Inulin-type fructans and whey protein both modulate appetite but only fructans alter gut microbiota in adults with overweight/obesity: A randomized controlled trial, Mol Nutr Food Res (2017)

[39] V. Demonfort Nkamga, B. Henrissat, M. Drancourt, Archaea: Essential inhabitants of the human digestive microbiota., Human Microbiome Journal. 3 (2017) 1-8.

[40] E. V. Lamoureux, S. A. Grandy, M. G. I. Langille, Moderate Exercise Has Limited but Distinguishable Effects on the Mouse Microbiome, mSystems 2 (4) (2017)

[41] J. M. Allen, L. J. Mailing, G. M. Niemiro, R. Moore, M. D. Cook, B. A. White, H. D. Holscher, J. A. Woods, Exercise Alters Gut Microbiota Composition and Function in Lean and Obese Humans, Med Sci Sports Exerc (2017)

[42] J. J. Faith, J. L. Guruge, M. Charbonneau, S. Subramanian, H. Seedorf, A. L. Goodman, J. C. Clemente, R. Knight, A. C. Heath, R. L. Leibel, M. Rosenbaum, J. I. Gordon, The long-term stability of the human gut microbiota, Science 341 (6141) (2013) 1237439.

[43] O. Cronin, O. O'Sullivan, W. Barton, P. D. Cotter, M. G. Molloy, F. Shanahan, Gut microbiota: implications for sports and exercise medicine, Br J Sports Med 51 (9) (2017) 700-701.

[44] J. P. Karl, L. M. Margolis, E. H. Madslien, N. E. Murphy, J. W. Castellani, Y. Gundersen, A. V. Hoke, M. W. Levangie, R. Kumar, N. Chakraborty, A. Gautam, R. Hammamieh, S. Martini, S. J. Montain, S. M. Pasiakos, Changes in intestinal microbiota composition and metabolism coincide with increased intestinal permeability in young adults under prolonged physiological stress, Am J Physiol Gastrointest Liver Physiol 312 (6) (2017) G559-G571.

[45] R. A. Koeth, Z. Wang, B. S. Levison, J. A. Buffa, E. Org, B. T. Sheehy, E. B. Britt, X. Fu, Y. Wu, L. Li, J. D. Smith, J. A. DiDonato, J. Chen, H. Li, G. D. Wu, J. D. Lewis, M. Warrier, J. M. Brown, R. M. Krauss, W. H. Tang, F. D. Bushman, A. J. Lusis, S. L. Hazen, Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis, Nat Med 19 (5) (2013) 576-85.

[46] W. H. Tang, Z. Wang, B. S. Levison, R. A. Koeth, E. B. Britt, X. Fu, Y. Wu, S. L. Hazen, Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk, N Engl J Med 368 (17) (2013) 1575-84.

[47] Z. Wang, E. Klipfell, B. J. Bennett, R. Koeth, B. S. Levison, B. Dugar, A. E. Feldstein, E. B. Britt, X. Fu, Y. M. Chung, Y. Wu, P. Schauer, J. D. Smith, H. Allayee, W. H. Tang, J. A. DiDonato, A. J. Lusis, S. L. Hazen, Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease, Nature 472 (7341) (2011) 57-63.

[48] B. J. Bennett, T. Q. de Aguiar Vallim, Z. Wang, D. M. Shih, Y. Meng, J. Gregory, H. Allayee, R. Lee, M. Graham, R. Crooke, P. A. Edwards, S. L. Hazen, A. J. Lusis, Trimethylamine-N-oxide, a metabolite associated with atherosclerosis, exhibits complex genetic and dietary regulation, Cell Metab 17 (1) (2013) 49-60.

[49] K. A. Romano, E. I. Vivas, D. Amador-Noguez, F. E. Rey, Intestinal microbiota composition modulates choline bioavailability from diet and accumulation of the proatherogenic metabolite trimethylamine-N-oxide, MBio 6 (2) (2015) e02481.

[50] H. J. Jang, D. M. Kim, K. B. Kim, J. W. Park, J. Y. Choi, J. H. Oh, K. D. Song, S. Kim, B. W. Cho, Analysis of metabolomic patterns in thoroughbreds before and after exercise, Asian-Australas J Anim Sci 30 (11) (2017) 1633-1642.

[51] G. Mithieux, Gut microbiota and host metabolism: what relationship?, Neuroendocrinology (2017)

[52] F. De Vadder, P. Kovatcheva-Datchary, C. Zitoun, A. Duchampt, F. Backhed, G. Mithieux, Microbiota-Produced Succinate Improves Glucose Homeostasis via Intestinal Gluconeogenesis, Cell Metab 24 (1) (2016) 151-7.

[53] N. Marungruang, J. Tovar, I. Bjorck, F. F. Hallenius, Improvement in cardiometabolic risk markers following a multifunctional diet is associated with gut microbial taxa in healthy overweight and obese subjects, Eur J Nutr (2017)

[54] J. U. Scher, A. Sczesnak, R. S. Longman, N. Segata, C. Ubeda, C. Bielski, T. Rostron, V. Cerundolo, E. G. Pamer, S. B. Abramson, C. Huttenhower, D. R. Littman, Expansion of intestinal Prevotella copri correlates with enhanced susceptibility to arthritis, Elife 2 (2013) e01202.

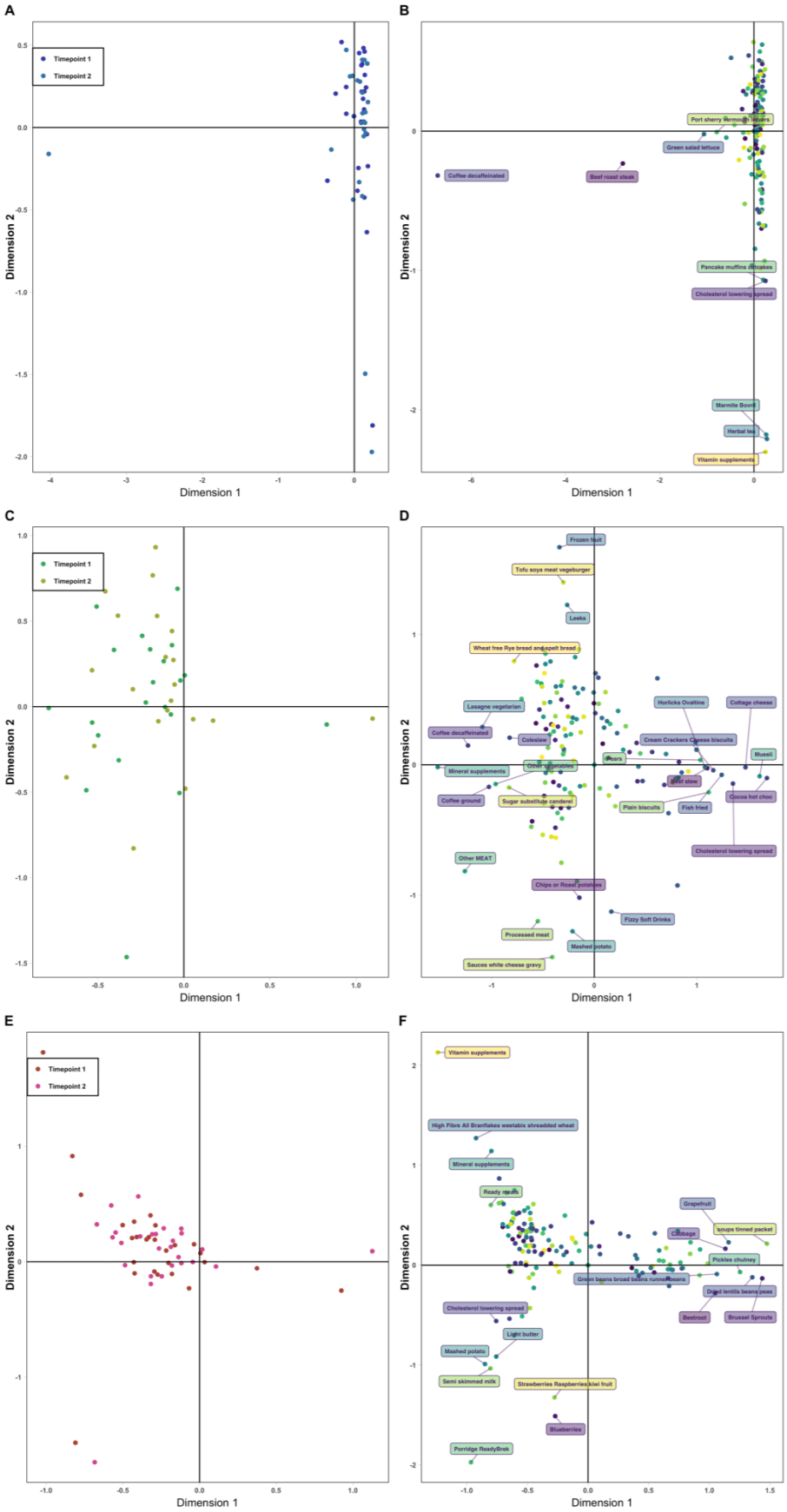
[55] A. Pianta, S. Arvikar, K. Strle, E. E. Drouin, Q. Wang, C. E. Costello, A. C. Steere, Evidence of the Immune Relevance of Prevotella copri, a Gut Microbe, in Patients With Rheumatoid Arthritis, Arthritis Rheumatol 69 (5) (2017) 964-975.

[56] C. Wen, Z. Zheng, T. Shao, L. Liu, Z. Xie, E. Le Chatelier, Z. He, W. Zhong, Y. Fan, L. Zhang, H. Li, C. Wu, C. Hu, Q. Xu, J. Zhou, S. Cai, D. Wang, Y. Huang, M. Breban, N. Qin, S. D. Ehrlich, Quantitative metagenomics reveals unique gut microbiome biomarkers in ankylosing spondylitis, Genome Biol 18 (1) (2017) 142.

[57] A. Z. Leite, N. C. Rodrigues, M. I. Gonzaga, J. C. C. Paiolo, C. A. de Souza, N. A. V. Stefanutto, W. P. Omori, D. G. Pinheiro, J. L. Brisotti, E. Matheucci Junior, V. S. Mariano, G. L. V. de Oliveira, Detection of Increased Plasma Interleukin-6 Levels and Prevalence of Prevotella copri and Bacteroides vulgatus in the Feces of Type 2 Diabetes Patients, Front Immunol 8 (2017) 1107.

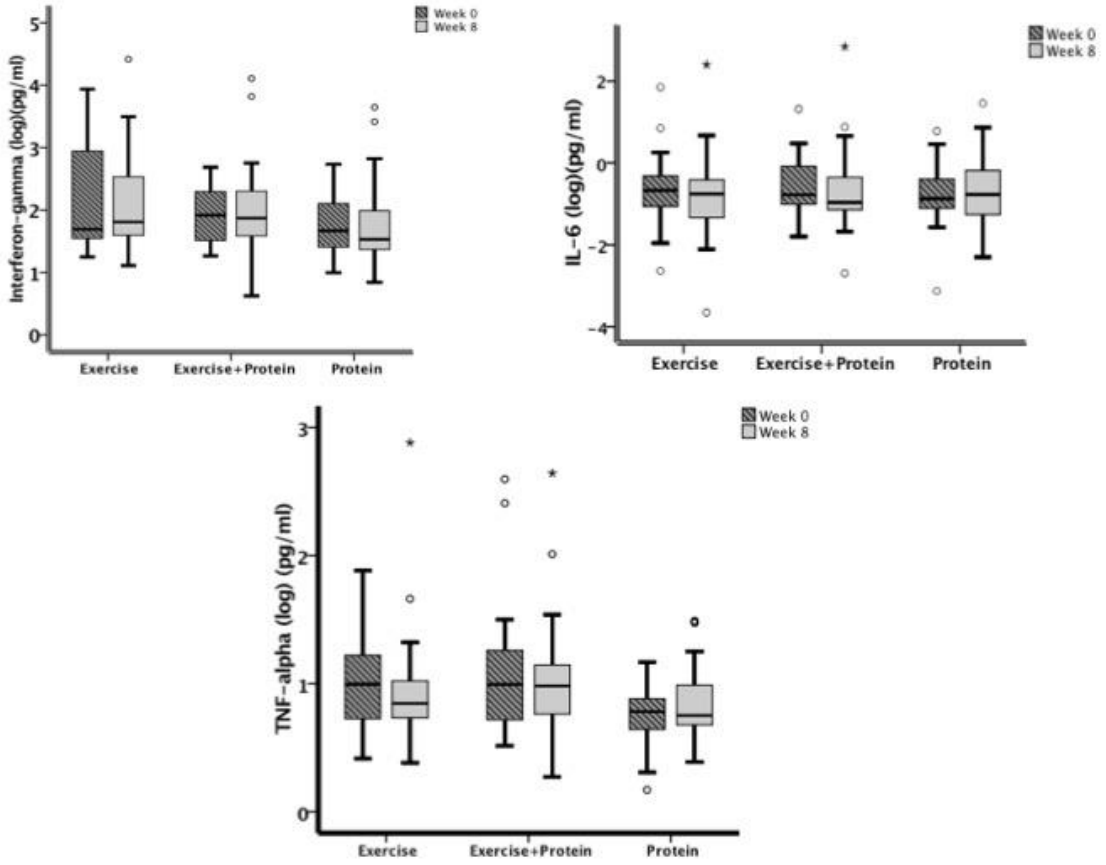
[58] C. T. Sempos, Invited commentary: some limitations of semiquantitative food frequency questionnaires, American Journal of Epidemiology 135 (10) (1992) 1127-1132.

6.7 | Supplementary Content

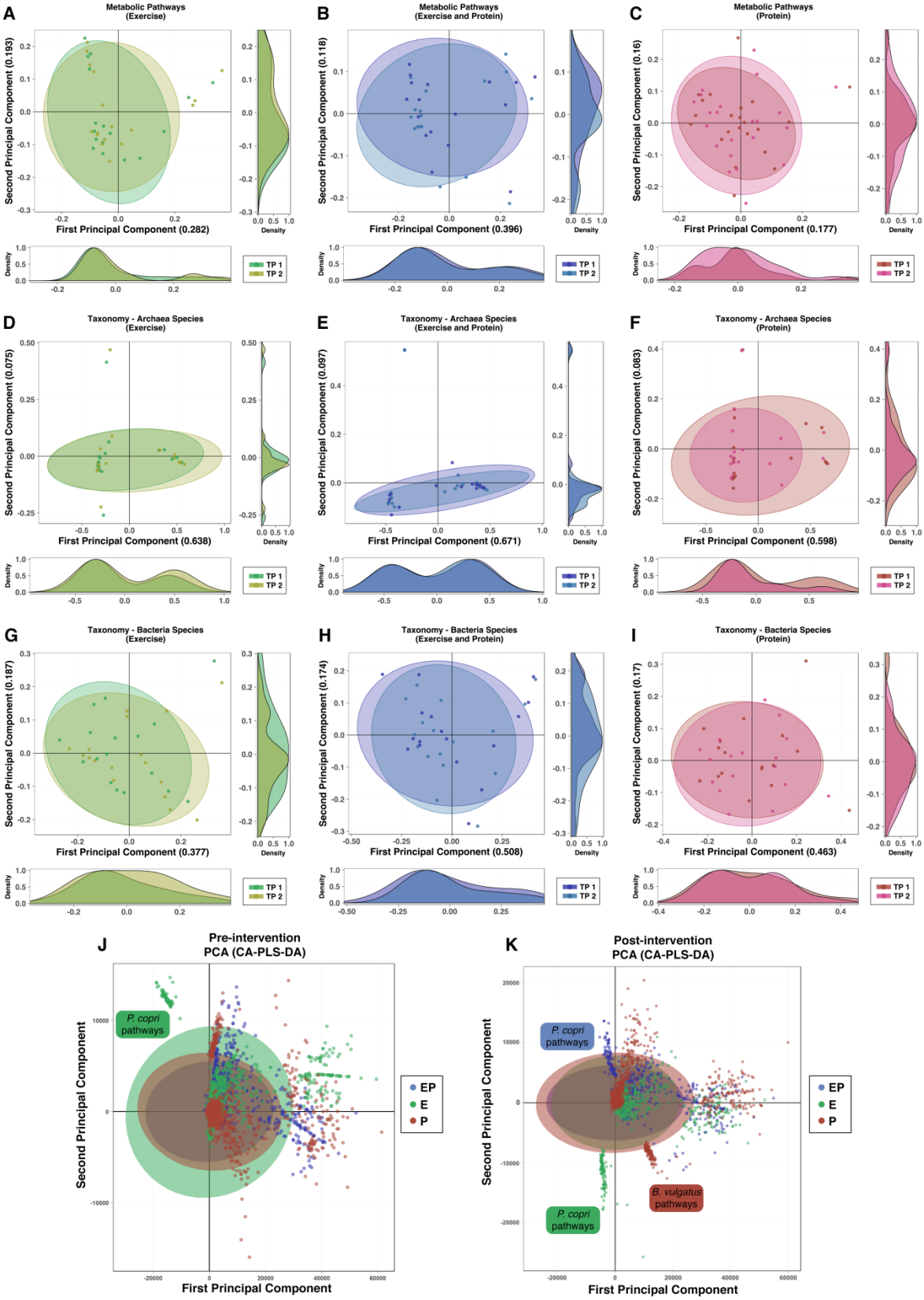


Supplementary Figure 6.1. Correspondence analysis of dietary intake of participants.

Supplementary Figure 6.1 continued. Self-reported dietary information from food frequency questionnaires (FFQ) for all participants was used to perform correspondence analysis of diet at the start and end of the intervention period. The exercise plus protein supplementation group (**A** and **B**), the exercise-only group (**C** and **D**), and the protein-only group (**E** and **F**) maintained diets that remained relatively unchanged during the intervention period. Participants were therefore undifferentiated before and after treatment in relation to diet.

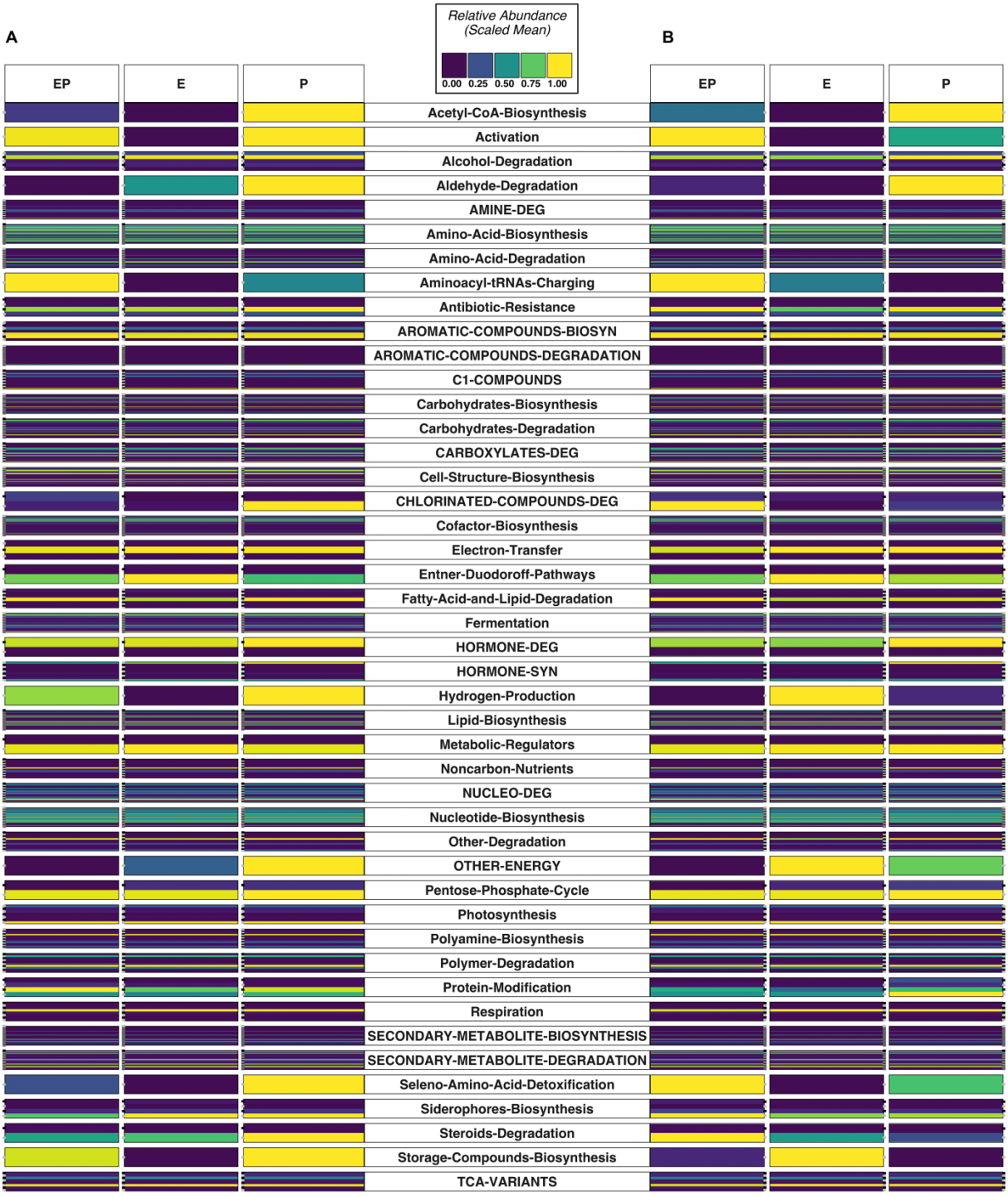


Supplementary Figure 6.2. Pre- and postintervention levels of the resting proinflammatory cytokines. Pre- and postintervention distribution of inflammatory cytokines (A) gamma interferon (IFN-γ), (B) IL-6, and (C) TNF-α. Dark central lines represent median values. Bars represent 95% confidence intervals. Outliers are denoted by hollow circles and extreme outliers by asterisks.



Supplementary Figure 6.3.

Supplementary Figure 6.3 continued. Pairwise analysis of microbial taxonomy and metabolic pathways and unsupervised feature detection of pathways. PCoA of metabolic pathways and species-level taxonomy measurement subset according to phylogenetic domain (Archaea and Bacteria), comparing measurements before and after intervention (time point 1 [TP1] and time point 2 [TP2], respectively). (**A** to **C**) Metabolic pathways for the (**A**) exercise-only group, (**B**) exercise plus protein supplementation group, and (**C**) protein-only group. (**D** to **F**) Archaea species for the (**D**) exercise-only group, (**E**) exercise plus protein supplementation group, and (**F**) protein-only group. (**G** to **I**) Bacteria species for the (**G**) exercise-only group, (**H**) exercise plus protein supplementation group, and (**I**) protein-only group. (**J** and **K**) Cross-validated partial least-squares-discriminant analysis (PLS-DA) of metabolic pathways from all three groups before (**J**) and after (**K**) the intervention period identified clusters of pathways related to specific taxa. (**J**) Prior to treatment, metabolic pathways associated with *Prevotella copri* were identified in the exercise group (**E**). (**K**) After intervention, each group contained a tightly associated cluster of pathways corresponding to either *Prevotella copri* or *Bacteroides vulgatus*. The two groups undergoing exercise (E and EP) had separate *Prevotella copri*-related clusters, while the protein-only-group cluster was composed of *Bacteroides vulgatus* pathways. Each experimental group was statistically assessed for alterations in diversity resulting from the intervention, with no significant separations detected. Statistical assessment of PCoA dissimilarity matrices was performed with the Adonis2 permutational multivariate analysis of variance (PERMANOVA) test. (**A** to **I**) Density plots were derived from kernel density estimates and scaled to a maximum estimated value of 1 and display concentrations of plotted data along the corresponding plot axis.



Supplementary Figure 6.4.

Supplementary Figure 6.4 continued. Categorical heat map of statistically significant metabolic pathways organized into low-level MetaCyc metabolic classifications. (**A** and **B**) Metabolic pathways of significantly ($P < 0.05$) varied relative abundances between groups before (**A**) and after (**B**) treatment were binned according to the MetaCyc database pathway classification. Scaled group means of pathway relative abundance values demonstrate shifts in the functional potential of the groups following the separate interventions. The Kruskal-Wallis test was used to calculate the *P* values used in the identification of pathways included in metabolic classification.

Inclusion Criteria
• Low physical activity level as per the International Physical Activity Questionnaire (Short form) • Not currently or recently (last 3 months) involved in regular or organized amateur sport or exercise • Non-smoker • Not on regular medications • Aged 18 to 40 • Body Mass Index between 22 to 35 (inclusive)
Exclusion Criteria
• Personal history of coronary artery disease, congenital heart disease or any cardiovascular disease • Family history of known coronary artery disease before 45 years of age • Uncontrolled hypertension (>140/90 mmHg) • Known renal or hepatic impairment • Type 1 or type 2 diabetes mellitus • Pulmonary disease – not including well-controlled, mild asthma • Primary or secondary immunodeficiency or autoimmune disorder • Current smoker or ex-smoker of less than 3 months duration • Psychiatric disorders including previous history of depression • A history of substance abuse • Current or recent involvement in another clinical research study • Known or suspected hypersensitivity to the dietary supplementation • Gastro-intestinal disease e.g. coeliac disease, inflammatory bowel disease, Irritable Bowel Syndrome • Previous significant gastro-intestinal surgery e.g. Total colectomy • Suspected or confirmed pregnancy

Supplementary Table 6.1. Inclusion and exclusion criteria for study entry.

Nutritional Information	Typical per 30g serving	Typical per 100g serving
Energy	125 kcal/521kJ	416 kcal/1743kJ
Protein	24 g	80 g
Carbohydrate *of which sugars*	1.5 g 0.9 g	5.0 g 3.0 g
Fat *of which saturates*	2.5 g 1.8 g	8.2 g 6.0 g
Dietary fibre	0.08 g	0.26 g
Salt	0.12 g	0.42 g

Supplementary Table 6.2. Nutritional content of the daily whey protein supplement (30grams) administered to participants in the EP and P study groups. g, grams; kCal, kilocalories; kJ, kilojoules.

	Intervention Groups			
	Exercise (E) (n=25)	Exercise + Protein (EP) (n=22)	Protein (P) (n=27)	p-value
Weight (kg)	-0.9 (-2.6, 0.9)	-0.8 (-1.6, 0.1)	-0.5 (-1.3, 0.6)	0.549
BMI (kg/m^2)	-0.3 (-0.9, 0.2)	-0.2 (-0.6, 0)	-1.1 (-0.4, 0.2)	0.419
Resting heart rate (BPM)	-5 (-16, 6) ∞	-5 (-9, 3) Ψ	4 (-3, 10)	0.005*
Systolic BP (mmHg)	-8 (-12, 1)	-8 (-16, 0)	-4 (-11, 0)	0.545
Diastolic BP (mmHg)	-5 (-12, 1)	-6 (-9, -2)	-5 (-8, 0)	0.785
Waist:Hip ratio	-0.01 (-0.03, 0.01)	-0.02 (-0.04, 0.01)	0 (-0.01, 0.04)	0.07
Body fat (%)	-1.3 (-2.4, -0.5) ∞	-0.8 (-1.7, -0.5) Ψ	0.5 (-0.2, 1)	<0.001*
Fat mass (kg)	-0.9 (-1.5, -2.7) ∞	-0.8 (-1.2, -0.4) Ψ	0.4 (-0.5, 0.9)	<0.001*
Fat mass (trunk) (kg)	-0.5 (-1, 02) ∞	-0.6 (-0.8, -1) Ψ	0.1 (-0.4, 0.6)	0.001*
Lean tissue mass (kg)	0.7 (0.3, 1.8) ∞	0.5 (-0.4, 1.1) Ψ	-0.2 (-0.9, 0.3)	0.001*
Weekly PA (METS)	1,159 (712, 1,964) ∞	1,265 (434, 2,487) Ψ	111 (-244, 634)	<0.001*
Weekly PA (kCals)	1,442 (818, 2,628) ∞	1,789 (571, 3,289) Ψ	184 (-418, 800)	<0.001*
Sitting time (hours/ week)	-5 (-17, 2)	-12 (-30, 1)	-5 (-18, 1)	0.407
Motorized transport (hours/week)	0 (-3.3, 2.8)	0 (-1, 1.3)	0.1 (-0.4, 5)	0.519

Supplementary Table 6.3

Supplementary Table 6.3 continued | Comparison of the postintervention changes (Δ) in clinical and anthropometric variables between groups. Between-group differences in postintervention changes were compared using Kruskal-Wallis tests (*P* values shown; *, $P < 0.05$). When significantly different a Mann-Whitney *U* test was applied to determine between which groups the difference existed. "∞" indicates a difference between the E and P groups; "Ψ" indicates a difference between the EP and P groups (for all data, $P < 0.05$). BMI, body mass index; BP, blood pressure; bpm, beats per minute; weekly PA, self-reported weekly physical activity expenditure assessed using the International Physical Activity Questionnaire; kCals, kilocalories; METS, metabolic equivalents.

Workload parameter	Exercise Group (E) (n=25)	Exercise + Protein Group (EP) (n=22)	p-value
Number of exercise sessions attended	21 (16, 23)	21 (20, 23)	0.317
Duration of aerobic training (mins)	671 (436, 728)	625 (539, 685)	0.983
Aerobic exercise energy expenditure (Cals)	6,043 (3754, 7411)	5,869 (5291, 7324)	0.654
Calories expended per body weight (Cals/kg)	69 (47, 90)	74 (62, 83)	0.685
Total number of repetitions	4,861 (4273, 6008)	4,874 (4281, 6136)	0.685
Total weight lifted (Tonnes)	171.6 (136.8, 205)	169.7 (111, 238.4)	0.701
Total weight lifted (kg) per kg of body weight	1,962 (1753, 2282)	2,288 (1464, 2746)	0.43

Supplementary Table 6.4. Measurement of aerobic and resistance training completed in the exercise and exercise plus protein intervention groups (8 weeks), with comparisons of workloads between the two groups. Median values and interquartile ranges (IQRs) are stated in the legend. Comparisons between groups were performed using Mann-Whitney *U* tests. Mins, minutes; Cals, calories expended.

	Exercise (E) Group (n=25)	Exercise + Protein (EP) Group (n=22)	Protein only (P) Group (n=27)	p-value
Interleukin 10 (pg/ml)	-0.05 (-0.21, 0.08)	0 (-0.23, 0.19)	0.02 (-0.11, 0.19)	0.385
Interleukin 6 (pg/ml)	-0.01 (-0.2, 0.2)	-0.04 (-0.28, 0.29)	0 (-0.17, 0.23)	0.893
Interleukin 8 (pg/ml)	-5.53 (-29.79, 1.37)	-20.14 (-124.47, 0.61) Ψ	-1.63 (-3.56, 1.38)	0.047*
TNF-α (pg/ml)	-0.1 (-0.66, 0.44)	-0.02 (-1.19, 0.77)	0.11 (-0.24, 0.34)	0.293
IFN-γ (pg/ml)	0.21 (-1.87, 3.39)	0.07 (-1.72, 2.83)	-0.41 (-1.57, 1.3)	0.962
CRP (mg/L)	0 (-0.5, 0)	0 (0, 0)	0 (0, 0)	0.71

Supplementary Table 6.5 | Comparison of postintervention changes (Δ) in inflammatory markers between groups. Legend: between-group differences in postintervention changes were compared using the Kruskal-Wallis test (*P* values shown; *, $P < 0.05$). Median changes and interquartile ranges are stated. Where data were significantly different, a Mann-Whitney *U* test was applied to determine between which groups the difference existed. "Ψ" denotes a significant difference between EP and P groups. TNF-α, tumour necrosis factor alpha; IFN-γ, interferon gamma; CRP, C-reactive protein.

Group	Domain	ANCOM Detected Features (FDR = 0.05)	Pre-treatment RA	Post-treatment RA	Whey Powder RA	Supplement Control RA
Exercise and Protein (EP)	Virus	s__Lactococcus_phage_Tuc2009	5.00E-04	2.10E-03	2.07E-03	0.00E+00
		s__Lactococcus_phage_TP901.1	1.67E-04	1.54E-03	9.81E-04	0.00E+00
		s__Lactococcus_phage_340	8.54E-03	2.44E-01	3.36E-03	5.97E-04
		s__Lactococcus_phage_jm2	8.41E-03	3.75E-02	1.87E-03	0.00E+00
		s__Lactococcus_phage_jm3	6.22E-03	3.58E-02	1.14E-03	0.00E+00
		s__Lactococcus_phage_P680	1.57E-02	1.40E-01	1.85E-03	0.00E+00
		s__Lactococcus_phage_phi7	8.34E-03	3.69E-02	1.16E-03	5.97E-04
		s__Streptococcus_phage_Alq132	1.81E-03	9.05E-03	8.40E-02	5.97E-03
		s__Streptococcus_phage_Sfi19	2.49E-03	4.53E-03	5.45E-02	2.39E-03
		s__Streptococcus_phage_DT1	8.56E-03	1.98E-02	2.07E-01	1.01E-02
		s__Streptococcus_phage_7201	3.59E-03	9.28E-03	1.58E-01	7.16E-03
		s__Streptococcus_phage_Abc2	3.86E-03	1.53E-02	1.23E-01	8.35E-03
		s__Lactococcus_phage_SK1	8.92E-04	8.26E-03	3.50E-04	0.00E+00
		s__Lactococcus_phage_bIL170	5.03E-03	4.16E-02	1.05E-03	5.97E-04
		s__Lactococcus_phage_P008	2.54E-03	4.05E-02	1.12E-03	0.00E+00
		s__Lactococcus_phage_712	1.87E-03	1.88E-02	1.81E-03	0.00E+00
		s__Lactococcus_phage_jj50	1.06E-03	1.35E-02	4.03E-04	0.00E+00
		s__Lactococcus_phage_bIBB29	6.20E-03	3.82E-02	8.33E-04	0.00E+00
	Bacteria	s__Vibrio_anguillarum	6.44E-06	1.33E-05	2.00E-06	5.00E-06
		s__Fervidobacterium_pennivorans	6.73E-06	1.08E-05	0.00E+00	1.25E-06
		s__Streptococcus_thermophilus	7.43E-04	1.28E-03	3.09E-01	5.64E-04
		s__Lactococcus_lactis	1.61E-04	3.45E-04	2.99E-01	6.35E-04
	Archaea	No significant OTUs detected	NA	NA	NA	NA
Exercise (E)	Virus	No significant OTUs detected	NA	NA	NA	NA
	Bacteria	s__Borrelia_hermsii	4.51E-06	2.32E-06	0.00E+00	0.00E+00
		s__Mycoplasma_pneumoniae	9.44E-07	6.61E-07	0.00E+00	0.00E+00

Group	Domain	ANCOM Detected Features (FDR = 0.05)	Pre-treatment RA	Post-treatment RA	Whey Powder RA	Supplement Control RA
	Archaea	No significant OTUs detected	NA	NA	NA	NA
Protein (P)	Virus	s__Streptococcus_phage_20617	7.11E-03	3.20E-03	9.97E-03	5.97E-04
		s__Lactococcus_phage_c2	1.05E-02	1.93E-02	4.62E-02	2.39E-03
		s__Lactococcus_phage_bIL67	2.29E-02	1.23E-02	1.83E-02	1.19E-03
		s__Lactococcus_phage_Tuc2009	2.45E-04	3.01E-03	2.07E-03	0.00E+00
		s__Lactococcus_phage_TP901.1	2.41E-04	1.89E-03	9.81E-04	0.00E+00
		s__Lactococcus_phage_r1t	6.31E-03	2.65E-03	3.14E-03	0.00E+00
		s__Lactococcus_phage_340	3.00E-02	2.33E-01	3.36E-03	5.97E-04
		s__Lactococcus_phage_jm2	1.61E-02	4.39E-02	1.87E-03	0.00E+00
		s__Lactococcus_phage_jm3	1.59E-02	4.12E-02	1.14E-03	0.00E+00
		s__Lactococcus_phage_P680	2.36E-02	1.45E-01	1.85E-03	0.00E+00
		s__Lactococcus_phage_phi7	1.48E-02	3.92E-02	1.16E-03	5.97E-04
		s__Streptococcus_phage_TP.778L	2.24E-04	9.90E-03	1.65E-02	1.19E-03
		s__Streptococcus_phage_Alq132	1.42E-03	1.09E-02	8.40E-02	5.97E-03
		s__Streptococcus_phage_Sfi19	1.43E-03	5.05E-03	5.45E-02	2.39E-03
		s__Streptococcus_phage_DT1	5.31E-03	1.83E-02	2.07E-01	1.01E-02
		s__Streptococcus_phage_7201	1.90E-03	1.09E-02	1.58E-01	7.16E-03
		s__Streptococcus_phage_Abc2	2.22E-03	1.66E-02	1.23E-01	8.35E-03
		s__Lactococcus_phage_SK1	5.05E-03	1.05E-02	3.50E-04	0.00E+00
		s__Lactococcus_phage_bIL170	1.42E-02	4.61E-02	1.05E-03	5.97E-04
		s__Lactococcus_phage_P008	1.24E-02	4.55E-02	1.12E-03	0.00E+00
		s__Lactococcus_phage_712	9.28E-03	2.24E-02	1.81E-03	0.00E+00
		s__Lactococcus_phage_jj50	8.44E-03	1.39E-02	4.03E-04	0.00E+00
		s__Lactococcus_phage_bIBB29	9.42E-03	3.93E-02	8.33E-04	0.00E+00
	Bacteria	s__Streptococcus_thermophilus	1.12E-03	2.21E-03	3.09E-01	5.64E-04
	Archaea	No significant OTUs detected	NA	NA	NA	NA

Table component A

p < 0.05
pFDR < 0.05

Urine Metabolites						
Metabolite	**E** Wilcoxon Rank	pFDR	**EP** Wilcoxon Rank	pFDR	**P** Wilcoxon Rank	pFDR
Allantoin	0.824	0.917	0.545	0.601	0.197	0.331
Carnitine	0.068	0.103	0.176	0.408	0.855	0.927
Citrate	0.105	0.299	1.000	0.862	0.747	0.862
Glycine	0.019	0.258	0.210	0.601	0.345	0.493
Hippurate	0.924	0.954	0.750	0.736	0.527	0.660
Leucine	0.030	0.248	0.679	0.917	0.623	0.959
PAG	0.059	0.164	0.775	0.954	0.023	0.093
Proline Betaine	0.633	0.617	0.702	0.507	0.584	0.533
Succinate	0.566	0.478	0.321	0.667	0.252	0.455
Trans Aconitate	0.874	0.991	0.371	0.125	0.004	0.046
TMAO	0.042	0.062	0.305	0.125	0.107	0.172
Valine	0.129	0.650	0.198	0.435	0.264	0.783

Faecal Metabolites						
Metabolite	**E** Wilcoxon Rank	pFDR	**EP** Wilcoxon Rank	pFDR	**P** Wilcoxon Rank	pFDR
Dimethylamine (DMA)	0.046	0.356	1.000	0.880	0.846	0.913
Glutamate	0.919	0.881	0.424	0.507	0.023	0.007
Methylamine (MA)	0.812	0.567	0.545	0.771	0.754	0.648
Phenylacetate	0.338	0.517	0.483	0.522	0.754	0.648
Serine	0.973	0.921	0.129	0.333	0.160	0.285
Trimethylamine (TMA)	0.865	0.746	0.424	0.569	0.777	0.862
Tyrosine	0.973	1.000	0.337	0.736	0.445	0.711

p < 0.05
pFDR < 0.05

Urine Metabolite Group Comparison							
Metabolite	Kruskal Wallis	Mann Whitney U					
	E vs EP vs P	**E vs P - p value**	**E vs P pFDR**	**E vs EP - p value**	**E vs EP pFDR**	**EP vs P - p value**	**EP vs P pFDR**
Allantoin	0.385	0.291	0.484	0.807	0.937	0.208	0.544
Carnitine	0.325	0.173	0.484	0.862	0.940	0.242	0.559
Citrate	0.392	0.272	0.484	0.211	0.544	0.957	0.957
Glycine	0.062	0.050	0.201	0.037	0.255	0.703	0.937
Hippurate	0.761	0.429	0.562	0.754	0.937	0.785	0.937
Leucine	0.579	0.468	0.562	0.299	0.559	0.887	0.940
PAG	0.006	0.001	0.014	0.311	0.559	0.057	0.255
Proline Betaine	0.959	0.878	0.878	0.771	0.937	0.922	0.948
Succinate	0.327	0.244	0.484	0.585	0.842	0.200	0.544
Trans Aconitate	0.087	0.322	0.484	0.522	0.784	0.014	0.169
TMAO	0.012	0.006	0.036	0.195	0.544	0.046	0.255
Valine	0.118	0.775	0.845	0.055	0.255	0.106	0.426

Faecal Metabolite Group Comparison							
Metabolite	Kruskal Wallis	Mann Whitney U					
	E vs EP vs P	**E vs P - p value**	**E vs P pFDR**	**E vs EP - p value**	**E vs EP pFDR**	**EP vs P - p value**	**EP vs P pFDR**
Dimethylamine (DMA)	0.508	0.316	0.890	0.329	0.890	0.991	0.991
Glutamate	0.028	0.100	0.890	0.656	0.891	0.005	0.105
Methylamine (MA)	0.696	0.485	0.890	0.764	0.891	0.492	0.964
Phenylacetate	0.623	0.363	0.890	0.691	0.891	0.551	0.964
Serine	0.607	0.428	0.890	0.366	0.890	0.902	0.991
Trimethylamine (TMA)	0.803	0.852	0.942	0.493	0.890	0.727	0.991
Tyrosine	0.445	0.609	0.891	0.524	0.890	0.202	0.706

Table component E

	Measure	Group	Metabolite			
			Allantoin	Carnitine	Citrate	Glycine
Urine Metabolites	Percent Change	E	0.040	-0.304	0.113	0.160
		EP	0.017	-0.323	-0.022	-0.186
		P	-0.145	0.013	-0.017	-0.096
	Mean change (SD)	E	1730 (21066)	-187501 (437691)	186718 (575441)	294032 (801563)
		EP	699 (20191)	-248992 (816086)	-33345 (472913)	-490924 (1411643)
		P	-6977 (27858)	6776 (586099)	-25697 (578360)	-186108 (634242)
	Median change (IQR)	E	625 (22071)	-115046 (442083)	181281 (654468)	261722 (370602)
		EP	2774 (16783)	-111501 (523672)	72494 (530470)	-148915 (956488)
		P	-4587 (32684)	13595 (443350)	64534 (506420)	-28847 (863822)

	Measure	Group	Metabolite			
			Hippurate	Leucine	PAG	Proline Betaine
Urine Metabolites	Percent Change	E	-0.007	0.048	-0.196	-0.185
		EP	-0.111	-0.011	-0.010	-0.127
		P	0.002	0.013	0.286	0.380
	Mean change (SD)	E	-4342 (271038)	2795 (8881)	-43452 (99096)	-73172 (406198)
		EP	-87427 (361172)	-727 (11098)	-2168 (129790)	-36254 (281728)
		P	1409 (476036)	777 (10687)	54206 (84419)	87703 (696375)
	Median change (IQR)	E	18143 (330942)	3821 (6034)	-55796 (109741)	-46720 (182241)
		EP	9770 (311037)	1544 (9325)	-10789 (152512)	-1626 (260571)
		P	-106522 (502221)	488 (11009)	28683 (132259)	-8146 (139201)

	Measure	Group	Metabolite			
			Succinate	Trans Aconitate	TMAO	Valine
Urine Metabolites	Percent Change	E	-0.082	0.050	-0.518	0.025
		EP	-0.032	0.008	0.005	-0.070
		P	0.138	-0.154	0.060	0.039
	Mean change (SD)	E	-22954 (113554)	1394 (14816)	-4324845 (10001137)	1181 (7793)
		EP	-9841 (109029)	219 (13674)	15418 (4520714)	-3787 (10604)
		P	35728 (138615)	-3965 (8235)	224051 (5442531)	1836 (7951)
	Median change (IQR)	E	-6659 (107283)	-476 (15042)	-1581160 (10243900)	1253 (5662)
		EP	-48579 (71425)	2013 (8208)	-449763 (1333768)	-906 (10053)
		P	27134 (206848)	-2690 (6076)	746835 (2277378)	717 (12098)

	Measure	Group	Metabolite			
			Dimethyla-mine (DMA)	Glutamate	Methyla-mine (MA)	Phenylacetate
Faecal Metabolites	Percent Change	E	-0.051	-0.013	-0.003	0.139
		EP	-0.008	0.076	-0.140	0.049
		P	-0.020	-0.172	0.158	-0.025
	Mean change (SD)	E	-3064 (6725)	-5301 (111833)	-925 (31567)	7368 (24144)
		EP	-472 (11748)	19772 (77762)	-12390 (47611)	2933 (17281)
		P	-1985 (18660)	-55203 (100712)	9971 (38476)	-1222 (17968)
	Median change (IQR)	E	-1636 (7024)	-13452 (123078)	-3237 (34478)	-159 (21834)
		EP	313 (10737)	3183 (93826)	1828 (49493)	1573 (17643)
		P	1205 (19734)	-56108 (85236)	-2431 (30736)	-895 (26770)

	Measure	Group	Metabolite		
			Serine	Trimethylamine (TMA)	Tyrosine
Faecal Metabolites	Percent Change	E	0.009	-0.148	0.008
		EP	0.081	0.051	0.117
		P	0.070	0.338	-0.054
	Mean change (SD)	E	2589 (51039)	-11542 (60276)	-103 (61576)
		EP	14876 (38326)	4361 (58358)	14693 (47246)
		P	12775 (44332)	24266 (74743)	-9190 (47972)
	Median change (IQR)	E	-3920 (46006)	2694 (38644)	-1613 (73144)
		EP	9085 (48256)	10580 (43674)	872 (54821)
		P	15010 (45823)	-3214 (45008)	-9094 (56306)

Table component F

Supplementary Table 6.6. The following combined data: (**A**) ANCOM results for statistically varied taxa of subjects and for corresponding taxa in whey and control supplements with relative abundance values and a list of significantly varied metabolic pathways and the (**B**) associated MetaCyc metabolic categories as well as (**C** & **D**) targeted faecal and urinary metabolites with (**E** & **F**) corresponding paired statistical results. Additional components in appendix B. Supplementary table 6.6B omitted due to length.

Chapter 7

INFLUENCE OF EXERCISE ON THE INTESTINAL MICROBIOME OF IBD PATIENTS

7.1 | Abstract

The human intestinal microbiome has been shown to be influenced by numerous aspects of human behaviour. One such behaviour, exercise, has been shown in professional athletes to confer considerable influence on the microbiome, while moderate-intensity structured physical activity has modest, but benign, influence on the gut microbiome of individuals unaccustomed to exercise. Despite the limited influence on the microbiome, rich systemic benefits of physical activity are gained in exercise naïve populations. Physical exercise, accordingly, presents itself as a potentially valuable adjunct therapy in disease populations with sensitive or otherwise vulnerable gastrointestinal symptoms. Here, in a population of IBD patients, it is assessed whether structured exercise can be applied without detrimental impact on disease symptoms while simultaneously gaining the health benefits of exercise. It was shown that indeed, short-term exercise did not profoundly modify the gut microbiome of IBD patients, while disease symptoms were not exacerbated and measurements of body composition and cardiorespiratory fitness improved.

In addition to the author, the original material subject of this chapter was contributed to accordingly:

Clinical work: Dr. Owen Cronin

Metabolomic processing and data analysis: Dr. Nicholas C. Penney, Dr. Isabel Garcia Perez, & Professor Elaine Holmes

Presented material has not been published at the time of this thesis' submission.

7.2 | Introduction

As described previously in chapter 6, moderate structured physical activity has been shown to confer a modest, but benign influence on the gut microbiome of individuals unaccustomed to exercise, both in respect to taxonomic composition and metabolite production [1]. Despite the limited influence on the microbiome, rich systemic benefits of physical activity are gained in exercise naïve populations. As such, physical exercise presents itself as a potentially valuable adjunct therapy in disease populations with sensitive or otherwise vulnerable gastrointestinal symptoms. In order to assess whether exercise can be implemented in such populations without detrimental impact on disease symptoms while simultaneously gaining the health benefits of exercise, the following chapter explores a study in which a population of patients with inflammatory bowel disease (IBD) underwent a structured moderate intensity exercise regime. This study granted access for the first time at the influence of physical activity on the IBD intestinal microbiome.

7.3 | Study objective

The primary objective of this chapter was to assess the viability of structured physical exercise in physically inactive IBD patients. Specifically, the chapter aims to determine the safety of exercise in respect to the patients' disease state and quality of life, and to further assess the status of the patients' gut microbiome in response to the exercise regime.

7.4 | Methods

7.4.1 | Experimental model and ethical approval

An overview of the study design is including in the Results section as Figure 7.1. The study was approved by the Cork Clinical Research Ethics Committee (CREC) before it com-

menced. Details of patient recruitment and experimental group assignment are described subsequently.

7.4.2 | Study recruitment and safe participation

Study recruitment included both male and female participants with confirmed diagnosis of Crohn's Disease, Ulcerative Colitis, or Indeterminate Colitis, and an age between 18 and 40 years (inclusive). Participants were recruited as patients from the gastroenterology outpatients' clinic at Cork University Hospital, Cork City, Ireland. Prior to enrolment, interested participants were screened by inclusion and exclusion criteria (described respectively in Tables 7.1 and 7.2). Only participants with stabilized IBD were enrolled to control for the influence of variable disease activity on study outcomes. Consultation with the attending clinician or review of recent medical documentation was used to assess participant disease activity. Recruitment was further unbiased on previous or current therapeutic intervention of the participants' medical condition.

Participants actively engaged in exercise or physical activities beyond that of daily life within 1 month prior to study commencement were excluded to more accurately assess the intervention. Baseline levels of physical activity were assessed using the International Physical Activity Questionnaire short form [2]. If appropriate, a subsequent screening visit at the study site was arranged for further assessment of the exclusion criteria. Safe participation in the exercise program was ensured by medical screening of all participants using an adapted version of the safe participation questionnaire of the American College of Sports Medicine [3].

7.4.3 | Intervention group allocation

Volunteers eligible for participation were randomized with 1:1 sequentially alternating allocation into the 2 intervention groups (immediate intervention : control+crossover).

During the control period, participants were instructed to maintain their usual levels of light physical activity.

7.4.4 | Combined exercise intervention

Combined aerobic and resistance training was performed at the Mardyke Arena gymnasium at University College Cork, Ireland. Complete details of the exercise intervention are described in section 2.4.

7.4.5 | Compliance and withdrawal from the study

Compliance with the prescribed exercise program was monitored remotely by the investigators using a FitLinxx activity monitoring system (Activelinxx, Shelton, CT), as detailed in section 2.2. Participants noted to have not complied with the exercise regime for more than 7 consecutive days were withdrawn from the study.

7.4.6 | Measurement visits

As with chapter 6 (section 6.3.6) measurement visits took place at 2 sites: Cork University Hospital and the Mardyke Arena, University College Cork. Baseline measurement was conducted within the 4 days prior to the commencement of the intervention period and once more after the 8-week intervention. Participants were asked to refrain from the use of alcohol, medication, and moderate to vigorous physical activity for at least 24 hours prior to measurement. To minimize potential effects of diurnal variation, measurement visits took place between 7:00 a.m. and 10:00 a.m. Initially, participants attended the Department of Medicine research facility at Cork University Hospital and sat restfully in a quiet environment. Participants proceeded to participate in measurement of clinical variables, e.g., recording of weight, blood pressure, and heart rate, before undergoing phlebotomy by a trained nurse using universal precautions. Following phlebotomy, individuals underwent

a total body dual-energy X-ray absorptiometry (DEXA) scan to assess body composition. For further detail refer to sections 2.3.3 and 2.2.4 When possible, volunteers were asked to provide fresh urine and faecal samples, which were transported at room temperature to the Teagasc Food Research Centre, Moorepark, Fermoy, Co. Cork, where DNA extraction took place. Following completion of the body composition assessment, participants proceeded to the indoor track at the Mardyke Arena gymnasium to undergo a submaximal cardiorespiratory fitness assessment as described in section 2.3.2.

7.4.7 | Inflammatory cytokine measurement

Blood samples (4 ml) from participants were collected in serum separator clot activator blood collection tubes (Greiner Bio-One, Stonehouse, United Kingdom; reference no. 454071). The blood samples were used to measure cytokine levels as described in section 2.3.4. All plasma samples were measured in duplicate, and the mean cytokine concentration of the duplicates (in picograms per millilitre) was used for analysis.

7.4.8 | Disease activity and quality of life assessment

Participants' quality of life and gastrointestinal symptoms, as influenced by the exercise intervention, were assessed with appropriate and widely recognised scoring systems. At each measurement visit patients reported current disease activity using the Harvey-Bradshaw Index (for patients with CD) [4], Simple Colitis Activity index (for patients with UC) [5]. Patients with Ulcerative Colitis were also assessed using the Powell-Tuck Index [6]. Likewise, quality of life, mood and anxiety levels were assessed regardless of subtype of IBD. The Short Form-36 (SF-36)[7], Hospital Anxiety and Depression Scale (HADS)[8], Beck Depression Inventory (BDI)[9], and the State/Trait Anxiety Inventory (STAI) [10]were used to measure change in quality of life and mood over the course of the intervention period.

7.4.9 | DNA extraction and metagenomic sequencing of faecal microbiome

Fresh faecal samples donated by participants were used for the extraction of DNA. Details of this and the subsequent process of shotgun metagenomic sequencing library preparation are described in sections 2.5.1 and 2.5.2, respectively. An equimolar library pool of all samples was made prior to sequencing on an Illumina NextSeq 500 (chemistry V.2.0) sequencing platform (Teagasc sequencing facility). High-throughput sequencing was performed using the high-output 500/550 reagent kit.

7.4.10 | Bioinformatic processing of microbial metagenomic sequencing

Processing of metagenomic FASTQ sequence files proceeded as described in previous chapters and section 2.6. Functional profiling of high-quality processed reads was facilitated by use of the Human Microbiome Project (HMP) Unified Metabolic Analysis Network (HUMAnN2 V.0.99) pipeline [11]. Taxonomic profiling was facilitated with the Kaiju taxonomic assignment software tool (V.1.5.0) [12].

7.4.11 | Metabolomic sample preparation

Processing of faecal and urine samples for metabolomic profiling is described in section 2.7.

7.4.12 | Quantification and statistical analysis

Statistical analysis was carried out using the Statistical Package for the Social Sciences V.23 (SPSS, Inc., Chicago, Illinois) and the R statistical programing environment (V.3.3.2), as more elaborately presented in section 2.8.2 - 2.8.4. Nonparametric analyses were performed to accommodate non-normal distributions of clinical and demographic variables, as well as for the analysis of microbiome and metabolomics data. Clinical data are presented as medians and interquartile ranges (IQR), unless stated otherwise.

Between-group differences in baseline, follow-up, and postintervention changes (Δ) in clinical and demographic data were compared using the Kruskal-Wallis test. For significantly different results, a Mann-Whitney *U* test was performed to determine the groups between which this difference applied. Where stated, the Wilcoxon signed-rank test was used to compare baseline and postintervention values within the groups. A type I error rate of ≤0.05 was considered significant in all cases. Correction of *P* values relating to microbiome and metabolomic analysis was performed using the Benjamini-Hochberg false-discovery rate (FDR) [13] in the base *stats* package in R. Statistical assessment of dissimilarity matrices (Bray-Curtis) derived from microbial data was facilitated with the adonis2 function in the vegan R package (V.2.4-3) [14]. Identification of statistically relevant taxonomic features was performed with the analysis of composition of microbiomes (ANCOM) test as implemented in the R package of the same name (V.1.1-3) [15]. Measurements of alpha diversity and calculations of relative abundances were also performed with the vegan package. Relative-abundance data were generated separately for identified species within each phylogenetic domain (e.g., *Bacteria*).

For metabolomic analysis, the resulting ^{1}H-NMR and LC-MS data sets were imported into MatLab to conduct multivariate statistical analysis. Data were centred and scaled to account for the repeated-measures design and then modelled using partial-least-squares–discriminant analysis (PLS-DA) with Monte Carlo cross-validation (MCCV) [16]. The fit and predictability of the models obtained were determined and expressed as R2 and Q2 values, respectively.

7.4.13 | Data and software availability

The microbial DNA sequences have been deposited in the European Nucleotide Database (ENA) database under ID code PRJEB27623.

7.5 | Results

7.5.1 | Study overview

Male and female patients with diagnosed IBD (n = 17) aged between 30 and 37 years with BMI ranging from 23 to 25kg/m^2 were recruited for participation (Table 7.1). The study was conducted in accordance with the Declaration of Helsinki and was granted ethical approval by the Cork Clinical Research Ethics Committee (CREC) prior to commencement. Written informed consent was provided by all participants. To prospectively measure the effect of *de novo* exercise training on both the gut microbiome and participants' physiology, physical inactivity beyond that of daily life prior to 3 months of the study's beginning was required of patients for inclusion in the study (Supplementary Table 7.1). Participants were screened against specific criteria for exclusion, including history of cardiovascular disease and uncontrolled hypertension (Supplementary Table 7.2). Following recruitment, participants were randomly allocated to one of two groups: immediate 8 weeks of exercise (EX) or 8 weeks of normal activity followed by 8 weeks of exercise (CON) (Figure 7.1). During the 8 weeks of normal activity participants were asked to maintain their typical levels of activity. Throughout the study period all participants were requested to refrain from taking vitamin, dietary, and herbal supplements, and to maintain usual *ad libitum* dietary intake.

While engaged in the exercise program, participants underwent combined aerobic and resistance training. Aerobic exercise was standardized, progressive, and similar in energy expenditure to a "couch-to-5km-running" program. The intensity of aerobic exercise was moderate, being graded between 5 and 7 of 10 on the modified Borg rating of perceived exertion (RPE) scale [17]. Resistance training consisted of 7 machine-based resistance exercises. Starting weights were calculated at induction at 70% of the individual's one-repetition maximum value (1RM). Subjects were required to perform a minimum 3 sets of 8

repetitions. Resistance training was progressive with aims of increasing resistance weight by 15% to 20% over the 8-week period. At study entry all participants were generally classified as overweight with BMI > 25 and body fat composition > 30%. There were no significant differences in anthropometric and demographic characteristics between patients in the two groups at baseline (Table 7.1). Disease characteristics were similar between the EX and CON groups, with no significant variations (Table 7.2).

7.5.2 | Improvements in body composition and cardio-respiratory fitness following exercise training period

Of the 110 IBD patients that were invited to participate in the study, 17 were suitable for participation according to inclusion and exclusion guidelines and accepted enrolment (Figure 7.1 and Supplementary Tables 7.1 & 7.2). During the exercise intervention period, patients in the EX group underwent an average of 22 out of 24 possible gym sessions, averaging 689 minutes of cardio-respiratory exercise and 4,782 repetitions of resistance training. Statistically significant changes in estimated maximal oxygen consumption (VO_{2max}) and heart rate following completion of the Rockport one-mile walk test were identified at the conclusion of the exercise period (*P* values <0.05, Table 7.5B). Patients in the EX group demonstrated significant improvements in body composition compared to CON group counterparts. Total body fat (%) was significantly reduced (*P* value <0.01) and lean tissue mass was elevated for EX participants (*P* value <0.001, Table 7.3A).

Patient characteristics	Values[a] Exercise group (n = 13)	Control group (n = 7)	*P* value
Age (years)	33 (31, 36)	31 (31-36)	0.938
Sex (female)[b]	n = 4 (30.8%)	n = 1 (14.3%)	0.417
Height (cm)	172 (165, 179)	173 (167, 183)	0.485
Weight (kg)	84.5 (75.8, 97)	89 (70.5, 102.9)	0.757
BMI (kg/m^2)	28.1 (26.2, 32.4)	27.2 (24.5, 33.7)	0.938
Resting heart rate (bpm)	73 (64, 77)	76 (58, 83)	0.817
Systolic BP (mmHg)	123 (114, 131)	124 (119, 130)	0.485
Diastolic BP (mmHg)	73 (69, 86)	82 (70, 86)	0.311
Waist:Hip ratio	0.96 (0.85, 0.98)	0.97 (0.86, 1.00)	0.643
Body fat (%)	35.2 (30.9, 37.4)	34.2 (32.2, 35.7)	0.817
Fat mass (kg)	28.1 (22.8, 35.7)	29.9 (23.8, 34.7)	0.938
Fat mass (trunk) (kg)	14.7 (12.9, 21.3)	16.8 (13.1, 22.2)	0.699
Lean tissue mass (kg)	50.2 (44, 61.6)	52.69 (45.8, 63.8)	0.643
Weekly PA (IPAQ, METS)	708 (210, 1,036)	180 (0, 606)	0.056
Weekly PA (IPAQ, kCals)	1,017 (283, 1,399)	266.9 (0, 1,039)	0.097
Sitting time (h/week)	52.5 (26.8, 65)	59.5 (38.5, 82)	0.241
motorized transport (h/week)	7 (6.13, 10.5)	7 (3, 9.33)	0.438

Table 7.1 | Baseline measurements of patients' demographic and anthropometric characteristics. [a]Values represent medians (interquartile ranges) except where otherwise indicated. *P* values represent results of Wilcoxon-Mann-Whitney U or Chi-squared tests. BMI, body mass index; IPAQ, International Physical Activity Questionnaire; METS, metabolic equivalents.

[b]Data indicates use of Chi-squared test.

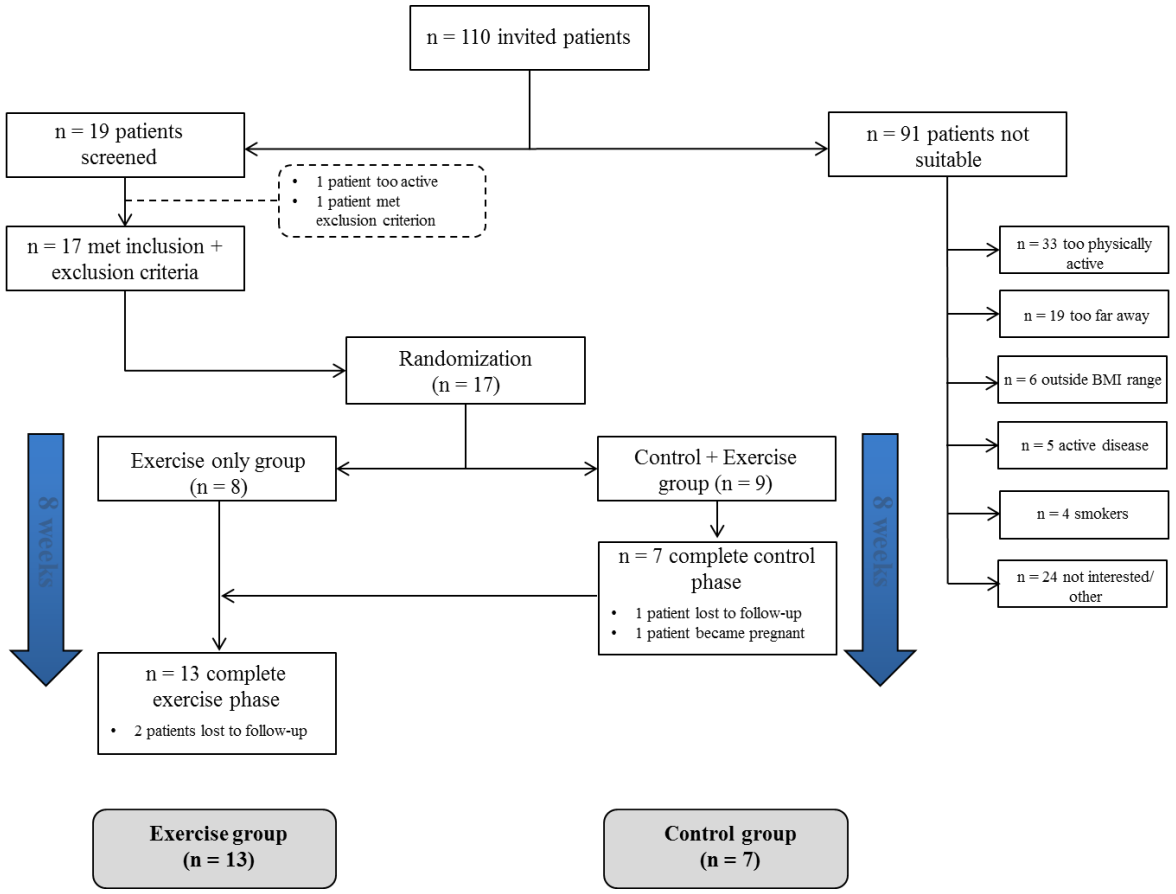


Figure 7.1 | Study outline. The details of study recruitment and allocation of participants into EX and CON groups are described in addition to details of patient drop out from the study. Complete details of exclusion and inclusion criteria are presented in Supplementary Tables 7.1 & 7.2.

Disease characteristics	Values[a] Exercise group (n=13)	Control group (n=7)	*P* value
Condition (%)[b]:			
Ulcerative Colitis	8 (61.5%)	5 (71.4%)	0.658
Crohn's Disease	5 (38.5%)	2 (28.6%)	
Age at diagnosis (years)	25 (25, 31)	26 (22, 31)	0.642
Baseline Disease activity score:			
Simple Colitis Index	1.5 (1, 2.75)	1 (0, 4)	0.833
Powell Tuck Score	1 (0.25, 2)	2 (0, 3)	0.833
Harvey Bradshaw Index	1 (0, 1.5)	1 (0, 1)	1.00
Previous surgery for condition[b]	Yes: 1 (7.7%)	Yes: 0	0.452
Previous appendicectomy[b]	Yes: 1 (7.7%)	Yes: 0	0.452
Current steroid use	None	None	
Steroid use in the last year (%)[b]	Yes: 5 (38.5%)	Yes: 3 (42.9%)	0.848
Current anti-TNF treatment[b]	Yes: 2 (15.4%)	Yes: 1 (14.3%)	0.948
Previous anti-TNF treatment	None	None	
Currently on immunomodulator	Yes: 100%	Yes: 100%	
Patients on azathioprine	N = 1	N = 1	
Patients on mesalazine	N = 11	N = 6	
Patients on 6-mercaptopurine	N = 4	N = 2	
Patients on PPIs	N = 4	N = 4	

Table 7.2 | Disease characteristics of EX and CON groups. [a]Values represent medians (interquartile ranges) except where otherwise indicated. *P* values represent results of Wilcoxon-Mann-Whitney U or Chi-squared tests. PPI, proton pump inhibitor.

[b]Data indicates use of Chi-squared test.

A	**Values**[a]		
Body composition profiles	**Exercise group (n=13)**	**Control group (n=7)**	***P* value**
Total body fat percentage (%)	-2.1 (-2.15, -0.45)	0.1 (-0.4, 1)	0.002
Fat mass (kg)	-1.52 (-2.32, -0.04)	-0.04 (-0.93, -0.04)	0.081
Trunk fat mass (kg)	-0.82 (-1.4, -0.002)	-0.13 (-0.8, 0.78)	0.115
Lean tissue mass (kg)	1.59 (0.68, 2.69)	-1.38 (-2.45, 0.26)	<0.001
B	**Values**[a]		
Fitness profile of EX	**Week 0**	**Week 8**	***P* value**
Estimated VO_{2max} (mls/kg/min)	43.41 (37.44, 46.37)	46.01 (42.05, 49.47)	0.03
Completion heart rate (Rockport test, BPM)	155 (147, 166)	145 (129, 150)	0.03

Table 7.3 | Fitness parameters of EX following exercise and in comparison to CON group.

(**A**) Comparison of EX and CON patients' body composition. EX patients showed significant reduction in total body fat percentage and significant increase in lean tissue mass. (**B**) Significant improvements were shown in patients following exercise intervention for both estimated maximal oxygen consumption and heart rate.

[a]Values represent medians (interquartile ranges) except where otherwise indicated. *P* values represent results of Wilcoxon-Mann-Whitney U test. BPM, beats per minute; VO_{2max}, estimated maximal oxygen consumption.

7.5.3 | Metagenomic assessment of gut microbiota after exercise and control periods

Species level taxonomic profiles of patients' gut microbiota showed modest alterations in α-diversity. Comparisons of Archaea species at the beginning and end of the control and exercise periods demonstrated minor increases in α-diversity (n.s.; Figure 7.2A & B). Mean α-diversity of Bacteria species increased significantly in participants following the control period (P = 0.015; Figure 7.2C), while a moderate decrease was present in the patients following exercise (n.s.; Figure 7.2D). Comparisons of within-group alterations of α-diversity (percent Δ) showed no significant differences between EX and CON groups (Figure 7.2E & F).

Assessment of taxonomic β-diversity with non-metric multidimensional scaling (NMDS) illustrated minimal changes in bacterial species for both the CON and EX groups between week 0 and week 8 (Figure 7.3A & B). Similarly for Archaea species, principal-coordinate analysis (PCoA) demonstrated little separation for both groups (Figure 7.3C & D). Taxonomic profiling of virus species did not report any viruses at the species level of phylogeny (data not presented). Statistical testing of taxonomic relative abundance did not identify any species that were significantly varied between the groups or measurement time points after correction for multiple comparisons (data not presented).

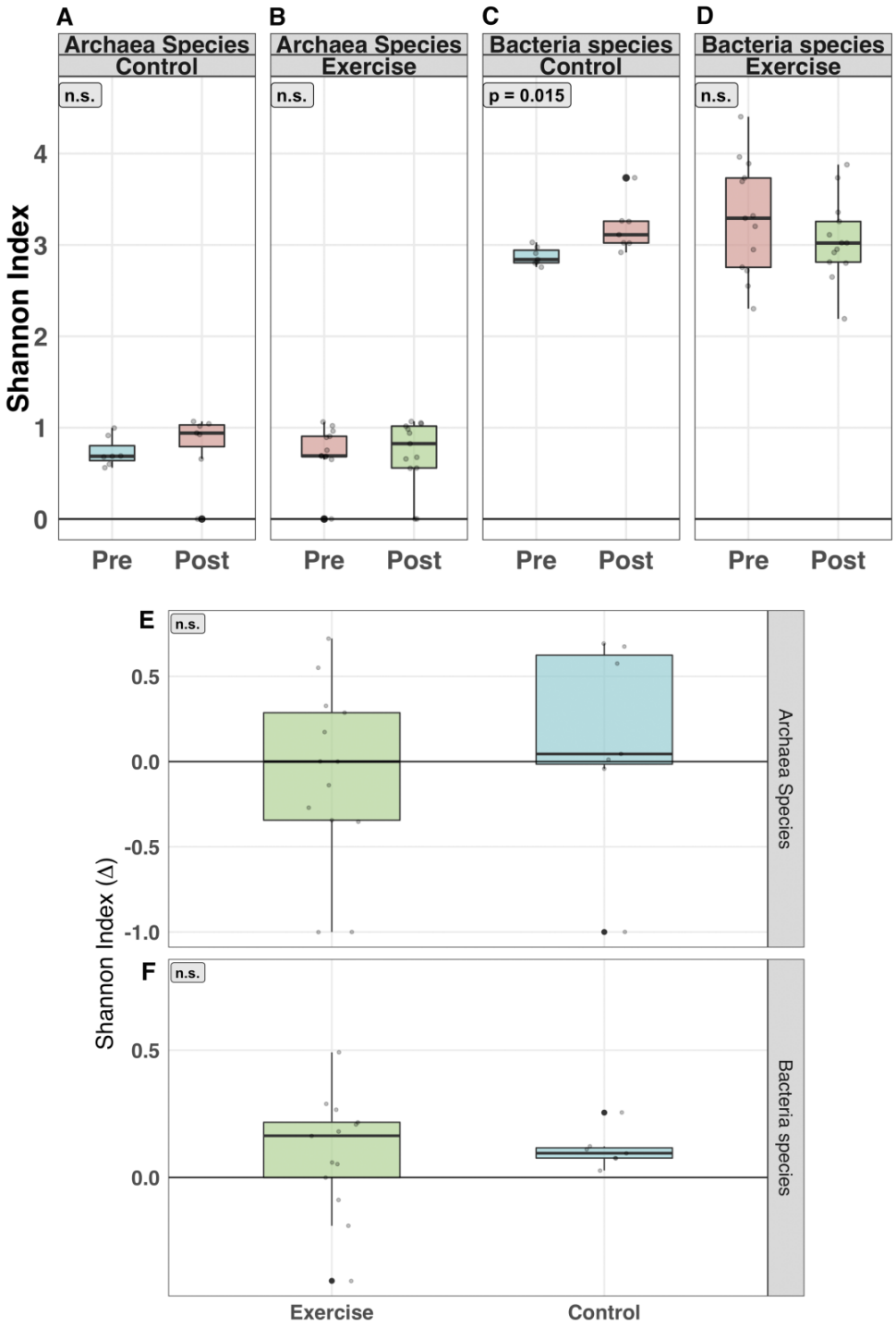


Figure 7.2 | *α*-diversity of taxonomic profiling. (**A-D**) Shannon *α*-diversity H-index of Bacteria and Archaea species at week 0 (pre) and week 8 (post). (**A** & **B**) Subtle median increases of Archaea *α*-diversity for patients in both treatments. (**C** & **D**) Bacterial *α*-diversity was significantly raised in the control group ($P = 0.015$), while a moderate median decrease was shown in the exercise group. (**E**) Percent change (Δ) of *α*-diversity for Archaea species shows a slight increase in the CON group compared to EX. (**F**) Both EX and CON show an increase in *α*-diversity for Bacteria species. *P* values were calculated from the Wilcoxon signed-rank test.

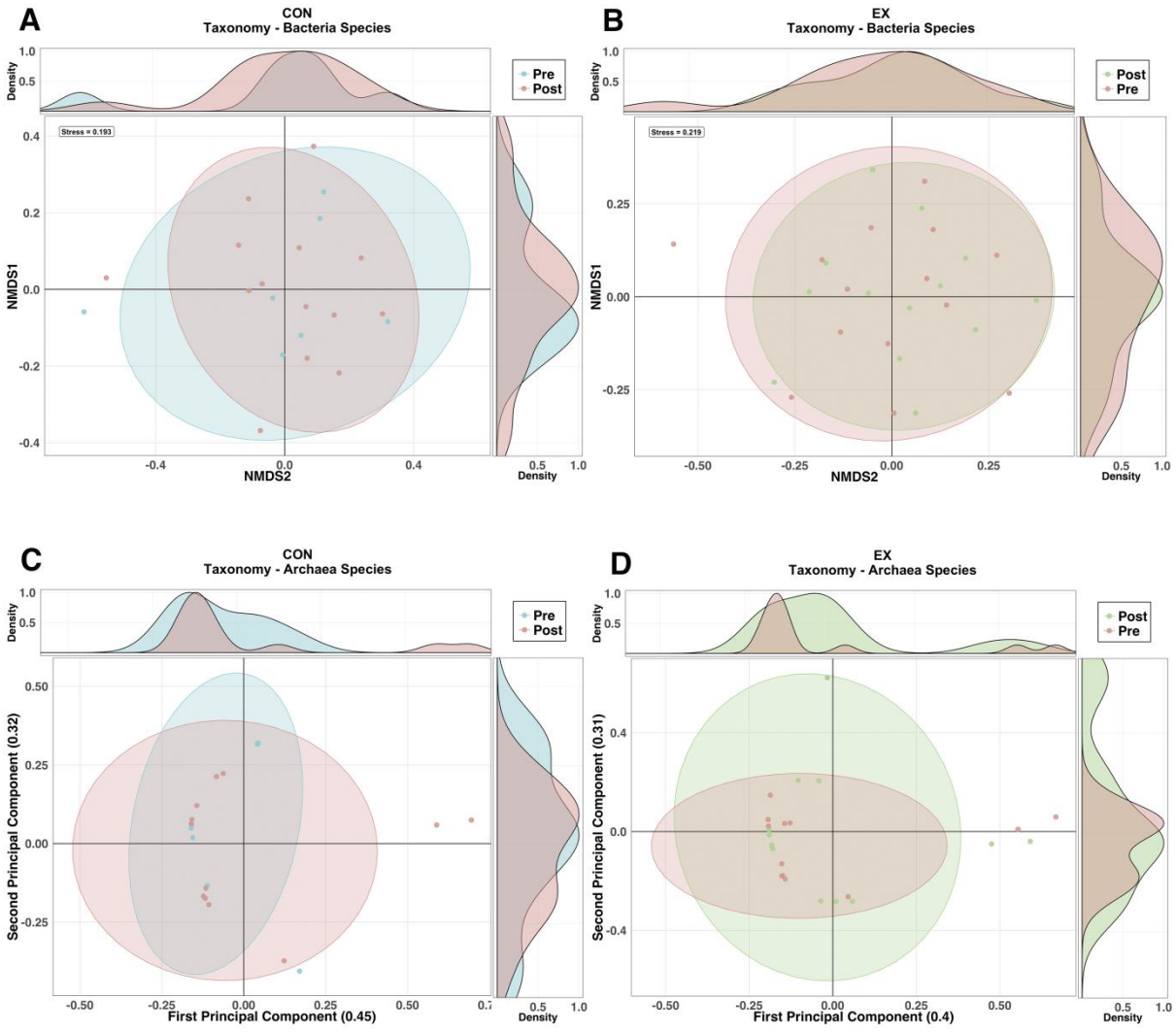


Figure 7.3 | β-diversity of taxonomic profiles. (**A** & **B**) Non-metric multidimensional scaling (NMDS) of relative abundance profiles for Bacteria species demonstrates minimal shifts in diversity for CON (**A**) and EX (**B**) following the 8 week treatment period. (**C** & **D**) Principal-coordinate analysis (PCoA) of Archaea relative abundance similarly shows subtle changes in the distribution of patients for both CON (**C**) and EX (**D**). Statistical assessment of dissimilarity matrices was performed with the Adonis2 permutational multivariate analysis of variance (PERMANOVA) test. Density plots along the axes of all panels display the concentrations of data points, and were generated from kernel density estimates and scaled to a maximum value of 1.

7.5.4 | Pathway assessment

General (e.g., fucose degradation) and taxonomically specific (e.g., fucose degradation in *Akkermansia muciniphila*) metabolic pathways models generated from metagenomic sequencing were compared, both between groups and time points. Statistical assessment of the metabolic pathways illustrated that no significant differences were present within the groups (i.e., week 0 vs. week 8) or between them (data not shown). Measured α-diversity (Shannon index) of pathway models demonstrated negligible variation of both the general pathways (Figure 7.4A & B) and species-specific pathways (Figure 7.4C & D) before and after intervention periods for both groups. Within-group alterations of pathway α-diversity (percent Δ) showed subtle but non-significant increases for both general and species-specific pathways in the EX group while a decrease was observed in the CON group for species-specific pathways (Figure 7.4E & F). Similarly to the assessment of taxonomic β-diversity, NMDS of both the species-specific (Figure 7.5A & B) and general pathways (Figure 7.5C & D) illustrated minimal separation of patients from week 0 to week 8.

7.5.5 | Metabolomic profiling

Untargeted metabolomic analysis of the full spectrum of urinary metabolites from patient samples presented no significant separations either within each group before and after intervention or between groups (data not presented).

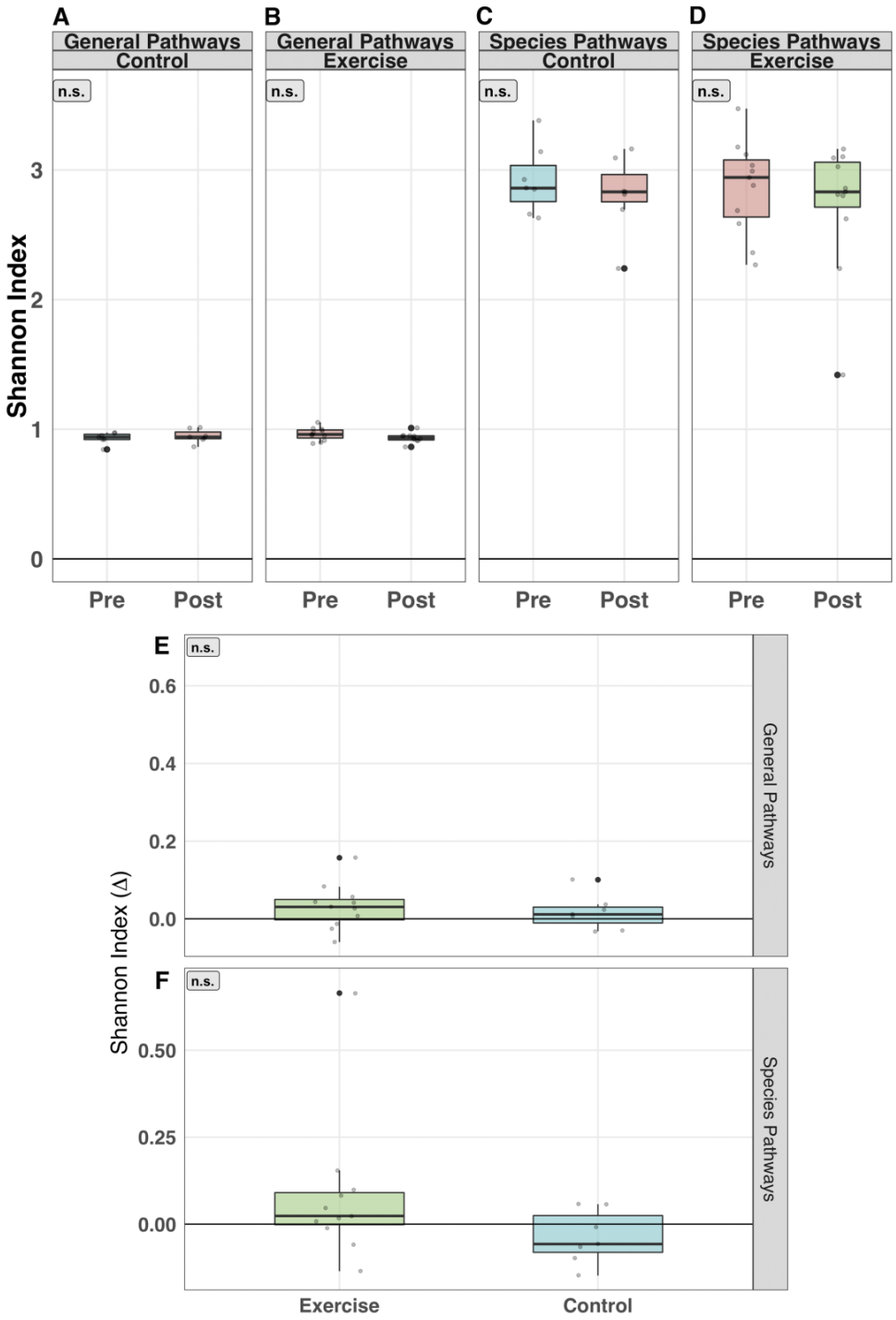


Figure 7.4 | α-diversity of metabolic pathways. (**A-D**) Shannon α-diversity H-index of general and species-specific metabolic pathways at week 0 (pre) and week 8 (post). (**A** & **B**) No significant changes detected for general pathways of patients in either treatment. (**C** & **D**) A slight decrease of species-specific pathways are shown in patients after 8 weeks of exercise. (**E**) Percent change (Δ) of α-diversity for general (**F**) and species-specific pathways shows a slight increase in the EX group for both, while a decrease is observed for the CON group in species-specific pathways. *P* values were calculated from the Wilcoxon signed-rank test.

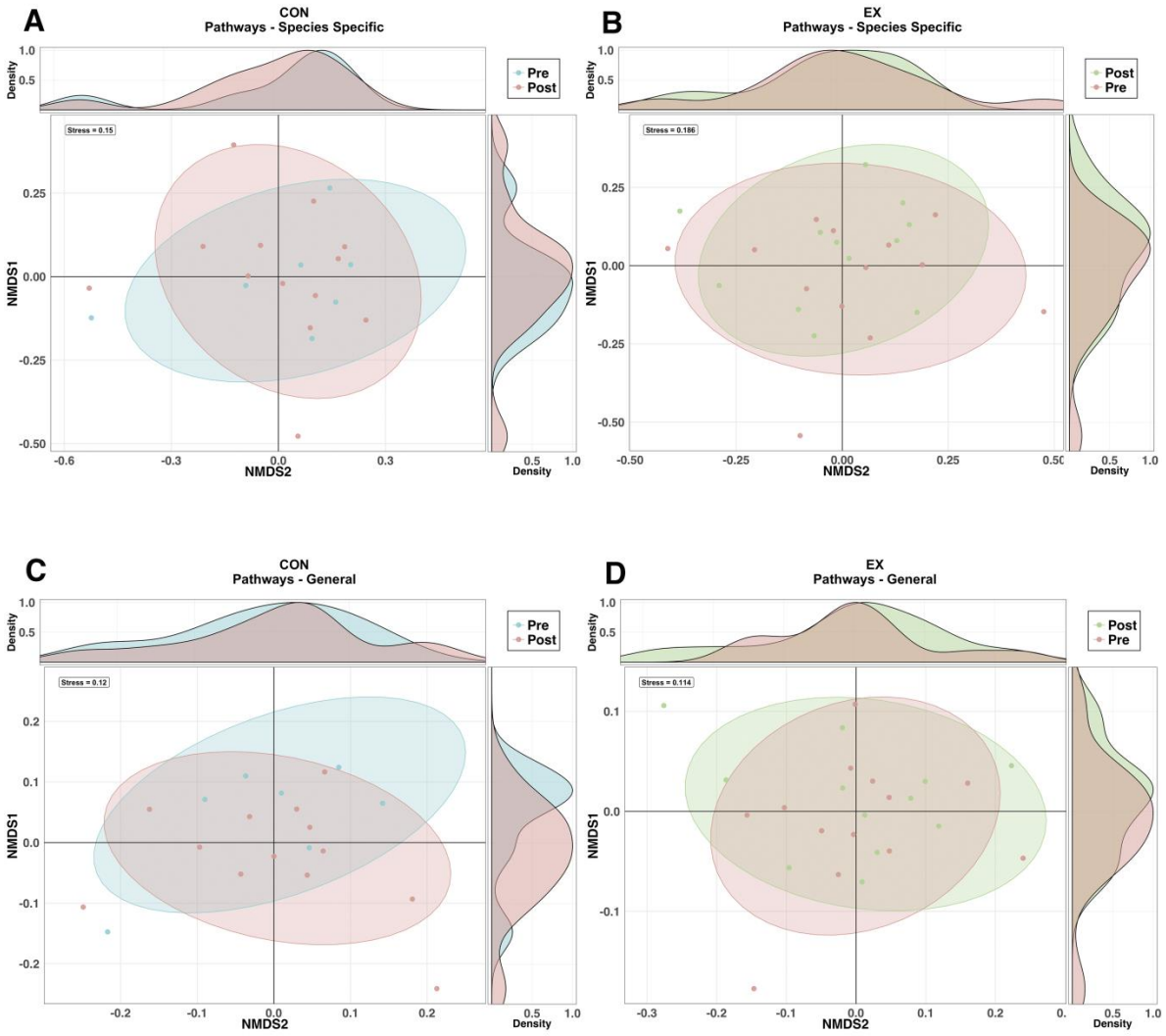


Figure 7.5 | β-diversity of metabolic pathways. Demonstrating the distribution of patients in relation to the metabolic pathway models with non-metric multidimensional scaling (NMDS), no significant separations were detected. (**A** & **B**) Species-specific pathways are minimally altered in diversity for CON (**A**) and EX (**B**) following the 8 week treatment period. (**C** & **D**) General pathways similarly had subtle changes in distribution for patients of both CON (**C**) and EX (**D**). Statistical assessment of dissimilarity matrices was performed with the Adonis2 permutational multivariate analysis of variance (PERMANOVA) test. Density plots along the axes of all panels display the concentrations of data points, and were generated from kernel density estimates and scaled to a maximum value of 1.

7.6 | Discussion

We have previously assessed the influence of the extremes of athleticism on the gut microbiome with elite athletes [18, 19], and the impact of exercise and or increased dietary protein intake on the gut microbiota of exercise naïve individuals [1]. While the athlete microbiome has presented a multitude of differences when compared to healthy but comparatively less-active controls, likely resulting from adaptations to lifelong rigorous physical activity, exercise naïve populations present more subtle changes in gut microbiome form and function. Accordingly, we had interest in assessing the influence of exercise on a human population with particularly vulnerable GI health.

The IBD patients recruited for this study were an ideal cohort to assess if an exercise intervention could be utilized without detrimental perturbations to the gut microbiome while still transferring the benefits of exercise. Our results indicate that similarly to our previous study, only subtle changes in the gut microbiome were transferred by structured exercise. Patients undergoing the short-term combined aerobic and resistance training presented no significant alterations in the composition or functional capacity of their gut microbiota. Slight increases in archaeal α-diversity were shown for both CON and EX treatments after 8 weeks, while the α-diversity of Bacteria species was significantly raised for the CON and marginally decreased for the EX patients (Figure 7.2). Our previous study of the influence of exercise on the microbiome illustrated a reduction in Archaea diversity resulting from a similar exercise regime implemented in this study, suggesting that intestinal Archaea are suppressed by exercise. Here it is shown that the median percent change (Δ) of archaeal species in EX patients is effectively at zero, compared to a slight increase in CON patients. As Archaea are potentially implicated in some intestinal disorders, and may modulate health-relevant compounds such as trimethylamine N-oxide

(TMAO), this study offers further evidence for the potential interaction of Archaea and exercise [20, 21].

The significant increase in bacterial α-diversity among CON patients is surprising, and may highlight a subconscious alteration in behaviour (e.g. dietary habits) of the patients in anticipation of beginning an exercise regime. Assessment of taxonomic β-diversity further supports the limited alteration of the patients' gut microbiome composition (Figure 7.3). Similarly, the functional capacity of the gut microbiota, presented as metabolic pathway models, demonstrated that the short-term moderate intensity exercise regime utilised in this study does not induce detrimental modulation of the gut microbiome's functional potential (Figures 7.4 & 7.5). General pathways were nearly unchanged in both CON and EX treatments, while those pathways that are associated to specific taxa did see a slight percent change (Δ) shift of increased diversity in EX and lowered diversity for CON (Figure 7.4).

Importantly, the impact of short-term combined aerobic and resistance exercise on the physiology of the study's patients was clearly demonstrated in improvements of measurements of cardiorespiratory fitness and body composition (Table 7.3). Specifically, during the EX intervention period patients had a significant reduction in total body fat percentage, while lean tissue mass was significantly increased. Significant improvements in patient heart rate and maximal oxygen consumption were additionally observed after the exercise period. This highlights, as previously outlined in our study of the influence of exercise on the microbiome of healthy but otherwise inactive individuals, that the benefits of structured exercise are not dependent on dramatic alteration of the gut microbiome.

7.7 | Conclusion

This prospective examination supports our previous findings that that short-to-medium-term combined physical exercise in physically inactive adults does not contribute to substantial alterations in either the diversity or functional potential of microbial gut populations. Furthermore, the study presents evidence in support of the utilisation of structured exercise in IBD patients as an adjunct therapy that will not negatively influence the gut microbiome.

7.8 | References

[1] O. Cronin, W. Barton, P. Skuse, N. C. Penney, I. Garcia-Perez, E. F. Murphy, T. Woods, H. Nugent, A. Fanning, S. Melgar, E. C. Falvey, E. Holmes, P. D. Cotter, O. O'Sullivan, M. G. Molloy, F. Shanahan, J. A. Gilbert, A Prospective Metagenomic and Metabolomic Analysis of the Impact of Exercise and/or Whey Protein Supplementation on the Gut Microbiome of Sedentary Adults, mSystems 3 (3) (2018) e00044-18.

[2] C. L. Craig, A. L. Marshall, M. Sjostrom, A. E. Bauman, M. L. Booth, B. E. Ainsworth, M. Pratt, U. Ekelund, A. Yngve, J. F. Sallis, P. Oja, International physical activity questionnaire: 12-country reliability and validity, Med Sci Sports Exerc 35 (8) (2003) 1381-95.

[3] L. S. Pescatello, ACSM's guidelines for exercise testing and prescripton, Wolters Kluwer/Lippincott Williams and Wilkins (Health)

[4] R. F. Harvey, J. M. Bradshaw, A Simple Index of Crohn's-Disease Activity, The Lancet 315 (8167) (1980) 514.

[5] R. S. Walmsley, R. C. S. Ayres, R. E. Pounder, R. N. Allan, A simple clinical colitis activity index, Gut 43 (1) (1998) 29-32.

[6] J. Powell-Tuck, R. L. Bown, J. E. Lennard-Jones, A Comparison of Oral Prednisolone Given as Single or Multiple Daily Doses for Active Proctocolitis, Scandinavian Journal of Gastroenterology 13 (7) (2010) 833-837.

[7] C. Anderson, S. Laubscher, R. Burns, Validation of the Short Form 36 (SF-36) Health Survey Questionnaire Among Stroke Patients, Stroke 27 (10) (1996) 1812-1816.

[8] I. Bjelland, A. A. Dahl, T. T. Haug, D. Neckelmann, The validity of the Hospital Anxiety and Depression Scale, Journal of Psychosomatic Research 52 (2) (2002) 69-77.

[9] A. T. Beck, R. A. Steer, R. Ball, W. Ranieri, Comparison of Beck Depression Inventories -IA and -II in psychiatric outpatients, J Pers Assess 67 (3) (1996) 588-97.

[10] C. D. Spielberger, F. Gonzalez-Reigosa, A. Martinez-Urrutia, L. F. Natalicio, D. S. Natalicio, The state-trait anxiety inventory, Revista Interamericana de Psicologia/Interamerican Journal of Psychology 5 (3 & 4) (2017)

[11] S. Abubucker, N. Segata, J. Goll, A. M. Schubert, J. Izard, B. L. Cantarel, B. Rodriguez-Mueller, J. Zucker, M. Thiagarajan, B. Henrissat, O. White, S. T. Kelley, B. Methe, P. D. Schloss, D. Gevers, M. Mitreva, C. Huttenhower, Metabolic reconstruction for metagenomic data and its application to the human microbiome, PLoS Comput Biol 8 (6) (2012) e1002358.

[12] P. Menzel, K. L. Ng, A. Krogh, Fast and sensitive taxonomic classification for metagenomics with Kaiju, Nat Commun 7 (2016) 11257.

[13] Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: a practical and powerful approach to multiple testing, Journal of the royal statistical society. Series B (Methodological) (1995) 289-300.

[14] J. Oksanen, F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. O'Hara, G. L. Simpson, P. Solymos, M. Henry, H. Stevens, E. Szoecs, H. Wagner, Vegan: Community Ecology Package. R package version 2.4-3,

[15] S. Mandal, W. Van Treuren, R. A. White, M. Eggesbo, R. Knight, S. D. Peddada, Analysis of composition of microbiomes: a novel method for studying microbial composition, Microb Ecol Health Dis 26 (2015) 27663.

[16] I. Garcia-Perez, J. M. Posma, R. Gibson, E. S. Chambers, T. H. Hansen, H. Vestergaard, T. Hansen, M. Beckmann, O. Pedersen, P. Elliott, J. Stamler, J. K. Nicholson, J. Draper, J. C. Mathers, E. Holmes, G. Frost, Objective assessment of dietary patterns by use of metabolic phenotyping: a randomised, controlled, crossover trial, Lancet Diabetes Endocrinol 5 (3) (2017) 184-195.

[17] G. Borg, Borg's perceived exertion and pain scales, Human kinetics

[18] W. Barton, N. C. Penney, O. Cronin, I. Garcia-Perez, M. G. Molloy, E. Holmes, F. Shanahan, P. D. Cotter, O. O'Sullivan, The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level, Gut (2017)

[19] S. F. Clarke, E. F. Murphy, O. O'Sullivan, A. J. Lucey, M. Humphreys, A. Hogan, P. Hayes, M. O'Reilly, I. B. Jeffery, R. Wood-Martin, D. M. Kerins, E. Quigley, R. P. Ross, P. W. O'Toole, M. G. Molloy, E. Falvey, F. Shanahan, P. D. Cotter, Exercise and associated dietary extremes impact on gut microbial diversity, Gut 63 (12) (2014) 1913-20.

[20] V. Demonfort Nkamga, B. Henrissat, M. Drancourt, Archaea: Essential inhabitants of the human digestive microbiota., Human Microbiome Journal. 3 (2017) 1-8.

[21] R. Lucas Lopez, M. J. Grande Burgos, A. Galvez, R. Perez Pulido, The human gastrointestinal tract and oral microbiota in inflammatory bowel disease: a state of the science review, APMIS 125 (1) (2017) 3-10.

7.9 | Supplementary Content

Inclusion Criteria
• 18– 40 year-old male and females • BMI between 22 and 35 • Confirmed diagnosis of Crohn's Disease or Ulcerative Colitis • Low physical activity levels – as defined by the International Physical Activity Questionnaire (IPAQ) (i.e. currently inactive) • No regular organized physical activity in the last month • Subject is willing to comply thoroughly with the basic, standard requirements of the training program • Subject is willing to refrain from taking any unspecified dietary, probiotics, multivitamin or herbal supplements (14-day washout period pre-commencement of exercise program) • Willingness to confine formal exercise to within the Mardyke Arena • Deemed fit f participate in exercise program for 8 weeks by the study investigators

Supplementary Table 7.1 | Inclusion criteria

Exclusion criteria:
• History of coronary artery disease, congenital heart disease or any cardiovascular disease, including a family history before 45 years of age • Uncontrolled hypertension (>140/90mmHg) (not white coat) • Course of antibiotics or bowel preparation for endoscopy one month prior to the commencement of the exercise program • History of psychiatric disorders or substance abuse • Any additional medical condition deemed exclusionary by the Principle Investigator • History of total colectomy • Currently treated with high dose corticosteroids or disease flare within 2 weeks of starting exercise program • Pregnant • History of faecal incontinence • Have a malignant disease or any concomitant end stage organ disease • Received any experimental drug or have been involved in an experimental trial in the last 30 days

Supplementary Table 7.2 | Exclusion criteria

Chapter 8

LONGITUDINAL EFFECTS OF PHYSICAL ACTIVITY ON THE GUT MICROBIOME

8.1 | Abstract

Previous investigations of the human gut microbiome have illustrated that physical exercise exerts influence on the taxonomic composition and metabolic potential of the microbiome in high-performance athletes. However, subsequent study of the microbiome in individuals unaccustomed to exercise, following an intervention period of structured moderate-intensity exercise, revealed modest alterations to the intestinal microbiome. These results raised the possibility that the absence of major alterations of the gut microbiome after the intervention was due to the limited exposure to the exercise stimulus. This theory is assessed here through an *N of 1* study, in which the gut microbiome of 2 participants, not regularly involved in structured physical activity beyond that of daily life, are observed over the course of 6 months in which an exercise routine of moderate-intensity has been initiated. Results indicate that diversity of the gut microbiome corresponds to training events and health challenges in the participants.

In addition to the author, the original material subject of this chapter was contributed to accordingly:

Clinical work: Dr. Owen Cronin

Metabolomic processing and data analysis: Dr. Nicholas C. Penney, Dr. Isabel Garcia Perez, & Professor Elaine Holmes

Presented material has not been published at the time of this thesis' submission.

8.2 | Introduction

As described in chapters 6 and 7, moderate structured physical activity has been shown to exert modest influence on the gut microbiome of individuals unaccustomed to exercise, both in respect to taxonomic composition and metabolic production [1]. Conversely, it has been shown both in chapter 3 and in separate studies that the gut microbiome of individuals engaged in regular long-term exercise shows divergence in taxonomic composition and functional potential when compared to the gut microbiome of comparatively less-active individuals [2-4]. Previous work has addressed the possibility that the absence of major alterations of the gut microbiome in participants undergoing short-term exercise interventions arises from the limited exposure to the exercise stimulus. The following chapter assesses this theory through description of a study in which the gut microbiome of 2 participants that are not regularly involved in structured physical activity beyond that of daily life are longitudinally assessed over the course of 6 months in which an exercise routine of moderate-intensity has been initiated This study marks the first such access to the interaction of physical exercise and the gut microbiome over a prolonged duration of time.

8.3 | Study objective

The primary objective of this chapter was to assess the influence of structured physical exercise in physically inactive but otherwise healthy adults over the course of 6 months of rigorous physical training. Specifically, the chapter aims to determine if the lack of alterations in the gut microbiome of participants in previous studies was due to the short durations of exercise interventions. Additionally, this study illustrates the usefulness of *N of 1* style study designs.

8.4 | Methods

8.4.1 | Experimental model and ethical approval

An overview of the study design is including in the Results section as Figure 8.1. Study approval was granted by the Cork Clinical Research Ethics Committee before commencing.

8.4.2 | Study recruitment and safe participation

Two male participants were recruited from the University College Cork via email. Both participants had illustrated interest in previous related studies and elected to be considered for inclusion in future studies. Prior to enrolment, both participants were screened by inclusion and exclusion criteria (described in Table 8.1).

At the discretion of the investigators, participants could have been removed from the study should the exercise intervention pose health risk or a health condition vulnerable to exercise arose. Baseline levels of physical activity were assessed using the International Physical Activity Questionnaire short form [5]. Safe participation in the study was ensured by medical screening of both participants using an adapted version of the safe participation questionnaire of the American College of Sports Medicine [6].

8.4.3 | Exercise intervention

Self-directed training was performed by both participants with the goal of increasing physical fitness to levels necessary for participation in an endurance-based sports competition. Specifically, the two participants were separately training for either a marathon or an Olympic distance triathlon. Training for the marathon consisted of regular aerobic exercise complimented with twice weekly resistance training while endurance aerobic training was exclusively the basis of preparation for the triathlon. Throughout the study's course, both participants tracked their respective physical activity with wearable

activity monitors (ActiGraph wGT3X-BT, ActiGraph, FL, U.S.). These activity monitors were worn continuously on the non-dominant wrist and were removed only for cleaning, during bathing, and downloading/recharging of the monitor.

8.4.4 | Measurement visits

The participants attended Cork University Hospital at two week intervals for the entirety of the study. For the first week of the epoch, participants were asked to maintain usual physical activity habits to obtain baseline levels of activity following which regular training was commenced. At each two week visit participant anthropometric measurements were taken (e.g. BMI and waist:hip ratio), in addition to fasting blood, fresh faecal and urine samples, and activity monitor data. Every other visit (i.e. monthly) additionally required completion of a food frequency questionnaire, the recording of heart rate variability using the HRV4training smartphone application [7], and a substrate utilization during exercise assessment at the exercise physiology laboratory in the University College Cork Mardyke Arena gymnasium. At the study's commencement, mid-point, and conclusion (0, 3, and 6 months respectively), body composition profile using total body DEXA scan, and cardiorespiratory fitness testing VO_{2max} were performed as described in sections 2.3.2 and 2.3.3, respectively. Blood samples were collected for analysis according to procedures presented in previous chapters and section 2.3.4. Volunteers were asked to provide fresh urine and faecal samples, which were transported at room temperature to Teagasc Food Research Centre, Moorepark, Fermoy, Co. Cork, where DNA extraction took place.

8.4.5 | VO_{2max} and body composition measurement

To establish baseline cardiorespiratory fitness of the participants, an assessment of maximal aerobic capacity was employed with an incremental protocol developed by the exercise physiology laboratory. Experience exercise physiologists supervised and conducted the

protocol in which each participant exercised to exhaustion. The test was not stopped until both an RPE of 20/20 and a continuous plateau in VO_{2max} for 60 seconds were achieved.

A Lunar iDXA machine (GE Healthcare, Madison, WI) at the Bone Densitometry Unit, Cork University Hospital, was used to measure body composition with the enCORE software (V.13.4, 2010). Quality control (QC) analysis was performed on the iDXA machine before use of the machine on each measurement day.

8.4.6 | Dietary data collection

Dietary data were collected by means of a 146-item food frequency questionnaire (FFQ) as outlined previously and in section 2.3.1 [8].

8.4.7 | DNA extraction and metagenomic sequencing of faecal

DNA was extracted from the donated fresh faecal samples received at the Teagasc Moorepark research facility following the methods presented in section 2.5.1. Libraries for metagenomic sequencing were prepared according to procedures in section 2.5.2, and were sequenced on an Illumina NextSeq 500 (chemistry V.2.0) sequencing platform (Teagasc sequencing facility). High-throughput sequencing was performed using the high-output 500/550 reagent kit.

8.4.8 | Bioinformatic processing of microbial metagenomic sequencing

Processing of metagenomic FASTQ sequence files adhered to approaches described in section 2.5.2, with high-quality processed reads used to generate functional profiles with the Human Microbiome Project (HMP) Unified Metabolic Analysis Network (HUMAnN2 V.0.99) pipeline [9]. Taxonomic profiling was facilitated with the Kaiju taxonomic assignment software tool (V.1.5.0) [10].

8.4.9 | Metabolomic sample preparation

The approaches employed in processing faecal and urine samples for metabolomic analysis are described in section 2.7.

8.4.10 | Quantification and statistical analysis

Statistical assessment of dissimilarity matrices (Bray-Curtis) derived from microbial data was facilitated with the adonis2 function in the vegan R package (V.2.4-3) [11]. Measurements of alpha diversity and calculations of relative abundances were also performed with the vegan package. Relative-abundance data were generated separately for identified species within each phylogenetic domain (e.g., *Bacteria*). Analysis of metabolic data proceeded according to methods described in section 2.8.4.

8.4.11 | Data and software availability

The microbial DNA sequences have been deposited in the European Nucleotide Database (ENA) database under ID code PRJEB27624.

8.5 | Results

8.5.1 | Study overview

Responding to study advertisement, 2 male adults aged 30 and 33 years with starting BMI of 28.6 and 31.7kg/m^2 were recruited for participation (Table 8.1). The study was conducted in accordance with the Declaration of Helsinki and was granted ethical approval by the Cork Clinical Research Ethics Committee (CREC) prior to commencement. Participants were screened against specific criteria for inclusion and exclusion, including history of cardiovascular disease and uncontrolled hypertension (Supplementary Table 8.1).

During the study the participants were engaged in self-directed improvements in physical fitness that was aimed towards achieving physical competency to participate in a competitive fitness event (triathlon and marathon competitions, respectively). Marathon training consisted of regular aerobic exercise and twice weekly resistance training while endurance aerobic training was exclusively the basis of preparation for the triathlon. Both participants tracked their respective physical activity with wearable activity monitors throughout the study's duration.

8.5.2 | Improvements in body composition and cardio-respiratory fitness

Both participants demonstrated incremental improvements in measurements of body composition and cardio-respiratory fitness throughout the duration of the study (Figure 8.2). BMI, waist circumference, and resting heart rate decreased for both participants while under observation. Additionally, estimated maximal oxygen consumption (VO_{2max}) increased. Total body fat (%) was also reduced.

Patient characteristics	Values					
	Participant 1 (Marathoner)			Participant 2 (Triathlete)		
	T_0	T_{14}	Δ	T_0	T_{14}	Δ
Age (years)	30	—	—	33	—	—
Height (cm)	181	—	—	182	—	—
Weight (kg)	93.8	89.2	**-4.6**	104.9	103.4	**-1.5**
BMI (kg/m^2)	28.6	27.2	**-1.4**	31.7	31.2	**-0.5**
Waist:Hip ratio	0.92	0.92	**0.0**	0.95	.91	**-0.04**
Body fat (%)	25.6	21.7	**-3.9**	34.7	34.5	**-0.2**
Fat mass (kg)	23.9	19.4	**-4.6**	36.3	35.7	**-0.6**
Fat mass (trunk) (kg)	14.8	11.7	**-3.1**	20.9	20.4	**-0.5**
Lean tissue mass (kg)	65.6	65.9	**0.2**	64.97	64.2	**-0.7**
Estimated VO_{2max} (mls/kg/min)	41.1	46.6	**5.5**	33.6	38	**4.4**
Max HR (bpm)	183	179	**-4**	196	179	**-17**
Resting HR (bpm)	69	50	**-19**	58	72	**-2**
Systolic BP (mmHg)	122	116	**-6**	128	127	**-1**
Diastolic BP (mmHg)	77	75	**-2**	87	72	**-15**
Weekly PA (IPAQ, METS)	891.5	—	—	646.5	—	—
Weekly PA (IPAQ, kCals)	1,393.7	—	—	1,130.3	—	—

Table 8.1 | Baseline measurements of participant demographic and anthropometric characteristics. Measurements presented for initial (**T_0**) and final (**T_{14}**) readings, along with the change between the two (**Δ**).

BMI, body mass index; IPAQ, International Physical Activity Questionnaire; METS, metabolic equivalents; HR, heart rate; bpm, beats per minute; PA; VO_{2max}, estimated maximal oxygen consumption.

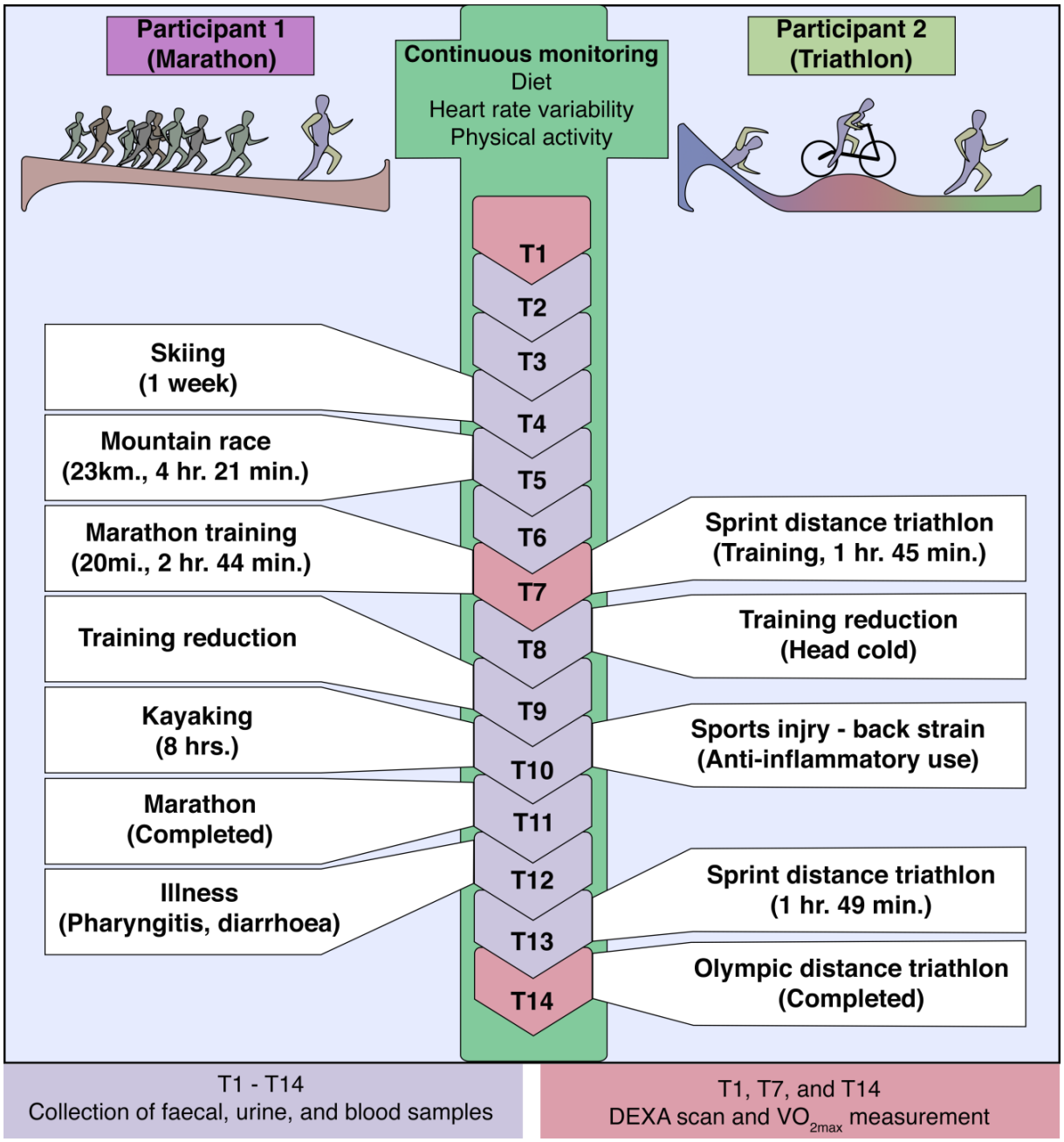


Figure 8.1 | Study overview. The two participants supplied blood, faecal, and urine samples at every 2 week long time point. DEXA scan and measurement of VO_{2max} occurred at the beginning (T0) mid-point (T7) and final time point (T14). Diet, heart rate variability, and physical activity were continuously monitored. VO_{2max}, estimated maximal oxygen consumption.

8.5.3 | Recording of dietary habits

Throughout the study period participants were requested to refrain from taking vitamin, dietary, and herbal supplements, and to maintain usual *ad libitum* dietary intake. While enrolled participants were also asked to record daily dietary habits via a smartphone application. Such dietary recordings were deconstructed into macronutrient values presented as daily averages in grams (Figure 8.3).

8.5.4 | Metagenomic assessment of gut microbiota

Species level taxonomic profiles of the participants' gut microbiota showed fluctuations that corresponded to health and training events. Notably, bacterial α-diversity peaked for the participant training for the marathon near the date of the event (Figure 8.4). Conversely when the same individual fell ill, α-diversity of bacterial species was drastically reduced. For the triathlon trainer, an increase in α-diversity coincided with intense training before reducing as training intensity was lowered. During a period of anti-inflammatory use resulting from a minor training injury, α-diversity reached its peak. For both individuals it was also shown that diversity of metabolic pathway models closely adhered to that of the taxonomic profiles.

Assessment of taxonomic β-diversity with non-metric multidimensional scaling (NMDS) illustrated strong divergence between the two participants (Figure 8.5). Taxonomic profiling of virus species did not report any viruses at the species level of phylogeny (data not presented).

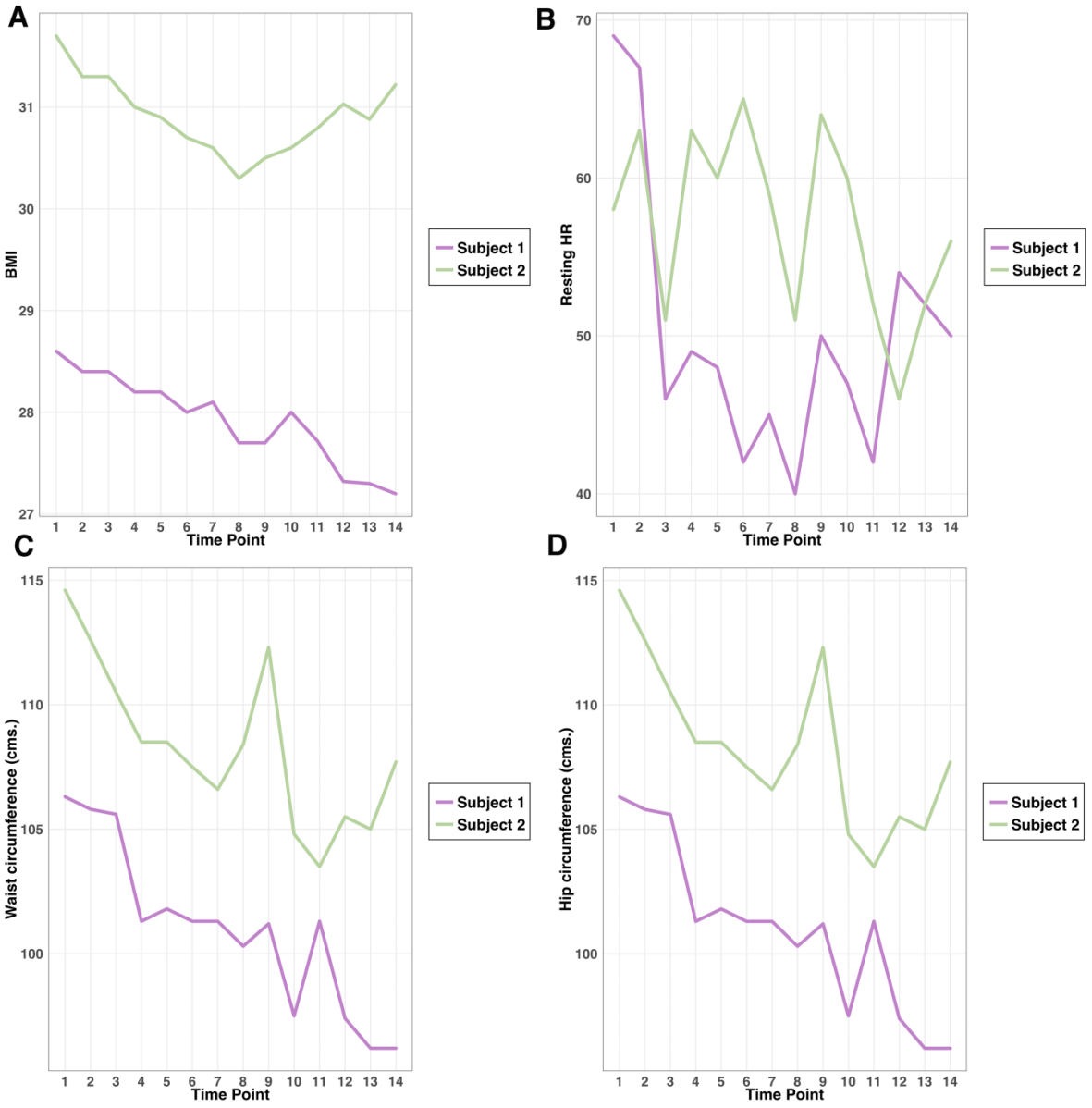


Figure 8.2 | Subject body composition and anthropometric data. Fitness parameters improved during the observation period. (**A**) BMI was shown to have an overall reduction over the study duration. (**B**) Resting heart rate was also demonstrated to lower for both participants. (**C**) Waist circumference was similarly reduced over the duration of the study, (**D**) just as hip circumference was also.

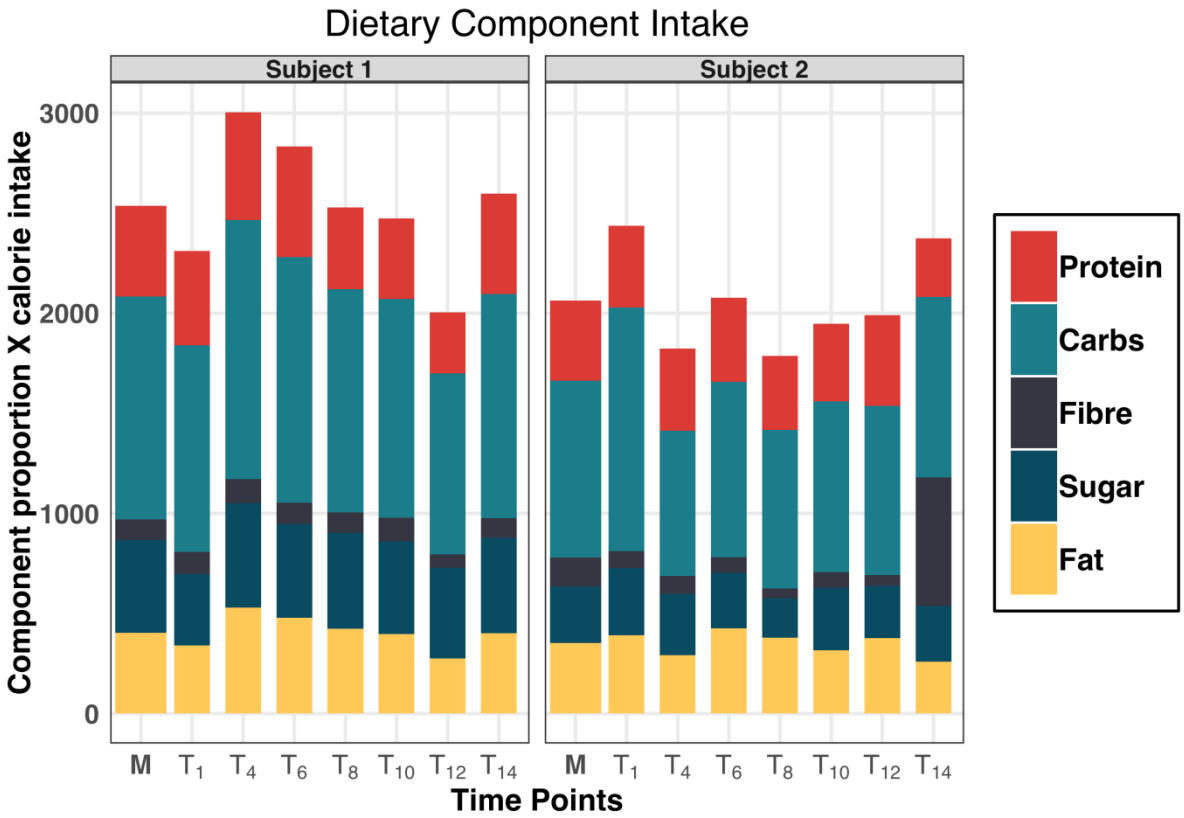


Figure 8.3 | Nutritional constituents of participants over study period. Daily averages of dietary constituents recorded over time intervals as grams per day (g./day), presented as proportions of time interval total intake (total g./interval), and scaled to average caloric intake per interval (calories/interval). Bar sums indicate average daily caloric intake for interval, with participants' averages for calories and dietary components indicated (**M**).

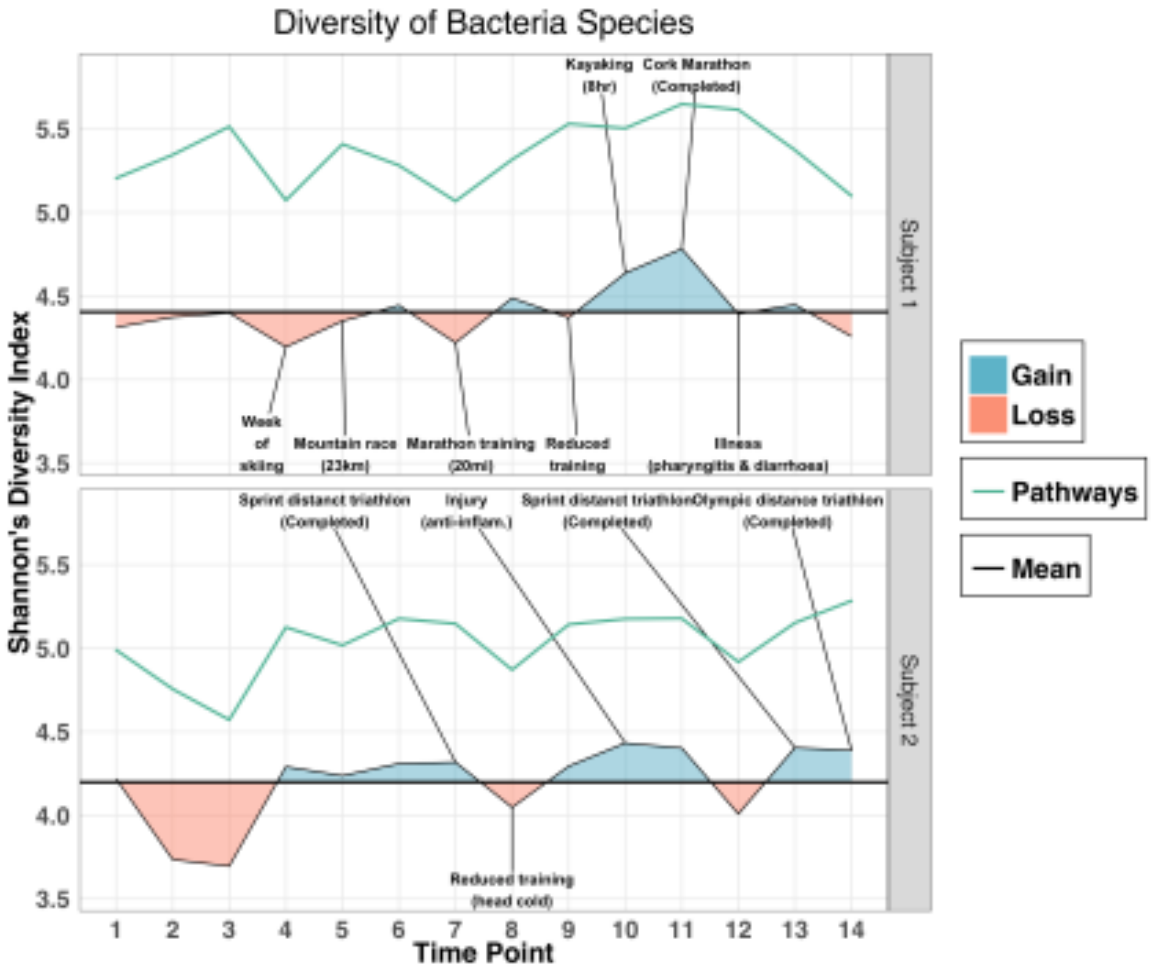


Figure 8.4 | α-diversity of bacterial species and pathways. Shannon α-diversity H-index of Bacteria species and metabolic pathway models for both participants over the duration of the study. Deviations of each participant's mean α-diversity highlight the potential influence of training and health events on the composition of the gut microbiome.

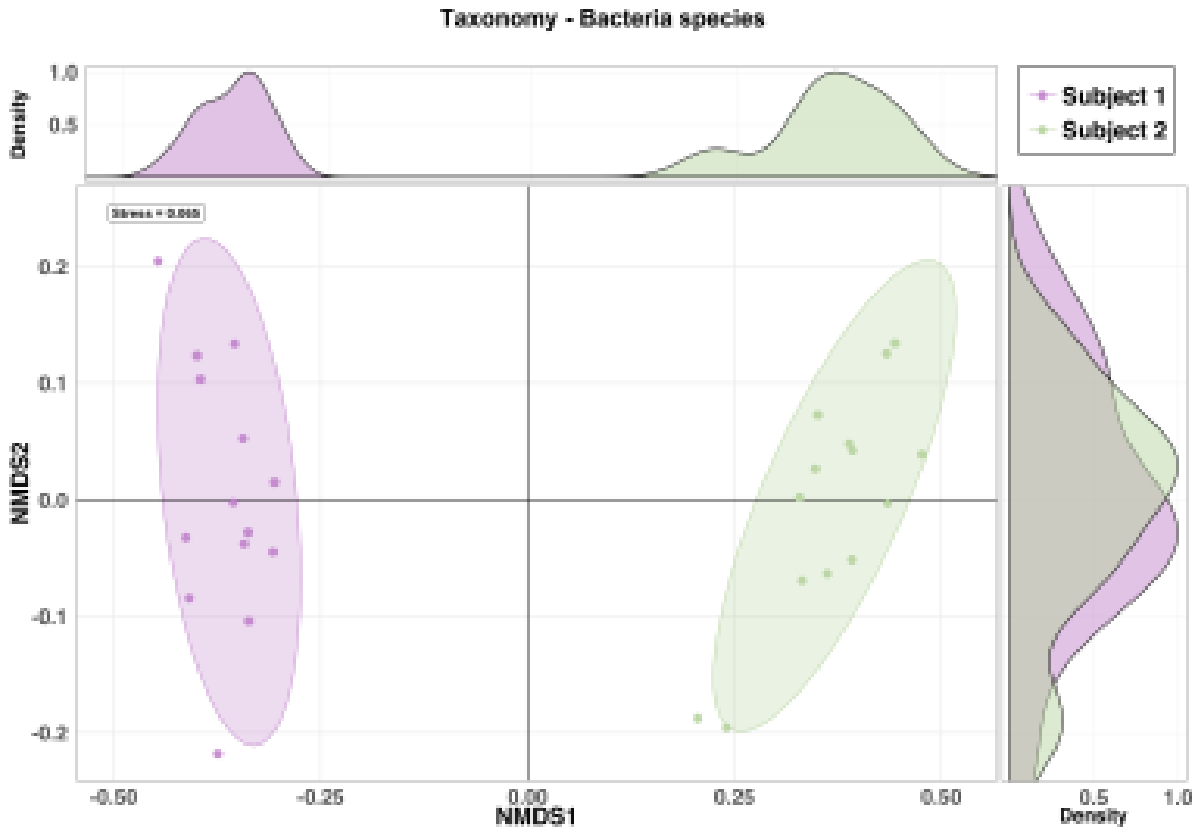


Figure 8.5 | β-diversity of bacterial taxonomy. Non-metric multidimensional scaling (NMDS) of relative abundance profiles for bacterial species demonstrates complete clustering of measurements according to individual. Density plots along the axes of all panels display the concentrations of data points, and were generated from kernel density estimates and scaled to a maximum value of 1.

8.5.5 | Pathway assessment

General (e.g., fucose degradation) and taxonomically specific (e.g., fucose degradation in *Akkermansia muciniphila*) metabolic pathways models generated from metagenomic sequencing were compared for both participants throughout the course of the study. To determine core metabolic functions that persisted within the participants despite both dramatically increasing physical activity levels, the full series of samples for each participant were compared. Statistical assessment of the metabolic pathways illustrated that 943 pathways were significantly varied between the two participants when all time points for each individual were collectively compared to one another. The collection of significantly varied pathways highlights core differences in the functional potential of the two participants (Supplementary Table 8.2). Assessment of β-diversity with NMDS of both the species-specific (Figure 8.6A) and general pathways (Figure 8.6B) illustrated dramatic clustering of pathways within each participant.

8.5.6 | Metabolomic profiling

Tracking alterations in metabolic profiles over the duration of the study with untargeted metabolomic analysis of the full spectrum of urinary and faecal metabolites from the participants demonstrated that the final time point of participant 1 strongly deviated from all previous time points (Figure 8.7A &C). The trajectory of time points for participant 2 did not have such variation, and the initial and final time points were closely related (Figure 8. 7B & D). Linear regression to BMI for urinary ^{1}H-NMR spectra from both participants identified *N*-methyl-nicotinate and hippurate positively associated to BMI, while phenylacetylglutamine was inversely associated to BMI (Figure 8.8).

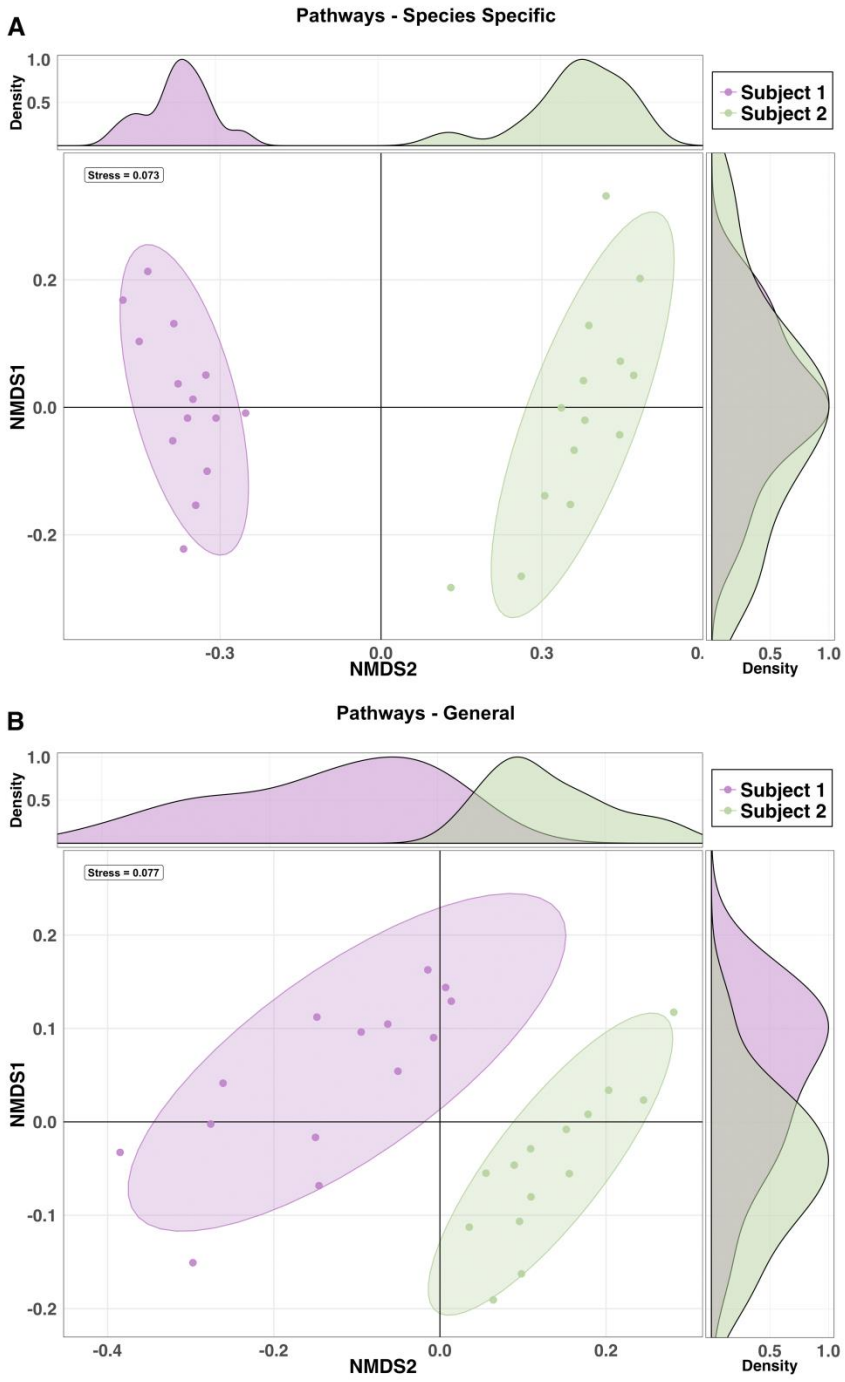


Figure 8.6 | β-diversity of metabolic pathways. Non-metric multidimensional scaling (NMDS) of relative abundance profiles for species-specific (**A**) and general (**B**) models of metabolic pathways demonstrates complete clustering of measurements according to individual. Density plots along the axes of all panels display the concentrations of data points, and were generated from kernel density estimates and scaled to a maximum value of 1.

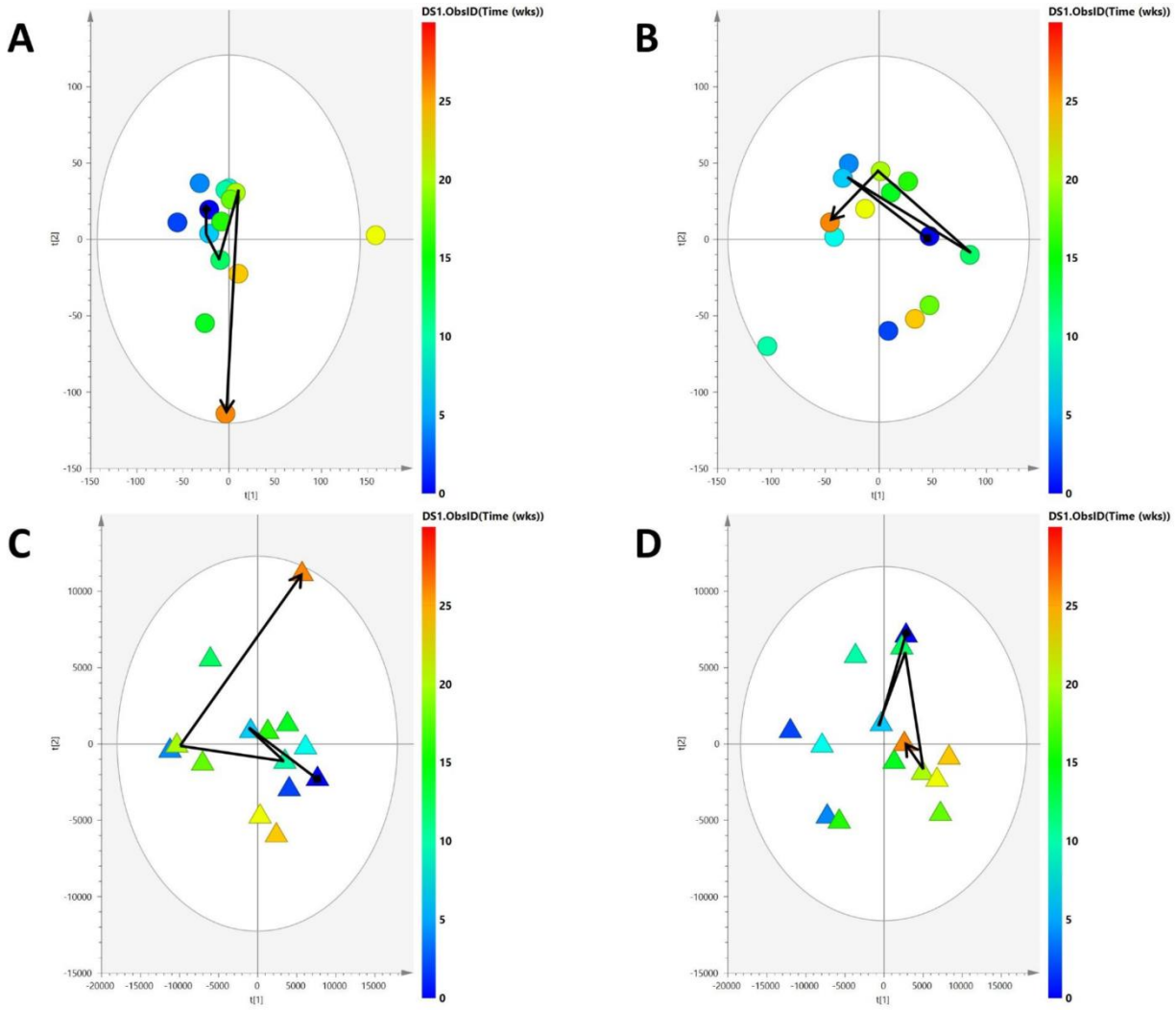
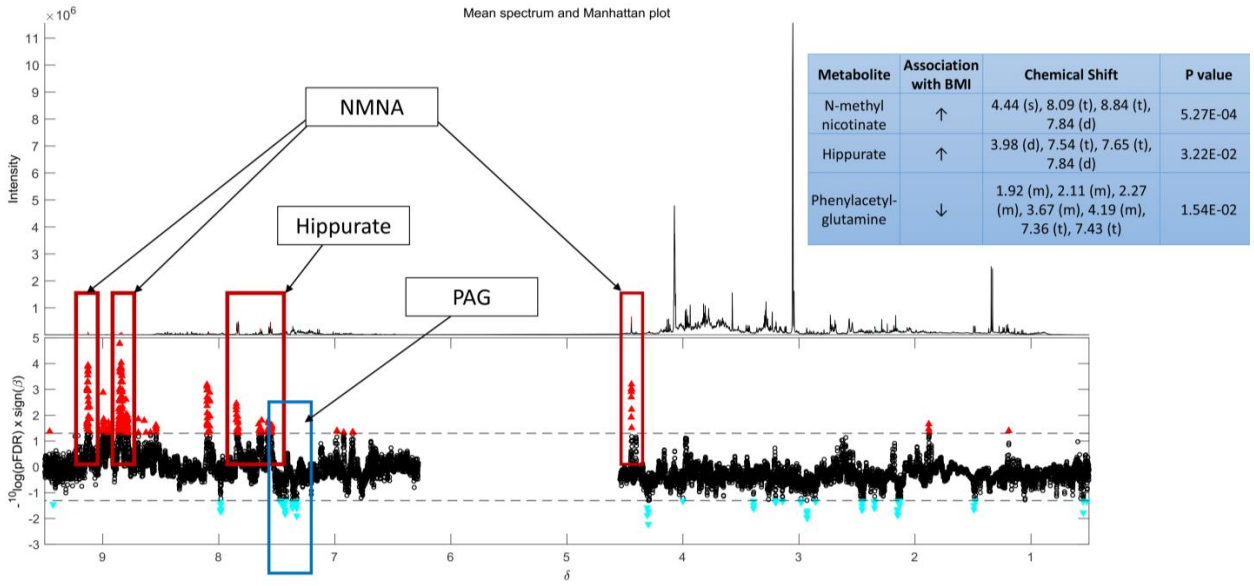


Figure 8.7 | Metabolite trajectories for both participants. PCA of ^{1}H-NMR spectra for both urinary and faecal metabolites. (**A & C**) Subject 1's urine and faecal profiles, respectively. Trajectory of subject 1's urinary profile illustrates a strong separation over time. (**C & D**) The urine and faecal metabolite profiles of subject 2 show less separation over the observation period.



Metabolite	Association with BMI	Chemical Shift	P value
N-methyl nicotinate	↑	4.44 (s), 8.09 (t), 8.84 (t), 7.84 (d)	5.27E-04
Hippurate	↑	3.98 (d), 7.54 (t), 7.65 (t), 7.84 (d)	3.22E-02
Phenylacetyl-glutamine	↓	1.92 (m), 2.11 (m), 2.27 (m), 3.67 (m), 4.19 (m), 7.36 (t), 7.43 (t)	1.54E-02

Figure 8.8 | Linear regression of ^{1}H-NMR urinary spectra to BMI. Significant metabolites (pFDR < 0.05) are highlighted in the Manhattan plot and mean spectrum. Spectra highlighted in red show an increase with bacterial diversity, while spectra highlighted in blue decrease with bacterial diversity. *N*-methyl-nicotinate and hippurate positively associated to BMI, while phenylacetylglutamine was inversely associated to BMI.

8.6 | Discussion

Our prior accomplishments with characterising the gut microbiome within the framework of physical exercise have included profiling the extremes of the paradigm with assessment of the elite athlete gut microbiome [2, 8], and also the examination of structured exercise as a novel stimulus in populations that are unaccustomed to rigorous physical activity [1]. Although profiling of the athlete gut microbiome presented numerous deviations in taxonomic composition, functional potential, and metabolic activity from a control population of healthy but sedentary adults, the microbiome alterations resulting from a short-term exercise intervention were more nuanced. We have theorised that the unique structure of the athlete microbiome may be the result of adaptations to lifelong engagement in rigorous physical activity and associated lifestyle (e.g. optimised diet) [12]. Additionally, the gut microbiome of adults has been shown to be resistant to dramatic alteration, and may not rapidly adapt to the systemic influences of exercise [13]. Here, we aimed to assess if indeed the gut microbiome is observably altered by rigorous exercise over a prolonged period of time. By tracking two participants over the course of a 6 month period, modification of the gut microbiome in response to well documented training events, dietary trends, and health events was achieved. Throughout the duration of the study, both participants were actively engaged in self-directed exercise. It was shown that this endeavour was successful for both individuals with regard to improvements in cardiorespiratory health and body composition (Table 8.1 and Figure 8.2).

This observational study demonstrated that diversity of bacterial species corresponded to the intensity of physical exertion, while also being influenced by injury and illness (Figure 8.4). Measurements of α-diversity for the participant training toward a marathon peaked at the sample collection point nearest to the date of the marathon that was

competed in, before reaching the lowest point when the individual became ill post-marathon. For the individual preparing for the Olympic triathlon, a period of transition between intense training and a reduction in training presented lowered diversity. Intriguingly, the level of diversity peaked with the occurrence of an injury and the subsequent use of anti-inflammatory medication. It has previously been shown that a wide range of pharmaceutical compounds can interact with microbes from the intestinal microbiome [14], which may provide a partial explanation of this dramatic increase in diversity. Metabolomic analysis identified trajectories of metabolite profiles for the two participants that were dissimilar, which may highlight the differences in type of exercise pursued by the two volunteers. Intriguingly, significant associations between urinary metabolites and BMI were detected. One such metabolite is phenylacetylglutamine (PAG), which has previously been associated to lean body composition [15]. This association is also highlighted here, as levels of urinary PAG decrease with BMI.

Assessment of β-diversity for both taxonomic and metabolic pathway profiles demonstrated that the participants maintained unique microbiome structures throughout the observation period (Figures 8.5 and 8.6). This intriguingly suggests that any alterations of the microbiome in response to exercise do not result in a specific microbiome structure. If this is indeed the case, individuals may retain their core microbiome when attempting to modify it with exercise, possibly with only certain aspects of its structure and function adapting to the exercise. Due to the observational nature of the study, advanced statistical analysis of the results was not possible, and subsequently the presented outcomes will require further studying to draw firmer conclusions.

8.7 | References

[1] O. Cronin, W. Barton, P. Skuse, N. C. Penney, I. Garcia-Perez, E. F. Murphy, T. Woods, H. Nugent, A. Fanning, S. Melgar, E. C. Falvey, E. Holmes, P. D. Cotter, O. O'Sullivan, M. G. Molloy, F. Shanahan, J. A. Gilbert, A Prospective Metagenomic and Metabolomic Analysis of the Impact of Exercise and/or Whey Protein Supplementation on the Gut Microbiome of Sedentary Adults, mSystems 3 (3) (2018) e00044-18.

[2] W. Barton, N. C. Penney, O. Cronin, I. Garcia-Perez, M. G. Molloy, E. Holmes, F. Shanahan, P. D. Cotter, O. O'Sullivan, The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level, Gut (2017)

[3] L. M. Petersen, E. J. Bautista, H. Nguyen, B. M. Hanson, L. Chen, S. H. Lek, E. Sodergren, G. M. Weinstock, Community characteristics of the gut microbiomes of competitive cyclists, Microbiome 5 (1) (2017) 98.

[4] X. Zhao, Z. Zhang, B. Hu, W. Huang, C. Yuan, L. Zou, Response of Gut Microbiota to Metabolite Changes Induced by Endurance Exercise, Frontiers in Microbiology 9 (2018)

[5] C. L. Craig, A. L. Marshall, M. Sjostrom, A. E. Bauman, M. L. Booth, B. E. Ainsworth, M. Pratt, U. Ekelund, A. Yngve, J. F. Sallis, P. Oja, International physical activity questionnaire: 12-country reliability and validity, Med Sci Sports Exerc 35 (8) (2003) 1381-95.

[6] L. S. Pescatello, ACSM's guidelines for exercise testing and prescripton, Wolters Kluwer/Lippincott Williams and Wilkins (Health)

[7] M. Altini, O. Amft, HRV4Training: Large-scale longitudinal training load analysis in unconstrained free-living settings using a smartphone application, Conf Proc IEEE Eng Med Biol Soc 2016 (2016) 2610-2613.

[8] S. F. Clarke, E. F. Murphy, O. O'Sullivan, A. J. Lucey, M. Humphreys, A. Hogan, P. Hayes, M. O'Reilly, I. B. Jeffery, R. Wood-Martin, D. M. Kerins, E. Quigley, R. P. Ross, P. W. O'Toole, M. G. Molloy, E. Falvey, F. Shanahan, P. D. Cotter, Exercise and associated dietary extremes impact on gut microbial diversity, Gut 63 (12) (2014) 1913-20.

[9] S. Abubucker, N. Segata, J. Goll, A. M. Schubert, J. Izard, B. L. Cantarel, B. Rodriguez-Mueller, J. Zucker, M. Thiagarajan, B. Henrissat, O. White, S. T. Kelley, B. Methe, P. D. Schloss, D. Gevers, M. Mitreva, C. Huttenhower, Metabolic reconstruction for metagenomic data and its application to the human microbiome, PLoS Comput Biol 8 (6) (2012) e1002358.

[10] P. Menzel, K. L. Ng, A. Krogh, Fast and sensitive taxonomic classification for metagenomics with Kaiju, Nat Commun 7 (2016) 11257.

[11] J. Oksanen, F. G. Blanchet, M. Friendly, R. Kindt, P. Legendre, D. McGlinn, P. R. Minchin, R. O'Hara, G. L. Simpson, P. Solymos, M. Henry, H. Stevens, E. Szoecs, H. Wagner, Vegan: Community Ecology Package. R package version 2.4-3,

[12] O. Cronin, O. O'Sullivan, W. Barton, P. D. Cotter, M. G. Molloy, F. Shanahan, Gut microbiota: implications for sports and exercise medicine, Br J Sports Med 51 (9) (2017) 700-701.

[13] J. J. Faith, J. L. Guruge, M. Charbonneau, S. Subramanian, H. Seedorf, A. L. Goodman, J. C. Clemente, R. Knight, A. C. Heath, R. L. Leibel, M. Rosenbaum, J. I. Gordon, The long-term stability of the human gut microbiota, Science 341 (6141) (2013) 1237439.

[14] L. Maier, M. Pruteanu, M. Kuhn, G. Zeller, A. Telzerow, E. E. Anderson, A. R. Brochado, K. C. Fernandez, H. Dose, H. Mori, K. R. Patil, P. Bork, A. Typas, Extensive impact of non-antibiotic drugs on human gut bacteria, Nature 555 (7698) (2018) 623-628.

[15] E. Holmes, J. V. Li, T. Athanasiou, H. Ashrafian, J. K. Nicholson, Understanding the role of gut microbiome-host metabolic signal disruption in health and disease, Trends Microbiol 19 (7) (2011) 349-59.

8.8 | Supplementary Content

Inclusion Criteria
• 18– 40 year-old (male) • Able to provide informed consent • Good overall medical health • Willing to increase physical activity in a structured manner and to be followed up and measured over a 6-month period • Considered low risk for cardiac events during exercise as guided by the AHA/ACSM's pre-participation questionnaire
Exclusion Criteria
• Any acute or chronic medical condition including chronic musculo-skeletal pain • Personal history of cardiovascular disease of any kind • Family history of cardiovascular disease before the age of 50 • History of total colectomy or significant abdominal surgery • Antibiotic in the last 6 weeks (relative exclusion criteria) • Any medical condition deemed exclusionary by the investigators • A history of substance abuse • Have been on an experimental medication in the last 30 days • On any regular medication

Supplementary Table 8.1 | Inclusion and Exclusion criteria.

See Appendix C for supplementary table 8.2

Chapter 9

SUMMARY AND CONCLUSION

9.1 | Summary of original research

The content presented in this thesis demonstrates substantial advancements made towards its primary aim to examine the capacity for influence of physical exercise on the human intestinal microbiome. Previous chapters have described observations stemming from a range of investigative frameworks that cross section host-microbe interactions within the context of exercise. Fortuitous access to unique populations ranging from professional athletes to IBD patients have enabled the opportunity to gain insights on numerous different aspects of physical activity's impact on the human gut microbiome. The remainder of this current chapter outlines final considerations of the distinct sections of this thesis, before describing implications for the field of microbiome research, and finally offering suggestions of pursuits to carry the *diet and exercise-microbiome paradigm* into the future.

9.1.1 | The athlete microbiome

Detailed in chapter 3, the gut microbiome of elite athletes has been described as presenting a unique functional capacity. Building upon previous analysis of the same cohort, it was shown that not only do the athletes have a more diverse gut microbiome, but the metabolic capacity is primed for enhanced energy usage (section 3.5.1, figure 3.3) [1-3]. Additionally an enrichment of faecal concentrations of SCFAs, compounds with widely purported health benefits, within this group suggests that there may be positive influence on health (section 3.5.2, figure 3.4) [4-6]. As stated previously, while elite athletes may be fit, their condition is not necessarily more healthy than non-athletic peers [7]. In addition to an increase of

beneficial metabolites, some, such as TMAO, are considered to be detrimental (section 3.5.2, supplementary figure 3.2) [8-12]. Furthermore, sensible speculation would be difficult to make on what the health impact would be of maintaining a gut microbiome with such irregular dynamics of energy expenditure and production without the nutritional resources provided by the dietary habits of high-performance athletes.

9.1.2 | Computational interrogations of the athlete microbiome

Two computational modelling approaches, presented separately in chapters 4 and 5, were used to further describe aspects of the athlete gut microbiome. In chapter 4 the original data from athletes and associated low and high BMI controls was applied to a novel mathematical modelling strategy that reclassified the participants based upon metabolomic profiles and ultimately predicted healthiness of diet (section 4.4.1, figure 4.1). Within these new participant classifications, no robust differences were detected in terms of the gut microbiome (section 4.4.4, and figure 4.3).

Additionally in chapter 5, a Flux Balance Analysis (FBA) based approach was used to elaborate on the metabolic function of gut microbiota while simultaneously validating the method for novel result discovery. With this approach it was shown that of the microbial models present within the original taxonomic profiles, retention of diversity dynamics is achieved (i.e. the modelling accurately represents the results of metagenomic sequencing, see section 5.5.1 and figure 5.1). Furthermore, metabolic reactions resulting from the approach show statistically significant variation between athletes and controls. Intriguingly, potential insight on the original result of elevated TMAO in the athletes was gained from the approach and may justify more elaborate experimentation (section 5.5.3, and supplementary table 5.1).

9.1.3 | Influence of short-term exercise and whey protein supplementation on the gut microbiome

An initial finding of the investigation into the athlete microbiome was that dietary protein consumption and creatine kinase, a biological product of muscle damage and subsequent proxy for fitness, correlated with microbial diversity. To assess if these factors do indeed influence the gut microbiome we took adults that were healthy but physically inactive and had them undergo a short-term exercise regime, increased protein consumption in the form of whey protein, or a combination of both [13].

Overall the identified effects of exercise on the gut microbiome were subtle. However, we observed that participants consuming the whey protein supplement experienced a significant alteration in the β-diversity of the gut virome (refer to section 6.4.3, figures 6.3 and 6.4, and supplementary table 6.6). We went on to profile the supplement itself with metagenomic sequencing, and determined that it was likely that a direct transmission of viral particles from the whey powder to the participants had occurred. To our knowledge, this marked the first presentation of such a finding. While this is an intriguing finding, it is important to note that examination of the virus component of the microbiome was not an explicit aim of the study. Accordingly, the investigative procedures used to analyse the virome were suboptimal, and therefor before strong conclusions are to be drawn on the matter a more tailored approach must be used to assess the virome within the study's parameters. For example, of the 523 million reads that were taxonomically assigned at species level, only 0.035% matched viruses. With such a low proportion of reads available for taxonomic assessment, it is well within reason that the virome of the participants in question was only partially characterised. To address this limitation, methods for the targeted extraction and bioinformatic analysis of virus genetic material should be used [14,

15]. Despite bearing this limitation, the conclusions drawn from the study were that the potential effects of physical exercise on the microbiome likely require a longer duration to become pronounced, and whey protein powder may directly alter dynamics of the gut microbiome.

9.1.4 | The vulnerable IBD gut microbiome and exercise

As demonstrated in our assessment of exercise and whey protein as a novel stimulus to the gut microbiome, short-term physical activity has minimal influence on the function and composition of intestinal microbiota [13]. With these findings it was apparent that the limited effect of short-term exercise on the microbiome may be desirable in populations with reduced GI health. IBD patients were subsequently recruited as a cohort to assess if an exercise intervention could be utilized without detrimental perturbations to the gut microbiome while still transferring the broad systemic benefits of exercise. As anticipated, subtle changes were presented in the gut microbiome after IBD patients underwent the structured exercise intervention (sections 7.5.3 - 7.5.4 and figures 7.2 – 7.3). Curiously, one of the few statistically significant changes identified in the microbiome profiling was an increase in bacterial diversity patients after a control period without exercise. This highlights the unexpected challenges resulting from utilising humans as an experimental model, and potentially resulted from subconscious behavioural changes in anticipation of starting exercise. Despite this, the enrolled patients undergoing exercise treatment succeeded in improving their body composition and cardiorespiratory fitness (section 7.5.2 and table 7.3). This suggests that in specific scenarios an adjunct therapy of short-term exercise can be recommended without concern of dramatic influences on the gut microbiome.

9.1.5 | Influence of prolonged exercise on the intestinal microbiome

In order to address the possibility that previous efforts to affect the gut microbiome with exercise were unyielding of dramatic alterations due to the limited duration of the exercise period, 2 individuals were examined over the course of 6 months while undergoing self-directed exercise. This observational study provided an opportunity to focus resources on 2 participants, subsequently granting a robust picture of microbiome fluctuations in response to training events and health challenges. While the conclusions to be drawn are limited due to the small *n* number that prohibits comprehensive statistical analysis, trends in gut microbiome dynamics, suggest that alterations of the microbiome were influenced by the participants' physical training (section 8.5.4 and figure 8.4).

The study was also designed to accommodate health disruptions for the participants (i.e. to proceed with sample and data collection), and accordingly captured dramatic loss of microbiome diversity when one individual fell ill. Curiously, at another point, following a minor training injury and the use of anti-inflammatory drugs, a substantial increase in diversity was observed. These examples highlight the usefulness of employing similar study designs, as the inevitable variability of human life can be better tracked and the influences of unexpected events on the research focus can be better understood.

9.2 | Implications for related research

Study of the human gut microbiome has led to recognition of its profound influence on human health. With broad health implications spanning from diabetes and inflammatory bowel disease to psychiatric disorder and some forms of cancer, the gut microbiome has garnered increasing interest as a modifiable factor for health and disease [16-24]. Accordingly, there is a mounting call for means to exert controlled alteration of the microbiome, both in terms of taxonomic composition and ultimately function. Considerable

evidence has been put forth for pharmaceuticals, antibiotics as well as compounds not intended to target microbes, diet, and probiotics as effective modulators of the gut microbiome [25]. The body of information presented in this thesis offers further support for the inclusion of physical exercise as a factor of influence on the microbiome. Based upon the various studies detailed here, there are grounds to continue exploring the microbiomes of different high-performance athletes, as well as continuing to pursue exercise as an intervention in disease populations. Moving forward in this regard, it is worth noting the *post hoc* power analysis performed on the studies presented in this thesis (see section 2.8.1). Sample size estimation and power analysis of genetic sequencing focused microbiome studies is challenging due to the multivariate nature of taxonomic or pathway measurements, and the consideration of sequencing factors (e.g. sequencing depth) [26, 27]. Despite the complexity of appropriate power analysis methods for the microbiome, and challenges to the conventions that guide such analysis, it is important to utilise this statistical approach to rationalise and guide microbiome research [28]. With the power analysis performed on measures of alpha diversity from the various studies described in this thesis, it was shown that largely the studies were adequately powered for the observation of changes in diversity. The notable exceptions, are the *N of 1* observational study described in chapter 8, which by its nature is completely underpowered, and the exercise intervention study in IBD patients presented in chapter 7. For this intervention study, the observed effect size of 0.05 at a significance level of 0.05, had a power of 29% with a sample size of 13. In this particular case there were challenges with participant recruitment and retention, which were only presented after the study had commenced. As a result of how underpowered the study was in regards to the detection of changes in diversity of the gut microbiome, results related to the microbiome must be approached with care. This study may however be used to move

similar research forward if it is treated as a pilot study, with more aggressive recruitment used in the future to achieve sufficient sample sizes.

9.3 | Relevant future efforts

Eluted to previously, an enormous range of aspects related to the interaction of exercise and the gut microbiome remain to be pursued. Our research group is actively engaged in advancing the breadth of microbiomes from diverse athletic populations characterised, with aim to identify microbial variations associated to differences in sports type. A potential desired outcome of this continued endeavour being the isolation of specific microbes that confer advantages to certain aspects of athleticism. Finally, the well-established monumental influence of diet on the microbiome has been a persistent challenge in isolating the effects of physical activity upon the microbiota of the intestine [4, 29-34]. The construction of an intervention study examining a cohort with a set and prescribed diet, while engaging in exercise will be essential in advancing the exercise and diet-microbiome paradigm—influences of physical activity and dietary nutrition on the human gut microbiome.

9.4 | References

[1] S. F. Clarke, E. F. Murphy, O. O'Sullivan, A. J. Lucey, M. Humphreys, A. Hogan, P. Hayes, M. O'Reilly, I. B. Jeffery, R. Wood-Martin, D. M. Kerins, E. Quigley, R. P. Ross, P. W. O'Toole, M. G. Molloy, E. Falvey, F. Shanahan, P. D. Cotter, Exercise and associated dietary extremes impact on gut microbial diversity, Gut 63 (12) (2014) 1913-20.

[2] B. A. Petriz, A. P. Castro, J. A. Almeida, C. P. Gomes, G. R. Fernandes, R. H. Kruger, R. W. Pereira, O. L. Franco, Exercise induction of gut microbiota modifications in obese, non-obese and hypertensive rats, BMC Genomics 15 (1) (2014) 511.

[3] M. Estaki, J. Pither, P. Baumeister, J. P. Little, S. K. Gill, S. Ghosh, Z. Ahmadi-Vand, K. R. Marsden, D. L. Gibson, Cardiorespiratory fitness as a predictor of intestinal microbial diversity and distinct metagenomic functions, Microbiome 4 (1) (2016) 42.

[4] A. Koh, F. De Vadder, P. Kovatcheva-Datchary, F. Backhed, From Dietary Fiber to Host Physiology: Short-Chain Fatty Acids as Key Bacterial Metabolites, Cell 165 (6) (2016) 1332-1345.

[5] V. K. Ridaura, J. J. Faith, F. E. Rey, J. Cheng, A. E. Duncan, A. L. Kau, N. W. Griffin, V. Lombard, B. Henrissat, J. R. Bain, M. J. Muehlbauer, O. Ilkayeva, C. F. Semenkovich, K. Funai, D. K. Hayashi, B. J. Lyle, M. C. Martini, L. K. Ursell, J. C. Clemente, W. Van Treuren, W. A. Walters, R. Knight, C. B. Newgard, A. C. Heath, J. I. Gordon, Gut microbiota from twins discordant for obesity modulate metabolism in mice, Science 341 (6150) (2013) 1241214.

[6] H. M. Hamer, D. M. Jonkers, A. Bast, S. A. Vanhoutvin, M. A. Fischer, A. Kodde, F. J. Troost, K. Venema, R. J. Brummer, Butyrate modulates oxidative stress in the colonic mucosa of healthy humans, Clin Nutr 28 (1) (2009) 88-93.

[7] P. B. Maffetone, P. B. Laursen, Athletes: Fit but Unhealthy?, Sports Med Open 2 (1) (2016) 24.

[8] Z. Wang, E. Klipfell, B. J. Bennett, R. Koeth, B. S. Levison, B. Dugar, A. E. Feldstein, E. B. Britt, X. Fu, Y. M. Chung, Y. Wu, P. Schauer, J. D. Smith, H. Allayee, W. H. Tang, J. A. DiDonato, A. J. Lusis, S. L. Hazen, Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease, Nature 472 (7341) (2011) 57-63.

[9] W. H. W. Tang, S. L. Hazen, Microbiome, trimethylamine N-oxide, and cardiometabolic disease, Translational Research (2016)

[10] W. H. W. Tang, Z. E. Wang, B. S. Levison, R. A. Koeth, E. B. Britt, X. M. Fu, Y. P. Wu, S. L. Hazen, Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk, New England Journal of Medicine 368 (17) (2013) 1575-1584.

[11] R. A. Koeth, Z. Wang, B. S. Levison, J. A. Buffa, E. Org, B. T. Sheehy, E. B. Britt, X. Fu, Y. Wu, L. Li, J. D. Smith, J. A. DiDonato, J. Chen, H. Li, G. D. Wu, J. D. Lewis, M. Warrier, J. M. Brown, R. M. Krauss, W. H. Tang, F. D. Bushman, A. J. Lusis, S. L. Hazen, Intestinal

microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis, Nat Med 19 (5) (2013) 576-85.

[12] B. J. Bennett, T. Q. de Aguiar Vallim, Z. Wang, D. M. Shih, Y. Meng, J. Gregory, H. Allayee, R. Lee, M. Graham, R. Crooke, P. A. Edwards, S. L. Hazen, A. J. Lusis, Trimethylamine-N-oxide, a metabolite associated with atherosclerosis, exhibits complex genetic and dietary regulation, Cell Metab 17 (1) (2013) 49-60.

[13] O. Cronin, W. Barton, P. Skuse, N. C. Penney, I. Garcia-Perez, E. F. Murphy, T. Woods, H. Nugent, A. Fanning, S. Melgar, E. C. Falvey, E. Holmes, P. D. Cotter, O. O'Sullivan, M. G. Molloy, F. Shanahan, J. A. Gilbert, A Prospective Metagenomic and Metabolomic Analysis of the Impact of Exercise and/or Whey Protein Supplementation on the Gut Microbiome of Sedentary Adults, mSystems 3 (3) (2018) e00044-18.

[14] A. N. Shkoporov, F. J. Ryan, L. A. Draper, A. Forde, S. R. Stockdale, K. M. Daly, S. A. McDonnell, J. A. Nolan, T. D. S. Sutton, M. Dalmasso, A. McCann, R. P. Ross, C. Hill, Reproducible protocols for metagenomic analysis of human faecal phageomes, Microbiome 6 (1) (2018) 68.

[15] A. H. A. Elbehery, J. Feichtmayer, D. Singh, C. Griebler, L. Deng, The Human Virome Protein Cluster Database (HVPC): A Human Viral Metagenomic Database for Diversity and Function Annotation, Front Microbiol 9 (2018) 1110.

[16] C. T. Brown, A. G. Davis-Richardson, A. Giongo, K. A. Gano, D. B. Crabb, N. Mukherjee, G. Casella, J. C. Drew, J. Ilonen, M. Knip, H. Hyoty, R. Veijola, T. Simell, O. Simell, J. Neu, C. H. Wasserfall, D. Schatz, M. A. Atkinson, E. W. Triplett, Gut microbiome metagenomics analysis suggests a functional model for the development of autoimmunity for type 1 diabetes, PLoS One 6 (10) (2011) e25792.

[17] A. V. Hartstra, K. E. Bouter, F. Backhed, M. Nieuwdorp, Insights into the role of the microbiome in obesity and type 2 diabetes, Diabetes Care 38 (1) (2015) 159-65.

[18] J. M. Norman, S. A. Handley, M. T. Baldridge, L. Droit, C. Y. Liu, B. C. Keller, A. Kambal, C. L. Monaco, G. Zhao, P. Fleshner, T. S. Stappenbeck, D. P. McGovern, A. Keshavarzian, E. A. Mutlu, J. Sauk, D. Gevers, R. J. Xavier, D. Wang, M. Parkes, H. W. Virgin, Disease-specific alterations in the enteric virome in inflammatory bowel disease, Cell 160 (3) (2015) 447-60.

[19] J. Halfvarson, C. J. Brislawn, R. Lamendella, Y. Vazquez-Baeza, W. A. Walters, L. M. Bramer, M. D'Amato, F. Bonfiglio, D. McDonald, A. Gonzalez, E. E. McClure, M. F. Dunklebarger, R. Knight, J. K. Jansson, Dynamics of the human gut microbiome in inflammatory bowel disease, Nat Microbiol 2 (2017) 17004.

[20] X. C. Morgan, T. L. Tickle, H. Sokol, D. Gevers, K. L. Devaney, D. V. Ward, J. A. Reyes, S. A. Shah, N. LeLeiko, S. B. Snapper, A. Bousvaros, J. Korzenik, B. E. Sands, R. J. Xavier, C.

Huttenhower, Dysfunction of the intestinal microbiome in inflammatory bowel disease and treatment, Genome Biol 13 (9) (2012) R79.

[21] P. J. Kennedy, A. B. Murphy, J. F. Cryan, P. R. Ross, T. G. Dinan, C. Stanton, Microbiome in brain function and mental health, Trends in Food Science & Technology 57 (2016) 289-301.

[22] M. Messaoudi, R. Lalonde, N. Violle, H. Javelot, D. Desor, A. Nejdi, J. F. Bisson, C. Rougeot, M. Pichelin, M. Cazaubiel, J. M. Cazaubiel, Assessment of psychotropic-like properties of a probiotic formulation (Lactobacillus helveticus R0052 and Bifidobacterium longum R0175) in rats and human subjects, Br J Nutr 105 (5) (2011) 755-64.

[23] A. D. Kostic, D. Gevers, C. S. Pedamallu, M. Michaud, F. Duke, A. M. Earl, A. I. Ojesina, J. Jung, A. J. Bass, J. Tabernero, J. Baselga, C. Liu, R. A. Shivdasani, S. Ogino, B. W. Birren, C. Huttenhower, W. S. Garrett, M. Meyerson, Genomic analysis identifies association of Fusobacterium with colorectal carcinoma, Genome Res 22 (2) (2012) 292-8.

[24] W. Chen, F. Liu, Z. Ling, X. Tong, C. Xiang, Human intestinal lumen and mucosa-associated microbiota in patients with colorectal cancer, PLoS One 7 (6) (2012) e39743.

[25] L. Maier, M. Pruteanu, M. Kuhn, G. Zeller, A. Telzerow, E. E. Anderson, A. R. Brochado, K. C. Fernandez, H. Dose, H. Mori, K. R. Patil, P. Bork, A. Typas, Extensive impact of non-antibiotic drugs on human gut bacteria, Nature 555 (7698) (2018) 623-628.

[26] D. Gevers, R. Knight, J. F. Petrosino, K. Huang, A. L. McGuire, B. W. Birren, K. E. Nelson, O. White, B. A. Methe, C. Huttenhower, The Human Microbiome Project: a community resource for the healthy human microbiome, PLoS Biol 10 (8) (2012) e1001377.

[27] B. J. Kelly, R. Gross, K. Bittinger, S. Sherrill-Mix, J. D. Lewis, R. G. Collman, F. D. Bushman, H. Li, Power and sample-size estimation for microbiome studies using pairwise distances and PERMANOVA, Bioinformatics 31 (15) (2015) 2461-8.

[28] P. Bacchetti, Current sample size conventions: flaws, harms, and alternatives, BMC Med 8 (2010) 17.

[29] A. Cotillard, S. P. Kennedy, L. C. Kong, E. Prifti, N. Pons, E. Le Chatelier, M. Almeida, B. Quinquis, F. Levenez, N. Galleron, S. Gougis, S. Rizkalla, J. M. Batto, P. Renault, A. N. R. M. consortium, J. Dore, J. D. Zucker, K. Clement, S. D. Ehrlich, Dietary intervention impact on gut microbial gene richness, Nature 500 (7464) (2013) 585-8.

[30] L. A. David, C. F. Maurice, R. N. Carmody, D. B. Gootenberg, J. E. Button, B. E. Wolfe, A. V. Ling, A. S. Devlin, Y. Varma, M. A. Fischbach, S. B. Biddinger, R. J. Dutton, P. J. Turnbaugh, Diet rapidly and reproducibly alters the human gut microbiome, Nature 505 (7484) (2014) 559-63.

[31] S. M. Finegold, H. R. Attebery, V. L. Sutter, Effect of diet on human fecal flora: comparison of Japanese and American diets, Am J Clin Nutr 27 (12) (1974) 1456-69.

[32] C. Human Microbiome Project, Structure, function and diversity of the healthy human microbiome, Nature 486 (7402) (2012) 207-14.

[33] S. Minot, R. Sinha, J. Chen, H. Li, S. A. Keilbaugh, G. D. Wu, J. D. Lewis, F. D. Bushman, The human gut virome: inter-individual variation and dynamic response to diet, Genome Res 21 (10) (2011) 1616-25.

[34] A. Zhernakova, A. Kurilshikov, M. J. Bonder, E. F. Tigchelaar, M. Schirmer, T. Vatanen, Z. Mujagic, A. V. Vila, G. Falony, S. Vieira-Silva, J. Wang, F. Imhann, E. Brandsma, S. A. Jankipersadsing, M. Joossens, M. C. Cenit, P. Deelen, M. A. Swertz, s. LifeLines cohort, R. K. Weersma, E. J. Feskens, M. G. Netea, D. Gevers, D. Jonkers, L. Franke, Y. S. Aulchenko, C. Huttenhower, J. Raes, M. H. Hofker, R. J. Xavier, C. Wijmenga, J. Fu, Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity, Science 352 (6285) (2016) 565-9.

Appendix A

CHAPTER 3 ADDITIONAL MATERIAL

Supplementary Table 3.1

MetaCyc Pathway	p Value			
	Overall	COH vs COL	COH vs EX	COL vs EX
PWY.6270.isoprene.biosynthesis.I	0.001	0.001	0.036	0.062
PWY.7316.dTDP.N.acetylviosamine.biosynthesis	0.001	0.002	0.002	0.659
PWY.6859.all.trans.farnesol.biosynthesis	0.002	0.005	0.005	0.289
PWY.7315.dTDP.N.acetylthomosamine.biosynthesis	0.002	0.003	0.003	0.848
PWY.7663.gondoate.biosynthesis.anaerobic.	0.002	0.455	0.012	0.006
PWY.5850.superpathway.of.menaquinol.6.biosynthesis.I	0.003	0.075	0.005	0.075
PWY.5860.superpathway.of.demethylmenaquinol.6.biosynthesis.	0.003	0.071	0.005	0.071
PWY.7392.taxadiene.biosynthesis.engineered.	0.003	0.005	0.013	0.201
REDCITCYC.TCA.cycle.VIII.helicobacter.	0.003	0.792	0.009	0.011
PWY.6060.malonate.degradation.II.biotin.dependent.	0.004	0.947	0.022	0.007
PWY.6322.phosphinothricin.tripeptide.biosynthesis	0.004	0.002	0.021	0.402
POLYISOPRENSYN.PWY.polyisoprenoid.biosynthesis.E.coli.	0.005	0.01	0.023	0.167
PWY.5101.L.isoleucine.biosynthesis.II	0.005	0.115	0.007	0.115
PWY0.1261.anhydromuropeptides.recycling	0.005	0.173	0.007	0.081
PWY.7221.guanosine.ribonucleotides.de.novo.biosynthesis	0.006	0.826	0.018	0.018
PWY.7560.methylerythritol.phosphate.pathway.II	0.006	0.015	0.978	0.008
PWY.6906.chitin.derivatives.degradation	0.007	0.021	0.013	0.289
PWY66.398.TCA.cycle.III.animals.	0.007	0.03	0.009	0.48
PWY.5757.fosfomycin.biosynthesis	0.009	0.005	0.045	0.339
PWY.7282.4.amino.2.methyl.5.phosphomethylpyrimidine.biosynthesis.yeast.	0.009	0.218	0.218	0.005
THISYNARA.PWY. superpathway.of.thiamin.diphosphate.biosynthesis.III.eukaryotes.	0.009	0.072	0.013	0.149
PWY0.1586.peptidoglycan.maturation.meso.diaminopimelate.containing.	0.01	0.028	0.012	0.86
GALACTARDEG.PWY.D.galactarate.degradation.I	0.012	0.301	0.017	0.084
GLUCARGALACTSU-PER.superpathway.of.D.glucarate.and.D.galactarate.degradation	0.012	0.301	0.017	0.084
PWY.5189.tetrapyrrole.biosynthesis.II.from.glycine.	0.013	0.482	0.017	0.084
PWY.5283.L.lysine.degradation.V	0.013	0.572	0.028	0.028
PWY.6737.starch.degradation.V	0.013	0.442	0.011	0.136
DENOVOPURINE2.PWY. superpathway.of.purine.nucleotides.de.novo.biosynthesis.II	0.014	0.173	0.021	0.136
NONOXIPENT.PWY. pentose.phosphate.pathway.non.oxidative.branch.	0.015	0.118	0.667	0.007

MetaCyc Pathway	p Value			
	Overall	COH vs COL	COH vs EX	COL vs EX
PWY.5104.L.isoleucine.biosynthesis.IV	0.015	0.166	0.026	0.109
PWY.6572.chondroitin.sulfate.degradation.I.bacterial.	0.015	0.019	0.019	0.883
PWY.7277.sphingolipid.biosynthesis.mammals.	0.016	0.097	0.016	0.224
ALLANTOINDEG.PWY. superpathway.of.allantoin.degradation.in.yeast	0.017	0.272	0.107	0.024
HSERMETANA.PWY.L.methionine.biosynthesis.III	0.017	0.107	0.018	0.296
LYSINE.DEG1.PWY.L.lysine.degradation.XI.mammalian.	0.017	0.068	0.031	0.176
OANTIGEN.PWY.O.antigen.building.blocks.biosynthesis.E.coli.	0.017	0.244	0.015	0.224
PWY.3001.superpathway.of.L.isoleucine.biosynthesis.I	0.017	0.023	0.023	0.499
PWY.6629.superpathway.of.L.tryptophan.biosynthesis	0.017	0.391	0.033	0.066
PWY66.389.phytol.degradation	0.017	0.014	0.533	0.043
METHGLYUT.PWY.superpathway.of.methylglyoxal.degradation	0.018	0.049	0.021	0.627
P23.PWY.reductive.TCA.cycle.I	0.018	0.344	0.153	0.017
COA.PWY.coenzyme.A.biosynthesis.I	0.019	0.033	0.868	0.024
UDPNAGSYN.PWY.UDP.N.acetyl.D.glucosamine.biosynthesis.I	0.019	0.253	0.019	0.206
PWY.5188.tetrapyrrole.biosynthesis.I.from.glutamate.	0.02	0.524	0.024	0.102
PWY.5918.superpathay.of.heme.biosynthesis.from.glutamate	0.02	0.162	0.024	0.162
X.PWY.6307.L.tryptophan.degradation.X.mammalian.via.tryptamine.	0.02	0.022	0.608	0.022
X.PWY.6358.superpathway.of.D.myo.inositol.1.4.5.trisphosphate.metabolism.	0.02	0.583	0.013	0.04
PWY.4242.pantothenate.and.coenzyme.A.biosynthesis.III	0.021	0.136	0.292	0.019
PWY.5862.superpathway.of.demethylmenaquinol.9.biosynthesis	0.021	0.21	0.035	0.124
PWY.6901.superpathway.of.glucose.and.xylose.degradation	0.021	0.072	0.868	0.016
PWY.4981.L.proline.biosynthesis.II.from.arginine.	0.022	0.025	0.497	0.025
PWY.5667.CDP.diacylglycerol.biosynthesis.I	0.022	0.149	0.023	0.27
PWY.5845.superpathway.of.menaquinol.9.biosynthesis	0.022	0.218	0.035	0.128
PWY.5971.palmitate.biosynthesis.II.bacteria.and.plants.	0.022	0.047	0.027	0.941
PWY0.1319.CDP.diacylglycerol.biosynthesis.II	0.022	0.149	0.023	0.27
PWY.5896.superpathway.of.menaquinol.10.biosynthesis	0.023	0.061	0.035	0.362
PWY.6313.serotonin.degradation	0.023	0.027	0.524	0.039
PWY.6435.4.hydroxybenzoate.biosynthesis.V	0.023	0.061	0.026	0.648
PWY.7094.fatty.acid.salvage	0.023	0.052	0.029	0.68
THISYN.PWY.superpathway.of.thiamin.diphosphate.biosynthesis.I	0.023	0.025	0.025	0.638
PWY.5136.fatty.acid.beta.oxidation.II.peroxisome.	0.025	0.107	0.022	0.517
PWY.5676.acetyl.CoA.fermentation.to.butanoate.II	0.025	0.136	0.022	0.41
PWY.7200.superpathway.of.pyrimidine.deoxyribonucleoside.salvage	0.025	0.809	0.043	0.043
PWY0.1241.ADP.L.glycero.beta.D.manno.heptose.biosynthesis	0.025	0.826	0.054	0.054
PWY.7644.heparin.degradation	0.026	0.644	0.083	0.083
PWY3DJ.35471.L.ascorbate.biosynthesis.IV	0.026	0.033	0.033	0.871
PWY.6491.D.galacturonate.degradation.III	0.027	0.048	0.033	0.935
PWY66.422.D.galactose.degradation.V.Leloir.pathway.	0.027	0.229	0.025	0.257
PWY.1501.mandelate.degradation.I	0.028	0.03	0.03	0.941

MetaCyc Pathway	p Value			
	Overall	COH vs COL	COH vs EX	COL vs EX
X.PRPP.PWY.superpathway.of.histidine.purine.and.pyrimidine.biosynthesis.	0.029	0.17	0.03	0.31
PWY.6124.inosine.5.phosphate.biosynthesis.II	0.029	0.312	0.312	0.016
PWY.6163.chorismate.biosynthesis.from.3.dehydroquinate	0.029	0.235	0.235	0.024
PWY.5392.reductive.TCA.cycle.II	0.03	0.843	0.062	0.058
TRPSYN.PWY.L.tryptophan.biosynthesis	0.03	0.416	0.059	0.081
PANTOSYN.PWY.pantothenate.and.coenzyme.A.biosynthesis.I	0.031	0.124	0.57	0.024
RIBOSYN2.PWY.flavin.biosynthesis.I.bacteria.and.plants.	0.031	0.04	0.04	0.617
PWY.6168.flavin.biosynthesis.III.fungi.	0.032	0.044	0.044	0.627
KETOGLUCONMET.PWY.ketogluconate.metabolism	0.035	0.055	0.047	0.93
PENTOSE.P.PWY.pentose.phosphate.pathway	0.035	0.281	0.618	0.015
PWY.6123.inosine.5.phosphate.biosynthesis.I	0.038	0.422	0.429	0.016
PWY.6165.chorismate.biosynthesis.II.archaea.	0.038	0.629	0.14	0.044
PWY.6892.thiazole.biosynthesis.I.E.coli.	0.038	0.044	0.824	0.044
FAO.PWY.fatty.acid.beta.oxidation.I	0.04	0.163	0.033	0.556
BRANCHED.CHAIN.AA.SYN.PWY.superpathway.of.branched.amino.acid.biosynthesis	0.041	0.187	0.077	0.158
XP4.PWY.superpathway.of.L.lysine.L.threonine.and.L.methionine.biosynthesis	0.041	0.17	0.057	0.233
X.P461.PWY.hexitol.fermentation.to.lactate.formate.ethanol.and.acetate.	0.042	0.042	0.708	0.042
X.PWY.7237.myo.chiro.and.scillo.inositol.degradation.	0.042	0.366	0.551	0.018
PWY.7389.superpathway.of.anaerobic.energy.metabolism.invertebrates.	0.042	0.194	0.041	0.377
PWY0.781.aspartate.superpathway	0.042	0.211	0.047	0.276
P122.PWY.heterolactic.fermentation	0.044	0.056	0.153	0.153
PWY66.375.leukotriene.biosynthesis	0.045	0.043	0.098	0.435
X.GLCMANNANAUT.PWY.superpathway.of.N.acetylglucosamine.N.acetylmannosamine.and.N.acetylneuraminate.degradation.	0.047	0.064	0.064	0.317
THREOCAT.PWY.superpathway.of.L.threonine.metabolism	0.048	0.259	0.046	0.332
FERMENTATION.PWY.mixed.acid.fermentation	0.049	0.281	0.281	0.051
P241.PWY.coenzyme.B.biosynthesis	0.049	0.186	0.051	0.393
PWY.5103.L.isoleucine.biosynthesis.III	0.049	0.262	0.102	0.128
PWY.5705.allantoin.degradation.to.glyoxylate.III	0.049	0.509	0.191	0.054
PWY.6121.5.aminoimidazole.ribonucleotide.biosynthesis.I	0.049	0.17	0.074	0.233

Supplementary Table 3.2

Colour of group mean and standard deviation corresponding to pathways: High, low, and mid values

MetaCyc Metabolic Category	Group Mean of Relative Abundance for Statistically Significant Pathways (P value <.05)			Standard Deviation		
Biosynthesis	High BMI	Low BMI	Athlete	High BMI	Low BMI	Athlete
BRANCHED-CHAIN-AA-SYN-PWY: superpathway of branched amino acid biosynthesis	6.49E+03	6.37E+03	6.13E+03	7.15E+02	6.03E+02	7.37E+02
COA-PWY: coenzyme A biosynthesis I	4.36E+03	3.99E+03	4.46E+03	6.67E+02	9.71E+02	8.56E+02
DENOVOPURINE2-PWY: superpathway of purine nucleotides de novo biosynthesis II	2.92E+03	3.09E+03	3.38E+03	5.55E+02	6.63E+02	4.43E+02
HSERMETANA-PWY: L-methionine biosynthesis III	6.99E+03	6.56E+03	6.33E+03	1.03E+03	1.15E+03	1.15E+03
LYSINE-AMINOAD-PWY: L-lysine biosynthesis IV	5.42E+01	6.26E+01	7.51E+01	2.44E+01	2.28E+01	2.58E+01
OANTIGEN-PWY: O-antigen building blocks biosynthesis (E. coli)	2.13E+03	2.39E+03	2.68E+03	6.07E+02	6.59E+02	5.73E+02
P241-PWY: coenzyme B biosynthesis	1.42E+02	1.74E+02	1.92E+02	5.13E+01	6.26E+01	6.73E+01
P4-PWY: superpathway of L-lysine, L-threonine and L-methionine biosynthesis I	2.34E+03	2.58E+03	2.73E+03	5.72E+02	5.68E+02	5.72E+02
PANTOSYN-PWY: pantothenate and coenzyme A biosynthesis I	4.47E+03	4.24E+03	4.59E+03	4.89E+02	6.56E+02	6.62E+02
POLYISOPRENSYN-PWY: polyisoprenoid biosynthesis (E. coli)	3.08E+03	2.42E+03	2.75E+03	8.13E+02	8.60E+02	7.32E+02
PWY0-1241: ADP-L-glycero-β-D-manno-heptose biosynthesis	1.80E+02	1.66E+02	2.84E+02	8.71E+01	8.56E+01	1.00E+02
PWY0-1319: CDP-diacylglycerol biosynthesis II	4.35E+03	4.10E+03	4.03E+03	3.83E+02	3.98E+02	5.11E+02
PWY0-1586: peptidoglycan maturation (meso-diaminopimelate containing)	6.69E+03	5.91E+03	6.05E+03	9.10E+02	9.47E+02	1.02E+03
PWY-3001: superpathway of L-isoleucine biosynthesis I	6.04E+03	5.85E+03	5.79E+03	4.62E+02	4.48E+02	4.51E+02
PWY3DJ-35471: L-ascorbate biosynthesis IV	2.43E+01	3.22E+01	3.71E+01	2.85E+01	2.88E+01	4.94E+01
PWY-4242: pantothenate and coenzyme A biosynthesis III	4.12E+03	3.92E+03	4.33E+03	4.34E+02	6.69E+02	6.65E+02
PWY-4981: L-proline biosynthesis II (from arginine)	1.47E+03	2.07E+03	1.63E+03	8.09E+02	8.24E+02	7.62E+02
PWY-5101: L-isoleucine biosynthesis II	4.87E+03	4.51E+03	4.28E+03	7.27E+02	7.35E+02	7.32E+02
PWY-5103: L-isoleucine biosynthesis III	6.14E+03	6.01E+03	5.79E+03	7.14E+02	5.98E+02	7.06E+02
PWY-5104: L-isoleucine biosynthesis IV	7.01E+03	6.81E+03	6.58E+03	7.62E+02	6.28E+02	6.86E+02
PWY-5188: tetrapyrrole biosynthesis I (from glutamate)	1.01E+03	1.07E+03	1.33E+03	3.54E+02	3.71E+02	3.70E+02
PWY-5189: tetrapyrrole biosynthesis II (from glycine)	7.03E+02	7.56E+02	9.62E+02	2.56E+02	2.74E+02	2.83E+02
PWY-5667: CDP-diacylglycerol biosynthesis I	4.35E+03	4.10E+03	4.03E+03	3.83E+02	3.98E+02	5.11E+02

MetaCyc Metabolic Category	Group Mean of Relative Abundance for Statistically Significant Pathways (P value <.05)			Standard Deviation		
PWY-5757: fosfomycin biosynthesis	1.12E+00	3.23E+00	3.32E+00	9.79E-01	1.89E+00	1.32E+00
PWY-5845: superpathway of menaquinol-9 biosynthesis	1.87E+02	2.07E+02	2.83E+02	1.85E+02	1.70E+02	2.88E+02
PWY-5850: superpathway of menaquinol-6 biosynthesis I	1.47E+02	1.67E+02	2.53E+02	1.88E+02	1.76E+02	2.81E+02
PWY-5860: superpathway of demethylmena-quinol-6 biosynthesis I	1.01E+02	1.15E+02	1.77E+02	1.33E+02	1.25E+02	2.09E+02
PWY-5862: superpathway of demethylmena-quinol-9 biosynthesis	1.28E+02	1.42E+02	1.97E+02	1.31E+02	1.21E+02	2.14E+02
PWY-5896: superpathway of menaquinol-10 biosynthesis	1.48E+02	1.71E+02	2.34E+02	1.89E+02	1.75E+02	2.81E+02
PWY-5918: superpathay of heme biosynthesis from glutamate	4.46E+02	5.02E+02	6.23E+02	1.66E+02	2.02E+02	2.38E+02
PWY-5971: palmitate biosynthesis II (bacteria and plants)	7.15E+03	6.31E+03	6.43E+03	1.04E+03	1.08E+03	1.17E+03
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I	6.47E+03	5.97E+03	5.84E+03	1.20E+03	1.15E+03	1.11E+03
PWY-6123: inosine-5'-phosphate biosynthesis I	5.83E+03	5.43E+03	5.83E+03	1.12E+03	1.11E+03	1.07E+03
PWY-6124: inosine-5'-phosphate biosynthesis II	6.31E+03	5.82E+03	6.34E+03	1.43E+03	1.27E+03	1.37E+03
PWY-6125: superpathway of guanosine nucleotides de novo biosynthesis II	3.80E+03	4.19E+03	4.30E+03	9.29E+02	1.04E+03	9.20E+02
PWY-6163: chorismate biosynthesis from 3-dehydroquinate	5.40E+03	5.23E+03	5.52E+03	5.17E+02	5.82E+02	5.92E+02
PWY-6165: chorismate biosynthesis II (archaea)	2.56E+01	2.27E+01	4.21E+01	1.85E+01	2.88E+01	2.95E+01
PWY-6168: flavin biosynthesis III (fungi)	5.04E+03	4.72E+03	4.65E+03	6.87E+02	7.35E+02	7.21E+02
PWY-6270: isoprene biosynthesis I	5.44E+03	4.57E+03	4.92E+03	8.10E+02	1.04E+03	8.11E+02
PWY-6322: phosphinothricin tripeptide biosynthesis	1.42E+00	3.97E+00	4.38E+00	1.33E+00	1.79E+00	1.82E+00
PWY-6358: superpathway of D-myo-inositol (1,4,5)-trisphosphate metabolism	6.46E-02	6.08E-02	2.52E-03	1.29E-01	1.82E-01	8.37E-02
PWY-6435: 4-hydroxybenzoate biosynthesis V	1.49E+03	1.77E+03	1.80E+03	4.35E+02	4.84E+02	4.95E+02
PWY-6629: superpathway of L-tryptophan biosynthesis	4.05E+02	4.44E+02	5.82E+02	2.01E+02	2.11E+02	2.69E+02
PWY66-375: leukotriene biosynthesis	4.16E-01	6.84E-01	7.22E-01	6.45E-01	6.36E-01	3.54E-01
PWY-6892: thiazole biosynthesis I (E. coli)	4.76E+03	4.39E+03	4.71E+03	7.61E+02	8.15E+02	9.12E+02

MetaCyc Metabolic Category	Group Mean of Relative Abundance for Statistically Significant Pathways (P value <.05)			Standard Deviation		
PWY-7094: fatty acid salvage	1.11E+03	1.37E+03	1.37E+03	3.38E+02	3.75E+02	3.96E+02
PWY-7200: superpathway of pyrimidine deoxyribonucleoside salvage	3.76E+02	3.61E+02	4.98E+02	4.51E+02	3.72E+02	4.28E+02
PWY-7221: guanosine ribonucleotides de novo biosynthesis	7.11E+03	6.86E+03	7.41E+03	1.67E+03	1.68E+03	1.64E+03
PWY-7277: sphingolipid biosynthesis (mammals)	0.00E+00	2.01E-01	7.95E-01	0.00E+00	4.90E-01	5.34E-01
PWY-7282: 4-amino-2-methyl-5-phosphomethylpyrimidine biosynthesis (yeast)	3.03E+03	3.26E+03	2.78E+03	6.46E+02	6.11E+02	5.81E+02
PWY-7315: dTDP-N-acetylthomosamine biosynthesis	1.98E+02	3.58E+02	4.20E+02	1.15E+02	1.31E+02	2.48E+02
PWY-7316: dTDP-N-acetylviosamine biosynthesis	4.01E+02	7.12E+02	7.25E+02	2.94E+02	2.87E+02	2.97E+02
PWY-7392: taxadiene biosynthesis (engineered)	4.21E+03	3.43E+03	3.72E+03	7.55E+02	9.15E+02	7.39E+02
PWY-7560: methylerythritol phosphate pathway II	5.41E+03	4.75E+03	5.36E+03	8.77E+02	9.93E+02	9.09E+02
PWY-7663: gondoate biosynthesis (anaerobic)	9.30E+03	9.55E+03	8.60E+03	8.96E+02	1.05E+03	1.23E+03
RIBOSYN2-PWY: flavin biosynthesis I (bacteria and plants)	5.00E+03	4.65E+03	4.62E+03	6.97E+02	7.53E+02	6.80E+02
THISYNARA-PWY: superpathway of thiamin diphosphate biosynthesis III (eukaryotes)	5.27E+03	4.82E+03	4.67E+03	1.00E+03	9.97E+02	9.02E+02
THISYN-PWY: superpathway of thiamin diphosphate biosynthesis I	5.15E+03	4.58E+03	4.58E+03	9.44E+02	8.68E+02	8.47E+02
TRPSYN-PWY: L-tryptophan biosynthesis	2.18E+02	2.35E+02	2.95E+02	1.50E+02	1.31E+02	1.88E+02
UDPNAGSYN-PWY: UDP-N-acetyl-D-glucosamine biosynthesis I	1.40E+03	1.62E+03	1.93E+03	5.50E+02	5.92E+02	5.39E+02
Polyamine Biosynthesis	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
UDPNAGSYN-PWY: UDP-N-acetyl-D-glucosamine biosynthesis I	1.40E+03	1.62E+03	1.93E+03	5.50E+02	5.92E+02	5.39E+02
Amino Acid Biosynthesis	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
BRANCHED-CHAIN-AA-SYN-PWY: superpathway of branched amino acid biosynthesis	6.49E+03	6.37E+03	6.13E+03	7.15E+02	6.03E+02	7.37E+02
HSERMETANA-PWY: L-methionine biosynthesis III	6.99E+03	6.56E+03	6.33E+03	1.03E+03	1.15E+03	1.15E+03
LYSINE-AMINOAD-PWY: L-lysine biosynthesis IV	5.42E+01	6.26E+01	7.51E+01	2.44E+01	2.28E+01	2.58E+01
P4-PWY: superpathway of L-lysine, L-threonine and L-methionine biosynthesis I	2.34E+03	2.58E+03	2.73E+03	5.72E+02	5.68E+02	5.72E+02

MetaCyc Metabolic Category	Group Mean of Relative Abundance for Statistically Significant Pathways (P value <.05)			Standard Deviation		
PWY-3001: superpathway of L-isoleucine biosynthesis I	6.04E+03	5.85E+03	5.79E+03	4.62E+02	4.48E+02	4.51E+02
PWY-4981: L-proline biosynthesis II (from arginine)	1.47E+03	2.07E+03	1.63E+03	8.09E+02	8.24E+02	7.62E+02
PWY-5101: L-isoleucine biosynthesis II	4.87E+03	4.51E+03	4.28E+03	7.27E+02	7.35E+02	7.32E+02
PWY-5103: L-isoleucine biosynthesis III	6.14E+03	6.01E+03	5.79E+03	7.14E+02	5.98E+02	7.06E+02
PWY-5104: L-isoleucine biosynthesis IV	7.01E+03	6.81E+03	6.58E+03	7.62E+02	6.28E+02	6.86E+02
PWY-6629: superpathway of L-tryptophan biosynthesis	4.05E+02	4.44E+02	5.82E+02	2.01E+02	2.11E+02	2.69E+02
TRPSYN-PWY: L-tryptophan biosynthesis	2.18E+02	2.35E+02	2.95E+02	1.50E+02	1.31E+02	1.88E+02
Aromatic Compound Biosynthesis	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
PWY-6163: chorismate biosynthesis from 3-dehydroquinate	5.40E+03	5.23E+03	5.52E+03	5.17E+02	5.82E+02	5.92E+02
PWY-6165: chorismate biosynthesis II (archaea)	2.56E+01	2.27E+01	4.21E+01	1.85E+01	2.88E+01	2.95E+01
PWY-6435: 4-hydroxybenzoate biosynthesis V	1.49E+03	1.77E+03	1.80E+03	4.35E+02	4.84E+02	4.95E+02
Carbohydrates Biosynthesis	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
PWY0-1241: ADP-L-glycero-β-D-manno-heptose biosynthesis	1.80E+02	1.66E+02	2.84E+02	8.71E+01	8.56E+01	1.00E+02
PWY-7315: dTDP-N-acetylthomosamine biosynthesis	1.98E+02	3.58E+02	4.20E+02	1.15E+02	1.31E+02	2.48E+02
PWY-7316: dTDP-N-acetylviosamine biosynthesis	4.01E+02	7.12E+02	7.25E+02	2.94E+02	2.87E+02	2.97E+02
Cell-Structure Biosynthesis	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
OANTIGEN-PWY: O-antigen building blocks biosynthesis (E. coli)	2.13E+03	2.39E+03	2.68E+03	6.07E+02	6.59E+02	5.73E+02
PWY0-1586: peptidoglycan maturation (meso-diaminopimelate containing)	6.69E+03	5.91E+03	6.05E+03	9.10E+02	9.47E+02	1.02E+03
PWY-7315: dTDP-N-acetylthomosamine biosynthesis	1.98E+02	3.58E+02	4.20E+02	1.15E+02	1.31E+02	2.48E+02
PWY-7316: dTDP-N-acetylviosamine biosynthesis	4.01E+02	7.12E+02	7.25E+02	2.94E+02	2.87E+02	2.97E+02
UDPNAGSYN-PWY: UDP-N-acetyl-D-glucosamine biosynthesis I	1.40E+03	1.62E+03	1.93E+03	5.50E+02	5.92E+02	5.39E+02
Cofactor-Biosynthesis	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**

MetaCyc Metabolic Category	Group Mean of Relative Abundance for Statistically Significant Pathways (P value <.05)			Standard Deviation		
COA-PWY: coenzyme A biosynthesis I	4.36E+03	3.99E+03	4.46E+03	6.67E+02	9.71E+02	8.56E+02
P241-PWY: coenzyme B biosynthesis	1.42E+02	1.74E+02	1.92E+02	5.13E+01	6.26E+01	6.73E+01
PANTOSYN-PWY: pantothenate and coenzyme A biosynthesis I	4.47E+03	4.24E+03	4.59E+03	4.89E+02	6.56E+02	6.62E+02
POLYISOPRENSYN-PWY: polyisoprenoid biosynthesis (E. coli)	3.08E+03	2.42E+03	2.75E+03	8.13E+02	8.60E+02	7.32E+02
PWY3DJ-35471: L-ascorbate biosynthesis IV	2.43E+01	3.22E+01	3.71E+01	2.85E+01	2.88E+01	4.94E+01
PWY-4242: pantothenate and coenzyme A biosynthesis III	4.12E+03	3.92E+03	4.33E+03	4.34E+02	6.69E+02	6.65E+02
PWY-5188: tetrapyrrole biosynthesis I (from glutamate)	1.01E+03	1.07E+03	1.33E+03	3.54E+02	3.71E+02	3.70E+02
PWY-5189: tetrapyrrole biosynthesis II (from glycine)	7.03E+02	7.56E+02	9.62E+02	2.56E+02	2.74E+02	2.83E+02
PWY-5845: superpathway of menaquinol-9 biosynthesis	1.87E+02	2.07E+02	2.83E+02	1.85E+02	1.70E+02	2.88E+02
PWY-5850: superpathway of menaquinol-6 biosynthesis I	1.47E+02	1.67E+02	2.53E+02	1.88E+02	1.76E+02	2.81E+02
PWY-5860: superpathway of demethylmena-quinol-6 biosynthesis I	1.01E+02	1.15E+02	1.77E+02	1.33E+02	1.25E+02	2.09E+02
PWY-5862: superpathway of demethylmena-quinol-9 biosynthesis	1.28E+02	1.42E+02	1.97E+02	1.31E+02	1.21E+02	2.14E+02
PWY-5896: superpathway of menaquinol-10 biosynthesis	1.48E+02	1.71E+02	2.34E+02	1.89E+02	1.75E+02	2.81E+02
PWY-5918: superpathay of heme biosynthesis from glutamate	4.46E+02	5.02E+02	6.23E+02	1.66E+02	2.02E+02	2.38E+02
PWY-6168: flavin biosynthesis III (fungi)	5.04E+03	4.72E+03	4.65E+03	6.87E+02	7.35E+02	7.21E+02
PWY-6892: thiazole biosynthesis I (E. coli)	4.76E+03	4.39E+03	4.71E+03	7.61E+02	8.15E+02	9.12E+02
PWY-7282: 4-amino-2-methyl-5-phosphomethylpyrimidine biosynthesis (yeast)	3.03E+03	3.26E+03	2.78E+03	6.46E+02	6.11E+02	5.81E+02
RIBOSYN2-PWY: flavin biosynthesis I (bacteria and plants)	5.00E+03	4.65E+03	4.62E+03	6.97E+02	7.53E+02	6.80E+02
THISYNARA-PWY: superpathway of thiamin diphosphate biosynthesis III (eukaryotes)	5.27E+03	4.82E+03	4.67E+03	1.00E+03	9.97E+02	9.02E+02
THISYN-PWY: superpathway of thiamin diphosphate biosynthesis I	5.15E+03	4.58E+03	4.58E+03	9.44E+02	8.68E+02	8.47E+02
Lipid Biosynthesis	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
PWY0-1319: CDP-diacylglycerol biosynthesis II	4.35E+03	4.10E+03	4.03E+03	3.83E+02	3.98E+02	5.11E+02

MetaCyc Metabolic Category	Group Mean of Relative Abundance for Statistically Significant Pathways (P value <.05)			Standard Deviation		
PWY-5667: CDP-diacylglycerol biosynthesis I	4.35E+03	4.10E+03	4.03E+03	3.83E+02	3.98E+02	5.11E+02
PWY-5971: palmitate biosynthesis II (bacteria and plants)	7.15E+03	6.31E+03	6.43E+03	1.04E+03	1.08E+03	1.17E+03
PWY-7094: fatty acid salvage	1.11E+03	1.37E+03	1.37E+03	3.38E+02	3.75E+02	3.96E+02
PWY-7277: sphingolipid biosynthesis (mammals)	0.00E+00	2.01E-01	7.95E-01	0.00E+00	4.90E-01	5.34E-01
PWY-7663: gondoate biosynthesis (anaerobic)	9.30E+03	9.55E+03	8.60E+03	8.96E+02	1.05E+03	1.23E+03
Hormone Biosynthesis	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
PWY66-375: leukotriene biosynthesis	4.16E-01	6.84E-01	7.22E-01	6.45E-01	6.36E-01	3.54E-01
Nucleotide Biosynthesis	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
DENOVOPURINE2-PWY: superpathway of purine nucleotides de novo biosynthesis II	2.92E+03	3.09E+03	3.38E+03	5.55E+02	6.63E+02	4.43E+02
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I	6.47E+03	5.97E+03	5.84E+03	1.20E+03	1.15E+03	1.11E+03
PWY-6123: inosine-5'-phosphate biosynthesis I	5.83E+03	5.43E+03	5.83E+03	1.12E+03	1.11E+03	1.07E+03
PWY-6124: inosine-5'-phosphate biosynthesis II	6.31E+03	5.82E+03	6.34E+03	1.43E+03	1.27E+03	1.37E+03
PWY-6125: superpathway of guanosine nucleotides de novo biosynthesis II	3.80E+03	4.19E+03	4.30E+03	9.29E+02	1.04E+03	9.20E+02
PWY-7200: superpathway of pyrimidine deoxyribonucleoside salvage	3.76E+02	3.61E+02	4.98E+02	4.51E+02	3.72E+02	4.28E+02
PWY-7221: guanosine ribonucleotides de novo biosynthesis	7.11E+03	6.86E+03	7.41E+03	1.67E+03	1.68E+03	1.64E+03
Secondary Metabolite Biosynthesis	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
PWY-5757: fosfomycin biosynthesis	1.12E+00	3.23E+00	3.32E+00	9.79E-01	1.89E+00	1.32E+00
PWY-6270: isoprene biosynthesis I	5.44E+03	4.57E+03	4.92E+03	8.10E+02	1.04E+03	8.11E+02
PWY-6322: phosphinothricin tripeptide biosynthesis	1.42E+00	3.97E+00	4.38E+00	1.33E+00	1.79E+00	1.82E+00
PWY-6358: superpathway of D-myo-inositol (1,4,5)-trisphosphate metabolism	6.46E-02	6.08E-02	2.52E-03	1.29E-01	1.82E-01	8.37E-02
PWY-7392: taxadiene biosynthesis (engineered)	4.21E+03	3.43E+03	3.72E+03	7.55E+02	9.15E+02	7.39E+02
PWY-7560: methylerythritol phosphate pathway II	5.41E+03	4.75E+03	5.36E+03	8.77E+02	9.93E+02	9.09E+02

MetaCyc Metabolic Category	Group Mean of Relative Abundance for Statistically Significant Pathways (P value <.05)			Standard Deviation		
Degradation	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
ALLANTOINDEG-PWY: superpathway of allantoin degradation in yeast	1.71E+02	1.47E+02	2.32E+02	1.43E+02	1.63E+02	1.87E+02
FAO-PWY: fatty acid β-oxidation I	1.52E+03	1.82E+03	1.89E+03	4.73E+02	5.23E+02	5.53E+02
GALACTARDEG-PWY: D-galactarate degradation I	2.70E+02	3.06E+02	3.97E+02	1.30E+02	1.38E+02	2.00E+02
GLCMANNANAUT-PWY: superpathway of N-acetylglucosamine, N-acetylmannosamine and N-acetylneuraminate degradation	9.84E+02	1.33E+03	1.17E+03	4.63E+02	4.78E+02	3.73E+02
KETOGLUCONMET-PWY: ketogluconate metabolism	2.51E+02	2.79E+02	3.08E+02	1.87E+02	1.84E+02	2.53E+02
LACTOSECAT-PWY: lactose and galactose degradation I	8.84E+02	9.39E+02	1.15E+03	4.59E+02	5.02E+02	5.00E+02
LYSINE-DEG1-PWY: L-lysine degradation XI (mammalian)	1.37E+00	2.63E+00	3.63E+00	2.06E+00	1.96E+00	1.90E+00
P23-PWY: reductive TCA cycle I	3.15E+02	2.53E+02	4.21E+02	2.11E+02	1.92E+02	2.58E+02
P461-PWY: hexitol fermentation to lactate, formate, ethanol and acetate	4.53E+02	6.32E+02	4.97E+02	4.25E+02	4.45E+02	5.60E+02
PWY0-1261: anhydromuropeptides recycling	1.18E+03	1.33E+03	1.71E+03	6.01E+02	6.12E+02	6.35E+02
PWY-1501: mandelate degradation I	2.98E+00	5.44E+00	5.69E+00	2.48E+00	2.04E+00	2.25E+00
PWY-5136: fatty acid β-oxidation II (peroxisome)	1.55E+03	1.86E+03	1.94E+03	4.73E+02	5.23E+02	5.52E+02
PWY-5283: L-lysine degradation V	4.04E-01	6.21E-01	3.05E-01	5.66E-01	4.48E-01	5.76E-01
PWY-5705: allantoin degradation to glyoxylate III	4.99E+01	3.52E+01	6.96E+01	5.12E+01	4.82E+01	1.15E+02
PWY-6307: L-tryptophan degradation X (mammalian, via tryptamine)	1.74E+03	2.22E+03	1.86E+03	6.16E+02	6.38E+02	6.32E+02
PWY-6313: serotonin degradation	1.53E+03	1.97E+03	1.65E+03	5.30E+02	5.67E+02	5.75E+02
PWY-6491: D-galacturonate degradation III	3.75E-01	7.57E-01	6.44E-01	1.06E+00	1.09E+00	9.72E-01
PWY-6572: chondroitin sulfate degradation I (bacterial)	1.39E+02	1.94E+02	1.92E+02	7.25E+01	6.75E+01	7.09E+01
PWY66-389: phytol degradation	2.13E+03	2.90E+03	2.38E+03	8.36E+02	9.29E+02	9.98E+02
PWY66-422: D-galactose degradation V (Leloir pathway)	6.06E+03	5.69E+03	5.44E+03	1.15E+03	1.22E+03	1.27E+03
PWY-6737: starch degradation V	6.69E+03	6.50E+03	6.07E+03	8.53E+02	9.71E+02	9.92E+02

MetaCyc Metabolic Category	Group Mean of Relative Abundance for Statistically Significant Pathways (P value <.05)			Standard Deviation		
PWY-6901: superpathway of glucose and xylose degradation	3.13E+03	3.88E+03	3.22E+03	1.28E+03	1.29E+03	1.24E+03
PWY-6906: chitin derivatives degradation	1.97E+01	3.57E+01	6.50E+01	2.98E+01	4.93E+01	5.00E+01
PWY-7237: myo-, chiro- and scillo-inositol degradation	3.52E+03	4.09E+03	3.02E+03	1.91E+03	1.79E+03	1.88E+03
PWY-7644: heparin degradation	7.16E-01	1.46E+00	1.42E+01	2.38E+00	1.34E+01	1.21E+01
THREOCAT-PWY: superpathway of L-threonine metabolism	8.66E+00	3.11E+01	2.54E+01	1.40E+01	1.28E+01	4.87E+01
Alcohol Degradation	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
PWY66-389: phytol degradation	2.13E+03	2.90E+03	2.38E+03	8.36E+02	9.29E+02	9.98E+02
Amine Degradation	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
ALLANTOINDEG-PWY: superpathway of allantoin degradation in yeast	1.71E+02	1.47E+02	2.32E+02	1.43E+02	1.63E+02	1.87E+02
GLCMANNANAUT-PWY: superpathway of N-acetylglucosamine, N-acetylmannosamine and N-acetylneuraminate degradation	9.84E+02	1.33E+03	1.17E+03	4.63E+02	4.78E+02	3.73E+02
PWY-5705: allantoin degradation to glyoxylate III	4.99E+01	3.52E+01	6.96E+01	5.12E+01	4.82E+01	1.15E+02
Amino Acid Degradation	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
LYSINE-DEG1-PWY: L-lysine degradation XI (mammalian)	1.37E+00	2.63E+00	3.63E+00	2.06E+00	1.96E+00	1.90E+00
PWY-5283: L-lysine degradation V	4.04E-01	6.21E-01	3.05E-01	5.66E-01	4.48E-01	5.76E-01
PWY-6307: L-tryptophan degradation X (mammalian, via tryptamine)	1.74E+03	2.22E+03	1.86E+03	6.16E+02	6.38E+02	6.32E+02
THREOCAT-PWY: superpathway of L-threonine metabolism	8.66E+00	3.11E+01	2.54E+01	1.40E+01	1.28E+01	4.87E+01
Aromatic Compound Degradation	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
PWY-1501: mandelate degradation I	2.98E+00	5.44E+00	5.69E+00	2.48E+00	2.04E+00	2.25E+00
C1 Compounds	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
P23-PWY: reductive TCA cycle I	3.15E+02	2.53E+02	4.21E+02	2.11E+02	1.92E+02	2.58E+02
Carbohydrates Degradation	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**

MetaCyc Metabolic Category	Group Mean of Relative Abundance for Statistically Significant Pathways (P value <.05)			Standard Deviation		
LACTOSECAT-PWY: lactose and galactose degradation I	8.84E+02	9.39E+02	1.15E+03	4.59E+02	5.02E+02	5.00E+02
PWY-6572: chondroitin sulfate degradation I (bacterial)	1.39E+02	1.94E+02	1.92E+02	7.25E+01	6.75E+01	7.09E+01
PWY66-422: D-galactose degradation V (Leloir pathway)	6.06E+03	5.69E+03	5.44E+03	1.15E+03	1.22E+03	1.27E+03
PWY-6737: starch degradation V	6.69E+03	6.50E+03	6.07E+03	8.53E+02	9.71E+02	9.92E+02
PWY-6901: superpathway of glucose and xylose degradation	3.13E+03	3.88E+03	3.22E+03	1.28E+03	1.29E+03	1.24E+03
PWY-6906: chitin derivatives degradation	1.97E+01	3.57E+01	6.50E+01	2.98E+01	4.93E+01	5.00E+01
PWY-7644: heparin degradation	7.16E-01	1.46E+00	1.42E+01	2.38E+00	1.34E+01	1.21E+01
Carboxylates Degradation	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
GALACTARDEG-PWY: D-galactarate degradation I	2.70E+02	3.06E+02	3.97E+02	1.30E+02	1.38E+02	2.00E+02
KETOGLUCONMET-PWY: ketogluconate metabolism	2.51E+02	2.79E+02	3.08E+02	1.87E+02	1.84E+02	2.53E+02
PWY-6491: D-galacturonate degradation III	3.75E-01	7.57E-01	6.44E-01	1.06E+00	1.09E+00	9.72E-01
Fatty Acid and Lipid Degradation	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
FAO-PWY: fatty acid β-oxidation I	1.52E+03	1.82E+03	1.89E+03	4.73E+02	5.23E+02	5.53E+02
PWY-5136: fatty acid β-oxidation II (peroxisome)	1.55E+03	1.86E+03	1.94E+03	4.73E+02	5.23E+02	5.52E+02
Hormone Degradation	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
PWY-6313: serotonin degradation	1.53E+03	1.97E+03	1.65E+03	5.30E+02	5.67E+02	5.75E+02
Polymer Degradation	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
PWY-6572: chondroitin sulfate degradation I (bacterial)	1.39E+02	1.94E+02	1.92E+02	7.25E+01	6.75E+01	7.09E+01
PWY-6737: starch degradation V	6.69E+03	6.50E+03	6.07E+03	8.53E+02	9.71E+02	9.92E+02
PWY-6906: chitin derivatives degradation	1.97E+01	3.57E+01	6.50E+01	2.98E+01	4.93E+01	5.00E+01
PWY-7644: heparin degradation	7.16E-01	1.46E+00	1.42E+01	2.38E+00	1.34E+01	1.21E+01
Secondary Metabolite Degradation	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**

MetaCyc Metabolic Category	Group Mean of Relative Abundance for Statistically Significant Pathways (P value <.05)			Standard Deviation		
GALACTARDEG-PWY: D-galactarate degradation I	2.70E+02	3.06E+02	3.97E+02	1.30E+02	1.38E+02	2.00E+02
KETOGLUCONMET-PWY: ketogluconate metabolism	2.51E+02	2.79E+02	3.08E+02	1.87E+02	1.84E+02	2.53E+02
P461-PWY: hexitol fermentation to lactate, formate, ethanol and acetate	4.53E+02	6.32E+02	4.97E+02	4.25E+02	4.45E+02	5.60E+02
PWY0-1261: anhydromuropeptides recycling	1.18E+03	1.33E+03	1.71E+03	6.01E+02	6.12E+02	6.35E+02
PWY-6491: D-galacturonate degradation III	3.75E-01	7.57E-01	6.44E-01	1.06E+00	1.09E+00	9.72E-01
PWY-7237: myo-, chiro- and scillo-inositol degradation	3.52E+03	4.09E+03	3.02E+03	1.91E+03	1.79E+03	1.88E+03
Energy Metabolism	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
FERMENTATION-PWY: mixed acid fermentation	5.86E+02	4.70E+02	7.72E+02	3.80E+02	4.17E+02	4.38E+02
NONOXIPENT-PWY: pentose phosphate pathway (non-oxidative branch)	6.31E+03	6.75E+03	6.23E+03	9.61E+02	8.75E+02	8.81E+02
P122-PWY: heterolactic fermentation	8.80E+02	1.21E+03	1.02E+03	4.15E+02	4.62E+02	4.17E+02
P461-PWY: hexitol fermentation to lactate, formate, ethanol and acetate	4.53E+02	6.32E+02	4.97E+02	4.25E+02	4.45E+02	5.60E+02
PENTOSE-P-PWY: pentose phosphate pathway	2.87E+03	3.58E+03	2.82E+03	1.42E+03	1.40E+03	1.37E+03
PWY-5676: acetyl-CoA fermentation to butanoate II	1.16E+03	1.40E+03	1.49E+03	3.58E+02	3.83E+02	4.32E+02
PWY-5913: TCA cycle VI (obligate autotrophs)	3.38E+02	2.58E+02	4.72E+02	2.69E+02	2.89E+02	3.04E+02
PWY66-398: TCA cycle III (animals)	3.47E+02	4.53E+02	4.71E+02	1.31E+02	1.63E+02	1.52E+02
PWY-6859: all-trans-farnesol biosynthesis	2.73E+03	2.06E+03	2.26E+03	5.95E+02	7.58E+02	6.16E+02
PWY-7389: superpathway of anaerobic energy metabolism (invertebrates)	3.64E+02	4.72E+02	5.50E+02	2.40E+02	2.50E+02	2.30E+02
REDCITCYC: TCA cycle VIII (helicobacter)	2.89E+02	2.74E+02	4.68E+02	1.74E+02	1.61E+02	2.17E+02
Fermentation	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
FERMENTATION-PWY: mixed acid fermentation	5.86E+02	4.70E+02	7.72E+02	3.80E+02	4.17E+02	4.38E+02
P122-PWY: heterolactic fermentation	8.80E+02	1.21E+03	1.02E+03	4.15E+02	4.62E+02	4.17E+02
P461-PWY: hexitol fermentation to lactate, formate, ethanol and acetate	4.53E+02	6.32E+02	4.97E+02	4.25E+02	4.45E+02	5.60E+02
PWY-5676: acetyl-CoA fermentation to butanoate II	1.16E+03	1.40E+03	1.49E+03	3.58E+02	3.83E+02	4.32E+02
PWY-7389: superpathway of anaerobic energy metabolism (invertebrates)	3.64E+02	4.72E+02	5.50E+02	2.40E+02	2.50E+02	2.30E+02
Pentose-Phosphate-Cycle	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**

MetaCyc Metabolic Category	Group Mean of Relative Abundance for Statistically Significant Pathways (P value <.05)			Standard Deviation		
NONOXIPENT-PWY: pentose phosphate pathway (non-oxidative branch)	6.31E+03	6.75E+03	6.23E+03	9.61E+02	8.75E+02	8.81E+02
PENTOSE-P-PWY: pentose phosphate pathway	2.87E+03	3.58E+03	2.82E+03	1.42E+03	1.40E+03	1.37E+03
TCA Variants	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
PWY-5913: TCA cycle VI (obligate autotrophs)	3.38E+02	2.58E+02	4.72E+02	2.69E+02	2.89E+02	3.04E+02
PWY66-398: TCA cycle III (animals)	3.47E+02	4.53E+02	4.71E+02	1.31E+02	1.63E+02	1.52E+02
REDCITCYC: TCA cycle VIII (helicobacter)	2.89E+02	2.74E+02	4.68E+02	1.74E+02	1.61E+02	2.17E+02
Metabolic-Clusters	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
OANTIGEN-PWY: O-antigen building blocks biosynthesis (E. coli)	2.13E+03	2.39E+03	2.68E+03	6.07E+02	6.59E+02	5.73E+02
Super-Pathways	**High BMI**	**Low BMI**	**Athlete**	**High BMI**	**Low BMI**	**Athlete**
ALLANTOINDEG-PWY: superpathway of allantoin degradation in yeast	1.71E+02	1.47E+02	2.32E+02	1.43E+02	1.63E+02	1.87E+02
BRANCHED-CHAIN-AA-SYN-PWY: superpathway of branched amino acid biosynthesis	6.49E+03	6.37E+03	6.13E+03	7.15E+02	6.03E+02	7.37E+02
DENOVOPURINE2-PWY: superpathway of purine nucleotides de novo biosynthesis II	2.92E+03	3.09E+03	3.38E+03	5.55E+02	6.63E+02	4.43E+02
GLCMANNANAUT-PWY: superpathway of N-acetylglucosamine, N-acetylmannosamine and N-acetylneuraminate degradation	9.84E+02	1.33E+03	1.17E+03	4.63E+02	4.78E+02	3.73E+02
HSERMETANA-PWY: L-methionine biosynthesis III	6.99E+03	6.56E+03	6.33E+03	1.03E+03	1.15E+03	1.15E+03
KETOGLUCONMET-PWY: ketogluconate metabolism	2.51E+02	2.79E+02	3.08E+02	1.87E+02	1.84E+02	2.53E+02
OANTIGEN-PWY: O-antigen building blocks biosynthesis (E. coli)	2.13E+03	2.39E+03	2.68E+03	6.07E+02	6.59E+02	5.73E+02
P461-PWY: hexitol fermentation to lactate, formate, ethanol and acetate	4.53E+02	6.32E+02	4.97E+02	4.25E+02	4.45E+02	5.60E+02
P4-PWY: superpathway of L-lysine, L-threonine and L-methionine biosynthesis I	2.34E+03	2.58E+03	2.73E+03	5.72E+02	5.68E+02	5.72E+02
PANTOSYN-PWY: pantothenate and coenzyme A biosynthesis I	4.47E+03	4.24E+03	4.59E+03	4.89E+02	6.56E+02	6.62E+02
PENTOSE-P-PWY: pentose phosphate pathway	2.87E+03	3.58E+03	2.82E+03	1.42E+03	1.40E+03	1.37E+03
POLYISOPRENSYN-PWY: polyisoprenoid biosynthesis (E. coli)	3.08E+03	2.42E+03	2.75E+03	8.13E+02	8.60E+02	7.32E+02
PWY-3001: superpathway of L-isoleucine biosynthesis I	6.04E+03	5.85E+03	5.79E+03	4.62E+02	4.48E+02	4.51E+02
PWY-4242: pantothenate and coenzyme A biosynthesis III	4.12E+03	3.92E+03	4.33E+03	4.34E+02	6.69E+02	6.65E+02
PWY-5676: acetyl-CoA fermentation to butanoate II	1.16E+03	1.40E+03	1.49E+03	3.58E+02	3.83E+02	4.32E+02
PWY-5705: allantoin degradation to glyoxylate III	4.99E+01	3.52E+01	6.96E+01	5.12E+01	4.82E+01	1.15E+02

MetaCyc Metabolic Category	Group Mean of Relative Abundance for Statistically Significant Pathways (P value <.05)			Standard Deviation		
PWY-5845: superpathway of menaquinol-9 biosynthesis	1.87E+02	2.07E+02	2.83E+02	1.85E+02	1.70E+02	2.88E+02
PWY-5850: superpathway of menaquinol-6 biosynthesis I	1.47E+02	1.67E+02	2.53E+02	1.88E+02	1.76E+02	2.81E+02
PWY-5860: superpathway of demethylmena-quinol-6 biosynthesis I	1.01E+02	1.15E+02	1.77E+02	1.33E+02	1.25E+02	2.09E+02
PWY-5862: superpathway of demethylmena-quinol-9 biosynthesis	1.28E+02	1.42E+02	1.97E+02	1.31E+02	1.21E+02	2.14E+02
PWY-5896: superpathway of menaquinol-10 biosynthesis	1.48E+02	1.71E+02	2.34E+02	1.89E+02	1.75E+02	2.81E+02
PWY-5918: superpathay of heme biosynthesis from glutamate	4.46E+02	5.02E+02	6.23E+02	1.66E+02	2.02E+02	2.38E+02
PWY-6125: superpathway of guanosine nucleotides de novo biosynthesis II	3.80E+03	4.19E+03	4.30E+03	9.29E+02	1.04E+03	9.20E+02
PWY-6165: chorismate biosynthesis II (archaea)	2.56E+01	2.27E+01	4.21E+01	1.85E+01	2.88E+01	2.95E+01
PWY-6270: isoprene biosynthesis I	5.44E+03	4.57E+03	4.92E+03	8.10E+02	1.04E+03	8.11E+02
PWY-6358: superpathway of D-myo-inositol (1,4,5)-trisphosphate metabolism	6.46E-02	6.08E-02	2.52E-03	1.29E-01	1.82E-01	8.37E-02
PWY-6629: superpathway of L-tryptophan biosynthesis	4.05E+02	4.44E+02	5.82E+02	2.01E+02	2.11E+02	2.69E+02
PWY-6901: superpathway of glucose and xylose degradation	3.13E+03	3.88E+03	3.22E+03	1.28E+03	1.29E+03	1.24E+03
PWY-7200: superpathway of pyrimidine deoxyribonucleoside salvage	3.76E+02	3.61E+02	4.98E+02	4.51E+02	3.72E+02	4.28E+02
PWY-7237: myo-, chiro- and scillo-inositol degradation	3.52E+03	4.09E+03	3.02E+03	1.91E+03	1.79E+03	1.88E+03
PWY-7277: sphingolipid biosynthesis (mammals)	0.00E+00	2.01E-01	7.95E-01	0.00E+00	4.90E-01	5.34E-01
PWY-7389: superpathway of anaerobic energy metabolism (invertebrates)	3.64E+02	4.72E+02	5.50E+02	2.40E+02	2.50E+02	2.30E+02
PWY-7392: taxadiene biosynthesis (engineered)	4.21E+03	3.43E+03	3.72E+03	7.55E+02	9.15E+02	7.39E+02
THISYNARA-PWY: superpathway of thiamin diphosphate biosynthesis III (eukaryotes)	5.27E+03	4.82E+03	4.67E+03	1.00E+03	9.97E+02	9.02E+02
THISYN-PWY: superpathway of thiamin diphosphate biosynthesis I	5.15E+03	4.58E+03	4.58E+03	9.44E+02	8.68E+02	8.47E+02
THREOCAT-PWY: superpathway of L-threonine metabolism	8.66E+00	3.11E+01	2.54E+01	1.40E+01	1.28E+01	4.87E+01

Supplementary Table 3.8

Athlete						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isobutyric	PWY-6147: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	5.72E-03	6.53E-04	1.71E-02	1.36E-01	6.53E-01
Isovaleric	PWY-5154: NO_NAME|g_Flavonifractor.s_Flavonifractor_plautii	3.87E-04	8.50E-04	1.16E-03	2.67E-01	7.25E-01
Valeric	PWY-7220: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	7.33E-03	1.48E-03	1.47E-02	1.22E-01	6.45E-01
Valeric	PWY-7222: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	7.33E-03	1.48E-03	1.47E-02	1.22E-01	6.45E-01
Isobutyric	PWY-3841: NO_NAME|g_Coprococcus.s_Coprococcus_catus	7.01E-03	1.51E-03	2.28E-02	1.25E-01	6.47E-01
Valeric	PWY-6386: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	6.13E-03	2.01E-03	1.23E-02	1.32E-01	6.51E-01
Valeric	PWY-7221: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	6.43E-03	3.63E-03	1.29E-02	1.29E-01	6.50E-01
Isovaleric	PWY-7400: NO_NAME|g_Flavonifractor.s_Flavonifractor_plautii	9.67E-04	3.68E-03	2.90E-03	2.26E-01	7.03E-01
Isobutyric	FERMENTATION-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	1.23E-04	4.50E-03	3.69E-04	3.14E-01	7.49E-01
Isobutyric	PWY-3841: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	1.60E-04	4.50E-03	4.80E-04	3.04E-01	7.43E-01
Valeric	PWY-5100: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	6.57E-03	4.70E-03	1.31E-02	1.28E-01	6.49E-01
Valeric	PWY-6163: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.86E-03	5.38E-03	1.17E-02	1.34E-01	6.53E-01
Valeric	PYRIDNUCSYN-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	6.35E-03	5.50E-03	1.27E-02	1.30E-01	6.50E-01
Valeric	RHAMCAT-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	7.04E-03	5.50E-03	1.41E-02	1.24E-01	6.47E-01
Isobutyric	PWY-6313: NO_NAME|g_Coprococcus.s_Coprococcus_catus	4.38E-03	5.67E-03	1.31E-02	1.50E-01	6.62E-01
Valeric	PWY-6385: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	4.93E-03	5.77E-03	9.87E-03	1.44E-01	6.58E-01
Valeric	PWY-2942: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.35E-03	5.77E-03	1.07E-02	1.39E-01	6.55E-01
Valeric	PEPTIDOGLYCANSYN-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.77E-03	5.77E-03	1.15E-02	1.35E-01	6.53E-01
Valeric	PANTO-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.83E-03	5.77E-03	1.17E-02	1.35E-01	6.53E-01

Athlete						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Valeric	DTDPRHAMSYN-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	9.72E-03	5.77E-03	1.94E-02	1.06E-01	6.36E-01
Valeric	GLUCONEO-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.10E-03	6.04E-03	1.02E-02	1.42E-01	6.57E-01
Valeric	PWY66-400: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.32E-03	6.04E-03	1.06E-02	1.40E-01	6.56E-01
Valeric	PWY-7229: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	7.62E-03	6.04E-03	1.52E-02	1.20E-01	6.44E-01
Valeric	PWY-6126: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	8.04E-03	6.04E-03	1.61E-02	1.17E-01	6.42E-01
Valeric	DAPLYSINESYN-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.24E-03	6.31E-03	1.05E-02	1.41E-01	6.56E-01
Valeric	PWY-6122: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	4.67E-03	6.85E-03	9.35E-03	1.47E-01	6.60E-01
Valeric	PWY-6277: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	4.67E-03	6.85E-03	9.35E-03	1.47E-01	6.60E-01
Valeric	PWY-6121: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.20E-03	6.85E-03	1.04E-02	1.41E-01	6.56E-01
Valeric	ANAGLYCOLYSIS-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.89E-03	7.12E-03	1.18E-02	1.34E-01	6.52E-01
Valeric	PWY-1042: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.60E-03	7.38E-03	1.12E-02	1.37E-01	6.54E-01
Valeric	PWY66-422: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.65E-03	7.38E-03	1.13E-02	1.37E-01	6.54E-01
Valeric	PWY-5505: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.67E-03	7.38E-03	1.13E-02	1.36E-01	6.54E-01
Valeric	PWY-7219: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.80E-03	7.38E-03	1.16E-02	1.35E-01	6.53E-01
Valeric	PWY-5686: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	6.16E-03	7.38E-03	1.23E-02	1.32E-01	6.51E-01
Valeric	PWY-7663: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	6.27E-03	7.38E-03	1.25E-02	1.31E-01	6.50E-01
Valeric	PWY-7664: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	6.40E-03	7.65E-03	1.28E-02	1.30E-01	6.50E-01
Valeric	PWY-6282: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	6.53E-03	7.65E-03	1.31E-02	1.28E-01	6.49E-01
Butyric	PWY-7039: NO_NAME	7.73E-04	7.90E-03	4.64E-03	2.36E-01	7.09E-01
Isobutyric	PWY-724: NO_NAME|g_Flavonifractor.s_Flavonifractor_plautii	1.00E-02	8.46E-03	3.01E-02	1.04E-01	6.34E-01

Athlete						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Valeric	PWY-5659: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.05E-03	8.73E-03	1.01E-02	1.43E-01	6.57E-01
Valeric	PWYG-321: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	6.02E-03	9.26E-03	1.20E-02	1.33E-01	6.52E-01
Valeric	PWY-6387: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	6.38E-03	9.26E-03	1.28E-02	1.30E-01	6.50E-01
Valeric	FASYN-ELONG-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	6.49E-03	9.26E-03	1.30E-02	1.29E-01	6.49E-01
Isobutyric	PWY-6630: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	6.47E-06	9.69E-03	1.94E-05	4.20E-01	7.98E-01
Isobutyric	COMPLETE-ARO-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	7.88E-06	9.69E-03	2.36E-05	4.14E-01	7.95E-01
Isobutyric	ARO-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	8.09E-06	9.69E-03	2.43E-05	4.13E-01	7.94E-01
Isobutyric	PWY-6163: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	8.44E-06	9.69E-03	2.53E-05	4.11E-01	7.94E-01
Isovaleric	PWY-6630: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	5.48E-07	9.69E-03	3.29E-06	4.94E-01	8.29E-01
Isovaleric	COMPLETE-ARO-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	6.97E-07	9.69E-03	4.18E-06	4.87E-01	8.27E-01
Isovaleric	ARO-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	7.21E-07	9.69E-03	4.32E-06	4.87E-01	8.26E-01
Isovaleric	PWY-6163: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	7.58E-07	9.69E-03	4.55E-06	4.85E-01	8.26E-01
Valeric	SER-GLYSYN-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	6.09E-03	9.80E-03	1.22E-02	1.32E-01	6.51E-01
Valeric	BIOTIN-BIOSYNTHESIS-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.05E-03	1.01E-02	1.01E-02	1.43E-01	6.57E-01
Proprionic	LACTOSECAT-PWY: NO_NAME|g_Streptococcus.s_Streptococcus_parasanguinis	2.96E-03	1.03E-02	1.78E-02	1.71E-01	6.73E-01
Valeric	GLUTORN-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	4.50E-03	1.03E-02	9.01E-03	1.49E-01	6.61E-01
Proprionic	COA-PWY: NO_NAME|g_Streptococcus.s_Streptococcus_parasanguinis	7.96E-03	1.09E-02	4.78E-02	1.17E-01	6.42E-01
Valeric	PWY-6519: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.53E-03	1.09E-02	1.11E-02	1.38E-01	6.54E-01
Valeric	PWY-6936: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	7.69E-03	1.09E-02	1.54E-02	1.19E-01	6.44E-01
Isobutyric	HOMOSER-METSYN-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	4.40E-03	1.10E-02	1.32E-02	1.50E-01	6.62E-01

Athlete						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isobutyric	HSERMETANA-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	4.83E-03	1.10E-02	1.45E-02	1.45E-01	6.59E-01
Isobutyric	PWY-7219: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	4.84E-03	1.10E-02	1.45E-02	1.45E-01	6.59E-01
Isobutyric	PWY-6897: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	6.44E-03	1.10E-02	1.93E-02	1.29E-01	6.49E-01
Isobutyric	PWY-6123: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	1.02E-02	1.10E-02	3.06E-02	1.03E-01	6.34E-01
Valeric	PWY-4984: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.02E-03	1.11E-02	1.00E-02	1.43E-01	6.58E-01
Valeric	CITRULBIO-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.12E-03	1.11E-02	1.02E-02	1.42E-01	6.57E-01
Valeric	ARGSYN-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.32E-03	1.17E-02	1.06E-02	1.40E-01	6.56E-01
Valeric	PWY-7400: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.34E-03	1.17E-02	1.07E-02	1.40E-01	6.56E-01
Valeric	PWY-5097: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	6.01E-03	1.17E-02	1.20E-02	1.33E-01	6.52E-01
Valeric	PWY-7388: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	6.12E-03	1.17E-02	1.22E-02	1.32E-01	6.51E-01
Valeric	PWY0-862: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	6.28E-03	1.17E-02	1.26E-02	1.31E-01	6.50E-01
Proprionic	PWY-7187: NO_NAME|g_Streptococcus.s_Streptococcus_parasanguinis	2.65E-03	1.27E-02	1.59E-02	1.77E-01	6.77E-01
Proprionic	PWY0-166: NO_NAME|g_Streptococcus.s_Streptococcus_parasanguinis	2.68E-03	1.27E-02	1.61E-02	1.76E-01	6.76E-01
Proprionic	PWY-7198: NO_NAME|g_Streptococcus.s_Streptococcus_parasanguinis	3.38E-03	1.27E-02	2.03E-02	1.64E-01	6.70E-01
Isobutyric	PWY0-1296: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	2.11E-04	1.32E-02	6.32E-04	2.93E-01	7.38E-01
Isobutyric	PEPTIDOGLYCANSYN-PWY: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	2.16E-04	1.32E-02	6.47E-04	2.92E-01	7.37E-01
Isobutyric	PWY0-1297: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	2.85E-04	1.32E-02	8.54E-04	2.80E-01	7.32E-01
Isobutyric	PWY-6385: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	2.89E-04	1.32E-02	8.67E-04	2.79E-01	7.31E-01
Isobutyric	PWY-6387: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	6.58E-04	1.32E-02	1.97E-03	2.43E-01	7.13E-01
Isobutyric	PWY-3001: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	2.19E-03	1.32E-02	6.56E-03	1.86E-01	6.82E-01

Athlete						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isobutyric	GLUCONEO-PWY: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	3.82E-03	1.32E-02	1.15E-02	1.58E-01	6.66E-01
Isobutyric	PWY-6317: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	7.76E-03	1.32E-02	2.33E-02	1.19E-01	6.43E-01
Valeric	PWY-5989: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	6.51E-03	1.33E-02	1.30E-02	1.29E-01	6.49E-01
Valeric	P108-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.23E-03	1.36E-02	1.05E-02	1.41E-01	6.56E-01
Isobutyric	PWY-5384: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	5.94E-05	1.46E-02	1.78E-04	3.43E-01	7.62E-01
Isobutyric	PWY-6151: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	1.01E-04	1.46E-02	3.03E-04	3.22E-01	7.52E-01
Isobutyric	PWY-5103: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	1.59E-04	1.46E-02	4.77E-04	3.04E-01	7.44E-01
Isobutyric	BRANCHED-CHAIN-AA-SYN-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	2.00E-04	1.46E-02	6.01E-04	2.95E-01	7.39E-01
Isobutyric	PWY-5686: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	3.43E-04	1.46E-02	1.03E-03	2.72E-01	7.28E-01
Isobutyric	ILEUSYN-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	3.60E-04	1.46E-02	1.08E-03	2.70E-01	7.27E-01
Isobutyric	PWY-3001: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	5.54E-04	1.46E-02	1.66E-03	2.51E-01	7.17E-01
Isobutyric	PWY-5100: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	7.45E-04	1.46E-02	2.23E-03	2.38E-01	7.10E-01
Isobutyric	SER-GLYSYN-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	9.04E-04	1.46E-02	2.71E-03	2.29E-01	7.05E-01
Isobutyric	P124-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	9.57E-04	1.46E-02	2.87E-03	2.26E-01	7.04E-01
Isobutyric	P161-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	9.64E-04	1.46E-02	2.89E-03	2.26E-01	7.04E-01
Isobutyric	THRESYN-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	1.70E-03	1.46E-02	5.09E-03	1.99E-01	6.89E-01
Isobutyric	PWY-5097: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	2.46E-03	1.46E-02	7.39E-03	1.80E-01	6.79E-01
Isobutyric	TCA: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	2.60E-03	1.46E-02	7.79E-03	1.78E-01	6.77E-01
Isobutyric	PWY0-1061: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	3.09E-03	1.46E-02	9.26E-03	1.69E-01	6.72E-01
Isobutyric	PWY-621: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	3.96E-03	1.46E-02	1.19E-02	1.56E-01	6.65E-01

Athlete						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isobutyric	PWY-7237: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	4.09E-03	1.46E-02	1.23E-02	1.54E-01	6.64E-01
Isobutyric	POLYISOPRENSYN-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	8.86E-03	1.46E-02	2.66E-02	1.11E-01	6.39E-01
Valeric	PWY-7237: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.98E-03	1.46E-02	1.20E-02	1.33E-01	6.52E-01
Valeric	ARO-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.75E-03	1.47E-02	1.15E-02	1.36E-01	6.53E-01
Valeric	COMPLETE-ARO-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.80E-03	1.47E-02	1.16E-02	1.35E-01	6.53E-01
Valeric	PANTOSYN-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.83E-03	1.53E-02	1.17E-02	1.35E-01	6.53E-01
Valeric	COA-PWY-1: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	4.48E-03	1.54E-02	8.95E-03	1.49E-01	6.61E-01
Valeric	PWY-7111: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.72E-03	1.57E-02	1.14E-02	1.36E-01	6.53E-01
Valeric	PWY-5667: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.75E-03	1.57E-02	1.15E-02	1.35E-01	6.53E-01
Valeric	PWY0-1319: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	5.75E-03	1.57E-02	1.15E-02	1.35E-01	6.53E-01
Valeric	PWY-5973: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	6.01E-03	1.79E-02	1.20E-02	1.33E-01	6.52E-01
Valeric	PWY-4242: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	4.07E-03	1.87E-02	8.15E-03	1.54E-01	6.64E-01
Valeric	COA-PWY: NO_NAME|g_Alistipes.s_Alistipes_sp_AP11	4.59E-03	1.89E-02	9.19E-03	1.48E-01	6.60E-01
Isobutyric	PWY-6385: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	4.51E-05	1.94E-02	1.35E-04	3.53E-01	7.67E-01
Isobutyric	PEPTIDOGLYCANSYN-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	4.67E-05	1.94E-02	1.40E-04	3.52E-01	7.66E-01
Isobutyric	PYRIDNUCSAL-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	6.14E-05	1.94E-02	1.84E-04	3.41E-01	7.62E-01
Isobutyric	PWY-6126: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	9.32E-05	1.94E-02	2.80E-04	3.25E-01	7.54E-01
Isobutyric	PWY-7229: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	1.24E-04	1.94E-02	3.72E-04	3.14E-01	7.48E-01
Isobutyric	PWY-6527: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	2.29E-04	1.94E-02	6.87E-04	2.89E-01	7.36E-01
Isobutyric	HISTSYN-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	2.79E-04	1.94E-02	8.36E-04	2.81E-01	7.32E-01

Athlete

SCFA	Pathway	p value	P correct BH	PPMC	LCI	UCI
Isovaleric	PWY-6385: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	6.07E-06	1.94E-02	3.64E-05	4.22E-01	7.99E-01
Isovaleric	PEPTIDOGLYCANSYN-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	6.33E-06	1.94E-02	3.80E-05	4.21E-01	7.98E-01
Isovaleric	PYRIDNUCSAL-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	8.87E-06	1.94E-02	5.32E-05	4.10E-01	7.93E-01
Isovaleric	PWY-6126: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	1.48E-05	1.94E-02	8.90E-05	3.92E-01	7.85E-01
Isovaleric	PWY-7229: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	2.11E-05	1.94E-02	1.26E-04	3.80E-01	7.80E-01
Isovaleric	PWY-6527: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	4.47E-05	1.94E-02	2.68E-04	3.53E-01	7.67E-01
Isovaleric	HISTSYN-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	5.68E-05	1.94E-02	3.41E-04	3.44E-01	7.63E-01
Valeric	PWY-6737: NO_NAME|g_Anaerotruncus.s_Anaerotruncus_colihominis	2.36E-02	1.94E-02	4.72E-02	5.16E-02	6.02E-01
Butyric	PWY-6122: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_3_1_46FAA	3.02E-02	2.15E-02	4.53E-02	3.54E-02	5.91E-01
Butyric	PWY-6277: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_3_1_46FAA	3.02E-02	2.15E-02	4.53E-02	3.54E-02	5.91E-01
Isobutyric	PWY66-422: NO_NAME|g_Coprococcus.s_Coprococcus_catus	4.98E-03	2.25E-02	2.99E-02	1.43E-01	6.58E-01
Isovaleric	PWY-724: NO_NAME|g_Flavonifractor.s_Flavonifractor_plautii	9.24E-03	2.34E-02	3.01E-02	1.09E-01	6.37E-01
Isovaleric	FERMENTATION-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	2.44E-05	2.41E-02	1.46E-04	3.75E-01	7.77E-01
Isovaleric	PWY-3841: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	3.80E-05	2.41E-02	2.28E-04	3.59E-01	7.70E-01
Isovaleric	PWY0-1296: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	4.29E-05	2.58E-02	2.58E-04	3.55E-01	7.68E-01
Isovaleric	PEPTIDOGLYCANSYN-PWY: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	4.72E-05	2.58E-02	2.83E-04	3.51E-01	7.66E-01
Isovaleric	PWY0-1297: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	6.32E-05	2.58E-02	3.79E-04	3.40E-01	7.61E-01
Isovaleric	PWY-6385: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	6.95E-05	2.58E-02	4.17E-04	3.37E-01	7.59E-01
Isovaleric	PWY-6387: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	1.86E-04	2.58E-02	1.11E-03	2.98E-01	7.40E-01
Isovaleric	PWY-3001: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	8.42E-04	2.58E-02	5.05E-03	2.32E-01	7.07E-01

Athlete						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isovaleric	GLUCONEO-PWY: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	1.65E-03	2.58E-02	9.92E-03	2.00E-01	6.90E-01
Isovaleric	PWY-6317: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	3.65E-03	2.58E-02	2.19E-02	1.60E-01	6.67E-01
Isobutyric	OANTIGEN-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_longum	1.00E-02	2.71E-02	3.00E-02	1.04E-01	6.35E-01
Isobutyric	PWY-5100: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	1.11E-02	2.81E-02	3.32E-02	9.83E-02	6.31E-01
Proprionic	PWY-7211: NO_NAME|g_Streptococcus.s_Streptococcus_parasanguinis	4.51E-03	2.91E-02	2.71E-02	1.49E-01	6.61E-01
Isovaleric	HOMOSER-METSYN-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	3.25E-03	2.99E-02	1.32E-02	1.66E-01	6.71E-01
Isovaleric	PWY-7219: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	3.49E-03	2.99E-02	1.45E-02	1.62E-01	6.69E-01
Isovaleric	HSERMETANA-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	3.71E-03	2.99E-02	1.45E-02	1.59E-01	6.67E-01
Isovaleric	PWY-6897: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	4.87E-03	2.99E-02	1.93E-02	1.45E-01	6.58E-01
Isovaleric	PWY-6123: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	8.29E-03	2.99E-02	3.06E-02	1.15E-01	6.41E-01
Isovaleric	PWY-5384: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	9.76E-06	3.39E-02	5.86E-05	4.07E-01	7.92E-01
Isovaleric	PWY-6151: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	3.24E-05	3.39E-02	1.94E-04	3.65E-01	7.73E-01
Isovaleric	PWY-5103: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	6.04E-05	3.39E-02	3.62E-04	3.42E-01	7.62E-01
Isovaleric	BRANCHED-CHAIN-AA-SYN-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	7.94E-05	3.39E-02	4.76E-04	3.32E-01	7.57E-01
Isovaleric	PWY-5686: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	1.35E-04	3.39E-02	8.13E-04	3.11E-01	7.47E-01
Isovaleric	ILEUSYN-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	1.41E-04	3.39E-02	8.45E-04	3.09E-01	7.46E-01
Isovaleric	PWY-3001: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	2.77E-04	3.39E-02	1.66E-03	2.81E-01	7.32E-01
Isovaleric	PWY-5100: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	4.07E-04	3.39E-02	2.23E-03	2.65E-01	7.24E-01
Isovaleric	SER-GLYSYN-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	4.29E-04	3.39E-02	2.57E-03	2.62E-01	7.23E-01
Isovaleric	P124-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	5.12E-04	3.39E-02	2.87E-03	2.55E-01	7.19E-01

Athlete						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isovaleric	P161-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	5.54E-04	3.39E-02	2.89E-03	2.51E-01	7.17E-01
Isovaleric	THRESYN-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	1.02E-03	3.39E-02	5.09E-03	2.23E-01	7.02E-01
Isovaleric	PWY-5097: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	1.53E-03	3.39E-02	7.39E-03	2.04E-01	6.92E-01
Isovaleric	TCA: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	1.67E-03	3.39E-02	7.79E-03	2.00E-01	6.89E-01
Isovaleric	PWY0-1061: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	1.84E-03	3.39E-02	9.26E-03	1.95E-01	6.87E-01
Isovaleric	PWY-621: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	2.49E-03	3.39E-02	1.19E-02	1.80E-01	6.78E-01
Isovaleric	PWY-7237: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	2.76E-03	3.39E-02	1.23E-02	1.74E-01	6.75E-01
Isovaleric	POLYISOPRENSYN-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_pseudocatenulatum	6.17E-03	3.39E-02	2.66E-02	1.32E-01	6.51E-01
Isobutyric	PWY-5188: NO_NAME|g_Coprococcus.s_Coprococcus_catus	4.05E-03	3.46E-02	1.22E-02	1.54E-01	6.64E-01
Proprionic	PWY-6386: NO_NAME|g_Streptococcus.s_Streptococcus_parasanguinis	3.02E-03	3.46E-02	1.81E-02	1.70E-01	6.73E-01
Isovaleric	ARGSYN-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_catus	1.63E-02	3.62E-02	4.62E-02	7.50E-02	6.17E-01
Butyric	PEPTIDOGLYCANSYN-PWY: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_3_1_46FAA	2.54E-02	3.72E-02	3.80E-02	4.69E-02	5.99E-01
Butyric	PWY-6387: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_3_1_46FAA	2.58E-02	3.72E-02	3.87E-02	4.57E-02	5.98E-01
Butyric	PWY-6385: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_3_1_46FAA	2.68E-02	3.72E-02	4.01E-02	4.34E-02	5.97E-01
Butyric	PWY-6386: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_3_1_46FAA	2.64E-02	3.89E-02	3.97E-02	4.42E-02	5.97E-01
Isovaleric	PWY-5100: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_1_1_57FAA	5.25E-03	3.92E-02	3.15E-02	1.41E-01	6.56E-01
Isovaleric	PWY-3841: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_7_1_58FAA	6.92E-03	4.08E-02	2.08E-02	1.25E-01	6.47E-01
Isovaleric	OANTIGEN-PWY: NO_NAME|g_Bifidobacterium.s_Bifidobacterium_longum	5.41E-03	4.15E-02	3.00E-02	1.39E-01	6.55E-01
Isobutyric	PWY-3841: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_7_1_58FAA	3.81E-03	4.16E-02	2.08E-02	1.58E-01	6.66E-01
Isobutyric	PWY-7199: NO_NAME|g_Flavonifractor.s_Flavonifractor_plautii	4.71E-03	4.89E-02	1.41E-02	1.46E-01	6.59E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Butyric	PWY0-881: NO_NAME	7.53E-05	4.52E-04	6.55E-01	4.80E-01	8.95E-01
Valeric	PWY-3841: NO_NAME|g_Blautia.s_Ruminococcus_gnavus	1.19E-04	7.15E-04	-6.80E-01	-8.89E-01	-4.56E-01
Isovaleric	SALVADEHYPOX-PWY: NO_NAME|unclassified	1.46E-04	8.74E-04	6.44E-01	4.45E-01	8.86E-01
Butyric	PWY0-881: NO_NAME|unclassified	1.51E-04	9.06E-04	4.22E-01	4.43E-01	8.85E-01
Isobutyric	SALVADEHYPOX-PWY: NO_NAME|unclassified	3.94E-04	1.18E-03	6.36E-01	3.87E-01	8.70E-01
Valeric	PWY-6124: NO_NAME|g_Blautia.s_Ruminococcus_gnavus	2.47E-04	1.48E-03	-6.84E-01	-8.78E-01	-4.15E-01
Isovaleric	PWY-5088: NO_NAME	2.79E-04	1.67E-03	6.47E-01	4.08E-01	8.76E-01
Butyric	POLYAMSYN-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	4.33E-04	2.60E-03	2.52E-01	3.81E-01	8.68E-01
Isovaleric	PWY-7219: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	4.67E-04	2.80E-03	6.07E-01	3.77E-01	8.67E-01
Butyric	PWY-6305: NO_NAME|g_Blautia.s_Ruminococcus_obeum	5.40E-04	3.24E-03	2.59E-01	3.68E-01	8.64E-01
Isovaleric	PWY-7229: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	5.53E-04	3.32E-03	6.30E-01	3.66E-01	8.64E-01
Isobutyric	PWY-7209: NO_NAME|unclassified	1.15E-03	3.46E-03	6.13E-01	3.18E-01	8.49E-01
Valeric	PWY-7209: NO_NAME|unclassified	8.09E-04	3.46E-03	6.70E-01	3.42E-01	8.56E-01
Isobutyric	PWY66-398: NO_NAME|unclassified	1.24E-03	3.71E-03	5.27E-01	3.13E-01	8.48E-01
Isovaleric	PWY66-398: NO_NAME|unclassified	8.57E-04	3.71E-03	5.32E-01	3.38E-01	8.55E-01
Isovaleric	PWY-5973: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	6.59E-04	3.95E-03	5.22E-01	3.55E-01	8.60E-01
Isovaleric	PWY-7209: NO_NAME|unclassified	2.05E-03	4.10E-03	6.16E-01	2.78E-01	8.36E-01
Isovaleric	GLUDEG-II-PWY: NO_NAME	6.91E-04	4.15E-03	6.51E-01	3.52E-01	8.60E-01
Isobutyric	URSIN-PWY: NO_NAME|unclassified	6.92E-04	4.15E-03	5.64E-01	3.52E-01	8.60E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isobutyric	PWY-5464: NO_NAME|unclassified	1.47E-03	4.40E-03	6.03E-01	3.01E-01	8.44E-01
Isovaleric	PWY-5464: NO_NAME|unclassified	1.22E-03	4.40E-03	5.84E-01	3.14E-01	8.48E-01
Isovaleric	PWY-5690: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	7.65E-04	4.59E-03	6.10E-01	3.45E-01	8.58E-01
Isobutyric	PWY-5088: NO_NAME	1.55E-03	4.65E-03	6.26E-01	2.98E-01	8.43E-01
Isobutyric	PWY-6595: NO_NAME|unclassified	1.73E-03	5.19E-03	6.17E-01	2.90E-01	8.40E-01
Isovaleric	PWY-6595: NO_NAME|unclassified	1.46E-03	5.19E-03	6.35E-01	3.02E-01	8.44E-01
Isovaleric	P108-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	9.24E-04	5.54E-03	5.77E-01	3.33E-01	8.54E-01
Isovaleric	PWY-7221: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	9.44E-04	5.67E-03	5.39E-01	3.31E-01	8.53E-01
Isovaleric	TCA: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	9.62E-04	5.77E-03	6.10E-01	3.30E-01	8.53E-01
Acetic	PWY-7237: NO_NAME|g_Bacteroides.s_Bacteroides_xylanisolvens	9.64E-04	5.79E-03	7.14E-01	3.30E-01	8.53E-01
Isovaleric	PWY-7400: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	9.93E-04	5.96E-03	5.93E-01	3.28E-01	8.52E-01
Isovaleric	ARGSYN-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	9.95E-04	5.97E-03	5.93E-01	3.28E-01	8.52E-01
Isovaleric	PWY-6122: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.00E-03	6.02E-03	5.25E-01	3.27E-01	8.52E-01
Isovaleric	PWY-6277: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.00E-03	6.02E-03	5.25E-01	3.27E-01	8.52E-01
Isobutyric	PWY-6606: NO_NAME|unclassified	2.02E-03	6.05E-03	5.69E-01	2.79E-01	8.37E-01
Isovaleric	PWY-6606: NO_NAME|unclassified	1.43E-03	6.05E-03	5.83E-01	3.03E-01	8.45E-01
Isovaleric	PWYG-321: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.02E-03	6.14E-03	5.22E-01	3.26E-01	8.52E-01
Isovaleric	PWY-7388: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.03E-03	6.21E-03	5.20E-01	3.25E-01	8.51E-01
Isovaleric	ARGSYNBSUB-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.03E-03	6.21E-03	5.93E-01	3.25E-01	8.51E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isobutyric	ILEUDEG-PWY: NO_NAME	1.31E-03	6.42E-03	5.49E-01	3.09E-01	8.46E-01
Isovaleric	ILEUDEG-PWY: NO_NAME	2.14E-03	6.42E-03	5.38E-01	2.75E-01	8.35E-01
Isovaleric	PWY-7663: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.09E-03	6.51E-03	5.22E-01	3.22E-01	8.50E-01
Isobutyric	PWY-7229: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	2.17E-03	6.52E-03	6.32E-01	2.74E-01	8.35E-01
Isovaleric	PWY-6121: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.09E-03	6.54E-03	5.39E-01	3.22E-01	8.50E-01
Isobutyric	PWY-7219: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	2.21E-03	6.64E-03	6.00E-01	2.72E-01	8.35E-01
Valeric	FASYN-INITIAL-PWY: NO_NAME|g_Eubacterium.s_Eubacterium_hallii	1.11E-03	6.68E-03	-5.82E-01	-8.50E-01	-3.20E-01
Isovaleric	PWY66-399: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.17E-03	7.00E-03	5.93E-01	3.17E-01	8.49E-01
Isovaleric	PWY-7664: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.21E-03	7.29E-03	5.22E-01	3.14E-01	8.48E-01
Butyric	FUCCAT-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	1.22E-03	7.34E-03	4.91E-01	3.14E-01	8.48E-01
Isovaleric	GLUTORN-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.25E-03	7.48E-03	6.04E-01	3.13E-01	8.48E-01
Isobutyric	P261-PWY: NO_NAME	1.82E-03	7.51E-03	6.25E-01	2.86E-01	8.39E-01
Isovaleric	P261-PWY: NO_NAME	2.50E-03	7.51E-03	5.73E-01	2.63E-01	8.32E-01
Isovaleric	PWY0-862: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.26E-03	7.55E-03	5.22E-01	3.12E-01	8.47E-01
Acetic	PWY-6993: NO_NAME	1.26E-03	7.56E-03	5.26E-01	3.12E-01	8.47E-01
Acetic	PWY-6993: NO_NAME|unclassified	1.27E-03	7.59E-03	5.26E-01	3.12E-01	8.47E-01
Isovaleric	PWY-7208: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.27E-03	7.61E-03	5.21E-01	3.12E-01	8.47E-01
Isovaleric	PWY-3781: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.27E-03	7.62E-03	5.09E-01	3.11E-01	8.47E-01
Isovaleric	FASYN-INITIAL-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.28E-03	7.66E-03	5.22E-01	3.11E-01	8.47E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isovaleric	PWY-6282: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.29E-03	7.75E-03	5.22E-01	3.10E-01	8.47E-01
Isovaleric	PWY-7383: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.33E-03	7.98E-03	5.93E-01	3.08E-01	8.46E-01
Butyric	PWY-7237: NO_NAME|g_Bacteroides.s_Bacteroides_xylanisolvens	2.69E-03	8.08E-03	6.14E-01	2.58E-01	8.30E-01
Isobutyric	PWY-5464: NO_NAME	3.01E-03	8.30E-03	5.95E-01	2.50E-01	8.27E-01
Isovaleric	PWY-5464: NO_NAME	2.55E-03	8.30E-03	5.58E-01	2.62E-01	8.31E-01
Valeric	PWY-5464: NO_NAME	4.15E-03	8.30E-03	6.14E-01	2.25E-01	8.19E-01
Isovaleric	FASYN-ELONG-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.38E-03	8.30E-03	5.15E-01	3.06E-01	8.45E-01
Isovaleric	PWY-181: NO_NAME	1.39E-03	8.33E-03	5.01E-01	3.05E-01	8.45E-01
Acetic	PWY-6823: NO_NAME	2.78E-03	8.34E-03	6.01E-01	2.55E-01	8.29E-01
Butyric	PWY-6823: NO_NAME	2.68E-03	8.34E-03	5.97E-01	2.58E-01	8.30E-01
Isovaleric	PWY-5505: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.40E-03	8.40E-03	5.27E-01	3.05E-01	8.45E-01
Isobutyric	ILEUDEG-PWY: NO_NAME|unclassified	1.74E-03	8.42E-03	5.00E-01	2.89E-01	8.40E-01
Isovaleric	ILEUDEG-PWY: NO_NAME|unclassified	2.81E-03	8.42E-03	4.91E-01	2.55E-01	8.29E-01
Isovaleric	PWY-4984: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.41E-03	8.46E-03	5.23E-01	3.04E-01	8.45E-01
Butyric	PWY-4242: NO_NAME|g_Streptococcus.s_Streptococcus_salivarius	1.43E-03	8.56E-03	3.69E-01	3.03E-01	8.45E-01
Butyric	PWY-7251: NO_NAME	1.43E-03	8.56E-03	3.69E-01	3.03E-01	8.45E-01
Butyric	PWY-7251: NO_NAME|unclassified	1.43E-03	8.56E-03	3.69E-01	3.03E-01	8.45E-01
Butyric	PWY66-388: NO_NAME	1.43E-03	8.56E-03	3.69E-01	3.03E-01	8.45E-01
Butyric	PWY66-388: NO_NAME|unclassified	1.43E-03	8.56E-03	3.69E-01	3.03E-01	8.45E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Butyric	PWY-6109: NO_NAME	1.43E-03	8.56E-03	3.69E-01	3.03E-01	8.45E-01
Butyric	PWY-6109: NO_NAME|unclassified	1.43E-03	8.56E-03	3.69E-01	3.03E-01	8.45E-01
Valeric	PWY-5109: NO_NAME	1.43E-03	8.57E-03	5.95E-01	3.03E-01	8.45E-01
Isobutyric	PWY-6121: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	2.91E-03	8.74E-03	5.33E-01	2.52E-01	8.28E-01
Butyric	NONOXIPENT-PWY: NO_NAME|g_Streptococcus.s_Streptococcus_salivarius	1.48E-03	8.89E-03	4.67E-01	3.01E-01	8.44E-01
Acetic	PWY-5667: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	1.48E-03	8.89E-03	5.01E-01	3.01E-01	8.44E-01
Acetic	PWY0-1319: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	1.48E-03	8.89E-03	5.01E-01	3.01E-01	8.44E-01
Valeric	ILEUDEG-PWY: NO_NAME|unclassified	4.49E-03	8.99E-03	6.31E-01	2.19E-01	8.17E-01
Isobutyric	PWY-6309: NO_NAME|unclassified	3.05E-03	9.16E-03	5.25E-01	2.48E-01	8.27E-01
Isovaleric	PWY-6309: NO_NAME|unclassified	2.17E-03	9.16E-03	4.96E-01	2.74E-01	8.35E-01
Valeric	ILEUDEG-PWY: NO_NAME	4.62E-03	9.24E-03	6.31E-01	2.17E-01	8.16E-01
Isobutyric	PWY-5973: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.14E-03	9.41E-03	5.08E-01	2.46E-01	8.26E-01
Butyric	PWY-7219: NO_NAME|g_Blautia.s_Ruminococcus_obeum	1.60E-03	9.60E-03	2.99E-01	2.95E-01	8.42E-01
Butyric	HOMOSER-METSYN-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	1.62E-03	9.74E-03	1.87E-01	2.94E-01	8.42E-01
Isobutyric	TCA: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.25E-03	9.75E-03	6.03E-01	2.44E-01	8.25E-01
Butyric	PWY-5189: NO_NAME|g_Blautia.s_Ruminococcus_obeum	1.64E-03	9.87E-03	2.99E-01	2.93E-01	8.41E-01
Butyric	PWY-2723: NO_NAME|unclassified	1.66E-03	9.99E-03	5.54E-01	2.93E-01	8.41E-01
Isobutyric	PWY-5690: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.35E-03	1.01E-02	6.03E-01	2.41E-01	8.24E-01
Isovaleric	ILEUSYN-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.71E-03	1.03E-02	5.09E-01	2.91E-01	8.41E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isovaleric	VALSYN-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.71E-03	1.03E-02	5.09E-01	2.91E-01	8.41E-01
Isovaleric	ANAGLYCOLYSIS-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.71E-03	1.03E-02	5.93E-01	2.90E-01	8.41E-01
Proprionic	PWY-6344: NO_NAME	1.72E-03	1.03E-02	6.02E-01	2.90E-01	8.40E-01
Valeric	FOLSYN-PWY: NO_NAME|unclassified	1.72E-03	1.03E-02	6.69E-01	2.90E-01	8.40E-01
Isovaleric	PWY-7111: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.76E-03	1.06E-02	5.12E-01	2.88E-01	8.40E-01
Isovaleric	PWY-7003: NO_NAME|unclassified	1.77E-03	1.06E-02	5.19E-01	2.88E-01	8.40E-01
Isovaleric	BRANCHED-CHAIN-AA-SYN-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.77E-03	1.06E-02	5.04E-01	2.88E-01	8.40E-01
Isobutyric	PWY-7221: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.61E-03	1.08E-02	5.32E-01	2.36E-01	8.22E-01
Isobutyric	PWY-6122: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.61E-03	1.08E-02	5.23E-01	2.36E-01	8.22E-01
Isobutyric	PWY-6277: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.61E-03	1.08E-02	5.23E-01	2.36E-01	8.22E-01
Isobutyric	PWY-7388: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.61E-03	1.08E-02	5.05E-01	2.36E-01	8.22E-01
Butyric	PWY-5973: NO_NAME|g_Veillonella.s_Veillonella_parvula	1.82E-03	1.09E-02	5.36E-01	2.86E-01	8.39E-01
Isovaleric	PWY-3841: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.83E-03	1.10E-02	5.86E-01	2.86E-01	8.39E-01
Acetic	PWY-5677: NO_NAME	2.57E-03	1.11E-02	4.81E-01	2.61E-01	8.31E-01
Valeric	PWY-5677: NO_NAME	3.70E-03	1.11E-02	5.92E-01	2.34E-01	8.22E-01
Butyric	GLUTORN-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	1.86E-03	1.12E-02	3.48E-01	2.85E-01	8.39E-01
Isobutyric	GLUTORN-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.73E-03	1.12E-02	5.99E-01	2.33E-01	8.22E-01
Isobutyric	PWY-7663: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.77E-03	1.13E-02	5.08E-01	2.32E-01	8.21E-01
Isobutyric	PWY-7242: NO_NAME|unclassified	3.77E-03	1.13E-02	4.62E-01	2.32E-01	8.21E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isovaleric	PWY-7242: NO_NAME|unclassified	2.24E-03	1.13E-02	4.22E-01	2.71E-01	8.34E-01
Butyric	MET-SAM-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	1.89E-03	1.14E-02	2.47E-01	2.83E-01	8.38E-01
Acetic	PWY-7414: NO_NAME|unclassified	1.92E-03	1.15E-02	6.18E-01	2.83E-01	8.38E-01
Isovaleric	PWY66-400: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.93E-03	1.16E-02	5.77E-01	2.82E-01	8.38E-01
Isobutyric	PWY-5081: NO_NAME|unclassified	5.04E-03	1.17E-02	4.52E-01	2.10E-01	8.13E-01
Isovaleric	PWY-5081: NO_NAME|unclassified	2.85E-03	1.17E-02	4.39E-01	2.53E-01	8.28E-01
Valeric	PWY-5081: NO_NAME|unclassified	5.83E-03	1.17E-02	5.53E-01	1.98E-01	8.09E-01
Isobutyric	PWY-7383: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.90E-03	1.17E-02	5.83E-01	2.30E-01	8.20E-01
Isovaleric	PWY-5179: NO_NAME|unclassified	1.96E-03	1.17E-02	3.59E-01	2.81E-01	8.37E-01
Acetic	PWY0-881: NO_NAME|unclassified	3.93E-03	1.18E-02	5.44E-01	2.29E-01	8.20E-01
Isovaleric	GLUCONEO-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.96E-03	1.18E-02	5.86E-01	2.81E-01	8.37E-01
Isovaleric	PWY-1042: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.97E-03	1.18E-02	5.77E-01	2.81E-01	8.37E-01
Isobutyric	PWYG-321: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.95E-03	1.19E-02	5.08E-01	2.29E-01	8.20E-01
Isobutyric	ARGSYNBSUB-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.95E-03	1.19E-02	5.83E-01	2.29E-01	8.20E-01
Isobutyric	PWY-6470: NO_NAME|unclassified	3.99E-03	1.20E-02	4.73E-01	2.28E-01	8.20E-01
Valeric	PWY-6470: NO_NAME|unclassified	3.70E-03	1.20E-02	7.66E-01	2.34E-01	8.22E-01
Isovaleric	PWY-5103: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	2.00E-03	1.20E-02	5.04E-01	2.80E-01	8.37E-01
Butyric	PWY-6823: NO_NAME|unclassified	2.00E-03	1.20E-02	5.35E-01	2.80E-01	8.37E-01
Isovaleric	PWY-5994: NO_NAME|unclassified	2.00E-03	1.20E-02	5.08E-01	2.79E-01	8.37E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Butyric	PWY-5188: NO_NAME|g_Blautia.s_Ruminococcus_obeum	2.00E-03	1.20E-02	1.58E-01	2.79E-01	8.37E-01
Valeric	PWY-6163: NO_NAME|g_Eubacterium.s_Eubacterium_hallii	2.02E-03	1.21E-02	-4.79E-01	-8.37E-01	-2.79E-01
Isobutyric	PWY66-399: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	4.08E-03	1.22E-02	5.83E-01	2.26E-01	8.19E-01
Isobutyric	FASYN-INITIAL-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	4.08E-03	1.22E-02	5.13E-01	2.26E-01	8.19E-01
Isobutyric	PWY-7400: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	4.10E-03	1.23E-02	5.83E-01	2.26E-01	8.19E-01
Isobutyric	ARGSYN-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	4.11E-03	1.23E-02	5.83E-01	2.26E-01	8.19E-01
Acetic	PWY-7414: NO_NAME	2.06E-03	1.24E-02	5.91E-01	2.77E-01	8.36E-01
Acetic	PWY-5304: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	2.08E-03	1.25E-02	5.01E-01	2.77E-01	8.36E-01
Isovaleric	PWY-6470: NO_NAME|unclassified	6.26E-03	1.25E-02	4.40E-01	1.92E-01	8.07E-01
Butyric	PWY-7221: NO_NAME|g_Blautia.s_Ruminococcus_obeum	2.09E-03	1.25E-02	3.09E-01	2.76E-01	8.36E-01
Isobutyric	PWY-5179: NO_NAME|unclassified	4.22E-03	1.27E-02	3.55E-01	2.24E-01	8.18E-01
Isobutyric	PWY-5994: NO_NAME|unclassified	4.23E-03	1.27E-02	5.26E-01	2.23E-01	8.18E-01
Isobutyric	FASYN-ELONG-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	4.27E-03	1.28E-02	5.02E-01	2.23E-01	8.18E-01
Valeric	PANTOSYN-PWY: NO_NAME|g_Eubacterium.s_Eubacterium_hallii	2.14E-03	1.28E-02	-5.42E-01	-8.35E-01	-2.75E-01
Isobutyric	PWY-3841: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	4.38E-03	1.31E-02	5.74E-01	2.21E-01	8.17E-01
Isobutyric	P108-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	4.42E-03	1.33E-02	5.68E-01	2.20E-01	8.17E-01
Butyric	PWY-6892: NO_NAME|g_Blautia.s_Ruminococcus_obeum	2.22E-03	1.33E-02	7.97E-02	2.72E-01	8.35E-01
Isobutyric	PWY-4981: NO_NAME|unclassified	4.45E-03	1.33E-02	5.45E-01	2.20E-01	8.17E-01
Isovaleric	PWY-4981: NO_NAME|unclassified	2.76E-03	1.33E-02	5.49E-01	2.56E-01	8.29E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isobutyric	PWY-7208: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	4.45E-03	1.34E-02	5.20E-01	2.19E-01	8.17E-01
Valeric	LACTOSECAT-PWY: NO_NAME|g_Eubacterium.s_Eubacterium_hallii	2.25E-03	1.35E-02	-4.20E-01	-8.34E-01	-2.71E-01
Isovaleric	PWY0-1061: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	2.26E-03	1.36E-02	5.86E-01	2.71E-01	8.34E-01
Isobutyric	PWY0-1061: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	4.53E-03	1.36E-02	5.74E-01	2.18E-01	8.16E-01
Isovaleric	PWY-6644: NO_NAME|unclassified	2.27E-03	1.36E-02	4.75E-01	2.70E-01	8.34E-01
Isobutyric	PWY-5505: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	4.63E-03	1.39E-02	5.18E-01	2.16E-01	8.16E-01
Butyric	PWY-7388: NO_NAME|g_Veillonella.s_Veillonella_parvula	2.32E-03	1.39E-02	5.38E-01	2.69E-01	8.33E-01
Isobutyric	GLUDEG-II-PWY: NO_NAME	4.66E-03	1.40E-02	6.31E-01	2.16E-01	8.16E-01
Isobutyric	PWY-7003: NO_NAME|unclassified	4.74E-03	1.42E-02	5.19E-01	2.15E-01	8.15E-01
Isobutyric	PWY-6644: NO_NAME|unclassified	4.75E-03	1.42E-02	4.63E-01	2.15E-01	8.15E-01
Butyric	PWY-6891: NO_NAME|g_Blautia.s_Ruminococcus_obeum	2.40E-03	1.44E-02	7.30E-02	2.66E-01	8.33E-01
Isovaleric	URSIN-PWY: NO_NAME|unclassified	4.80E-03	1.44E-02	5.42E-01	2.14E-01	8.15E-01
Valeric	ARGORNPROST-PWY: NO_NAME|unclassified	2.40E-03	1.44E-02	6.36E-01	2.66E-01	8.33E-01
Isobutyric	PWY-7664: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	4.81E-03	1.44E-02	5.08E-01	2.13E-01	8.15E-01
Valeric	PWY-6608: NO_NAME|unclassified	2.42E-03	1.45E-02	7.06E-01	2.66E-01	8.32E-01
Butyric	METSYN-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	2.43E-03	1.46E-02	2.47E-01	2.65E-01	8.32E-01
Butyric	PWY-5973: NO_NAME|g_Haemophilus.s_Haemophilus_parainfluenzae	2.43E-03	1.46E-02	2.90E-01	2.65E-01	8.32E-01
Isovaleric	PWY-5104: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	2.43E-03	1.46E-02	4.99E-01	2.65E-01	8.32E-01
Isobutyric	PWY-7654: NO_NAME|unclassified	3.49E-03	1.46E-02	5.53E-01	2.38E-01	8.23E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isovaleric	PWY-7654: NO_NAME|unclassified	4.87E-03	1.46E-02	5.32E-01	2.12E-01	8.14E-01
Valeric	PWY-6612: NO_NAME|unclassified	2.44E-03	1.47E-02	6.45E-01	2.65E-01	8.32E-01
Isobutyric	GLUDEG-I-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	4.95E-03	1.48E-02	5.80E-01	2.11E-01	8.14E-01
Isovaleric	GLUDEG-I-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	2.65E-03	1.48E-02	5.87E-01	2.59E-01	8.30E-01
Butyric	PWY-5667: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	2.48E-03	1.49E-02	1.44E-01	2.64E-01	8.32E-01
Butyric	PWY0-1319: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	2.48E-03	1.49E-02	1.44E-01	2.64E-01	8.32E-01
Isovaleric	PWY-6644: NO_NAME	2.48E-03	1.49E-02	5.16E-01	2.64E-01	8.32E-01
Isobutyric	PWY-5081: NO_NAME	6.12E-03	1.49E-02	4.91E-01	1.94E-01	8.08E-01
Isovaleric	PWY-5081: NO_NAME	3.65E-03	1.49E-02	4.74E-01	2.35E-01	8.22E-01
Valeric	PWY-5081: NO_NAME	7.47E-03	1.49E-02	5.27E-01	1.78E-01	8.02E-01
Valeric	PWY-5464: NO_NAME|unclassified	7.48E-03	1.50E-02	6.22E-01	1.78E-01	8.02E-01
Isobutyric	P163-PWY: NO_NAME|unclassified	5.03E-03	1.51E-02	5.19E-01	2.10E-01	8.13E-01
Isovaleric	P163-PWY: NO_NAME|unclassified	4.35E-03	1.51E-02	5.13E-01	2.21E-01	8.17E-01
Isobutyric	PWY0-862: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	5.06E-03	1.52E-02	5.08E-01	2.09E-01	8.13E-01
Acetic	ILEUSYN-PWY: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	2.54E-03	1.52E-02	4.93E-01	2.62E-01	8.31E-01
Acetic	VALSYN-PWY: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	2.54E-03	1.52E-02	4.93E-01	2.62E-01	8.31E-01
Isobutyric	PWY-6282: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	5.09E-03	1.53E-02	5.08E-01	2.09E-01	8.13E-01
Isobutyric	PWY-4984: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	5.10E-03	1.53E-02	5.13E-01	2.09E-01	8.13E-01
Acetic	PWY-6123: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	2.58E-03	1.55E-02	5.38E-01	2.61E-01	8.31E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isovaleric	PWY0-1586: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	2.58E-03	1.55E-02	5.18E-01	2.61E-01	8.31E-01
Butyric	PWY-7013: NO_NAME|g_Blautia.s_Ruminococcus_obeum	2.60E-03	1.56E-02	1.45E-02	2.60E-01	8.31E-01
Isobutyric	PWY-3781: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	5.20E-03	1.56E-02	5.01E-01	2.07E-01	8.12E-01
Isobutyric	PWY-7228: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	5.21E-03	1.56E-02	5.23E-01	2.07E-01	8.12E-01
Isovaleric	PWY-7228: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	2.70E-03	1.56E-02	5.20E-01	2.58E-01	8.30E-01
Butyric	PWY-6387: NO_NAME|g_Blautia.s_Ruminococcus_obeum	2.61E-03	1.56E-02	1.82E-01	2.60E-01	8.31E-01
Proprionic	PWY-7237: NO_NAME|g_Bacteroides.s_Bacteroides_xylanisolvens	7.85E-03	1.57E-02	5.98E-01	1.74E-01	8.00E-01
Butyric	PEPTIDOGLYCANSYN-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	2.64E-03	1.59E-02	2.36E-01	2.59E-01	8.30E-01
Valeric	SALVADEHYPOX-PWY: NO_NAME|unclassified	8.02E-03	1.60E-02	6.23E-01	1.72E-01	8.00E-01
Valeric	PWY-7654: NO_NAME|unclassified	8.03E-03	1.61E-02	5.56E-01	1.72E-01	8.00E-01
Isobutyric	P164-PWY: NO_NAME|unclassified	5.40E-03	1.62E-02	6.58E-01	2.04E-01	8.11E-01
Isovaleric	P164-PWY: NO_NAME|unclassified	4.68E-03	1.62E-02	6.83E-01	2.16E-01	8.15E-01
Isobutyric	ANAGLYCOLYSIS-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	5.41E-03	1.62E-02	5.83E-01	2.04E-01	8.11E-01
Isobutyric	PWY-7111: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	5.42E-03	1.63E-02	5.04E-01	2.04E-01	8.11E-01
Butyric	PWY-5347: NO_NAME|g_Blautia.s_Ruminococcus_obeum	2.71E-03	1.63E-02	2.47E-01	2.57E-01	8.30E-01
Butyric	PWY-6168: NO_NAME|g_Blautia.s_Ruminococcus_obeum	2.72E-03	1.63E-02	2.50E-01	2.57E-01	8.30E-01
Isobutyric	ILEUSYN-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	5.45E-03	1.63E-02	4.92E-01	2.04E-01	8.11E-01
Isobutyric	VALSYN-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	5.45E-03	1.63E-02	4.92E-01	2.04E-01	8.11E-01
Isobutyric	PWY-7654: NO_NAME	3.79E-03	1.64E-02	5.55E-01	2.32E-01	8.21E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isovaleric	PWY-7654: NO_NAME	5.46E-03	1.64E-02	5.30E-01	2.03E-01	8.11E-01
Isovaleric	PWY-4702: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	2.73E-03	1.64E-02	5.08E-01	2.57E-01	8.29E-01
Valeric	FAO-PWY: NO_NAME|unclassified	2.74E-03	1.64E-02	6.75E-01	2.57E-01	8.29E-01
Valeric	PWY-7654: NO_NAME	8.25E-03	1.65E-02	5.68E-01	1.70E-01	7.99E-01
Valeric	PWY-5306: NO_NAME|unclassified	2.76E-03	1.65E-02	3.67E-01	2.56E-01	8.29E-01
Butyric	PWY-7663: NO_NAME|g_Blautia.s_Ruminococcus_obeum	2.80E-03	1.68E-02	2.44E-01	2.55E-01	8.29E-01
Isobutyric	FUCCAT-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	5.62E-03	1.68E-02	5.67E-01	2.01E-01	8.10E-01
Isovaleric	FUCCAT-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.37E-03	1.68E-02	5.77E-01	2.41E-01	8.24E-01
Butyric	COA-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	2.82E-03	1.69E-02	1.61E-01	2.54E-01	8.29E-01
Isobutyric	GLUCONEO-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	5.63E-03	1.69E-02	5.73E-01	2.01E-01	8.10E-01
Valeric	PWY-6123: NO_NAME|g_Blautia.s_Ruminococcus_gnavus	2.82E-03	1.69E-02	-5.89E-01	-8.29E-01	-2.54E-01
Valeric	PWY-7229: NO_NAME|g_Blautia.s_Ruminococcus_gnavus	2.83E-03	1.70E-02	-6.07E-01	-8.29E-01	-2.54E-01
Isobutyric	ARGORNPROST-PWY: NO_NAME|unclassified	5.67E-03	1.70E-02	4.79E-01	2.00E-01	8.10E-01
Isovaleric	METH-ACETATE-PWY: NO_NAME|unclassified	2.84E-03	1.71E-02	5.36E-01	2.54E-01	8.29E-01
Isovaleric	PWY-1622: NO_NAME	2.85E-03	1.71E-02	5.16E-01	2.53E-01	8.28E-01
Butyric	PWY-5104: NO_NAME|g_Veillonella.s_Veillonella_atypica	2.87E-03	1.72E-02	4.60E-01	2.53E-01	8.28E-01
Isovaleric	PWY-5088: NO_NAME|unclassified	2.88E-03	1.73E-02	6.16E-01	2.53E-01	8.28E-01
Isobutyric	PWY-6309: NO_NAME	5.79E-03	1.74E-02	4.87E-01	1.99E-01	8.09E-01
Isovaleric	PWY-6309: NO_NAME	4.62E-03	1.74E-02	4.57E-01	2.17E-01	8.16E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Butyric	PWY-7024: NO_NAME|unclassified	2.90E-03	1.74E-02	2.31E-01	2.52E-01	8.28E-01
Isobutyric	BRANCHED-CHAIN-AA-SYN-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	5.80E-03	1.74E-02	4.88E-01	1.98E-01	8.09E-01
Valeric	PWY-6126: NO_NAME|g_Blautia.s_Ruminococcus_gnavus	2.91E-03	1.74E-02	-6.07E-01	-8.28E-01	-2.52E-01
Valeric	PROPFERM-PWY: NO_NAME|unclassified	2.98E-03	1.79E-02	5.61E-01	2.50E-01	8.27E-01
Isobutyric	PWY-6644: NO_NAME	5.96E-03	1.79E-02	5.03E-01	1.96E-01	8.09E-01
Butyric	PWYG-321: NO_NAME|g_Veillonella.s_Veillonella_parvula	3.01E-03	1.80E-02	5.42E-01	2.50E-01	8.27E-01
Valeric	PWY-7420: NO_NAME|unclassified	3.01E-03	1.81E-02	5.45E-01	2.49E-01	8.27E-01
Butyric	PWY-7208: NO_NAME|g_Blautia.s_Ruminococcus_obeum	3.01E-03	1.81E-02	2.53E-01	2.49E-01	8.27E-01
Isobutyric	PWY-5173: NO_NAME|unclassified	7.67E-03	1.82E-02	4.61E-01	1.76E-01	8.01E-01
Isovaleric	PWY-5173: NO_NAME|unclassified	7.76E-03	1.82E-02	4.27E-01	1.75E-01	8.01E-01
Valeric	PWY-5173: NO_NAME|unclassified	9.08E-03	1.82E-02	5.18E-01	1.62E-01	7.96E-01
Acetic	PWY-5109: NO_NAME	6.06E-03	1.82E-02	4.62E-01	1.95E-01	8.08E-01
Butyric	PWY-5101: NO_NAME|g_Blautia.s_Ruminococcus_obeum	3.05E-03	1.83E-02	3.69E-01	2.49E-01	8.27E-01
Isobutyric	PWY-181: NO_NAME	6.12E-03	1.84E-02	4.91E-01	1.94E-01	8.08E-01
Isobutyric	3-HYDROXYPHENYLACETATE-DEGRADATION-PWY: NO_NAME|unclassified	5.37E-03	1.85E-02	4.78E-01	2.05E-01	8.12E-01
Isovaleric	3-HYDROXYPHENYLACETATE-DEGRADATION-PWY: NO_NAME|unclassified	6.16E-03	1.85E-02	4.95E-01	1.94E-01	8.08E-01
Butyric	ARGININE-SYN4-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	3.08E-03	1.85E-02	1.52E-01	2.48E-01	8.26E-01
Butyric	FASYN-ELONG-PWY: NO_NAME|g_Veillonella.s_Veillonella_parvula	3.10E-03	1.86E-02	5.42E-01	2.47E-01	8.26E-01
Isobutyric	PWY66-399: NO_NAME	6.21E-03	1.86E-02	5.43E-01	1.93E-01	8.07E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isovaleric	PWY66-399: NO_NAME	4.39E-03	1.86E-02	5.08E-01	2.21E-01	8.17E-01
Butyric	PWY-6282: NO_NAME|g_Veillonella.s_Veillonella_parvula	3.13E-03	1.88E-02	5.42E-01	2.47E-01	8.26E-01
Isobutyric	PWY66-400: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	6.28E-03	1.88E-02	5.68E-01	1.92E-01	8.07E-01
Valeric	SALVADEHYPOX-PWY: NO_NAME|g_Blautia.s_Ruminococcus_gnavus	3.14E-03	1.89E-02	-4.82E-01	-8.26E-01	-2.46E-01
Acetic	PWY-6124: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	3.15E-03	1.89E-02	5.42E-01	2.46E-01	8.26E-01
Valeric	PWY-7420: NO_NAME	3.16E-03	1.90E-02	5.45E-01	2.46E-01	8.26E-01
Butyric	PWY-6313: NO_NAME|g_Veillonella.s_Veillonella_parvula	3.18E-03	1.91E-02	4.70E-01	2.45E-01	8.26E-01
Acetic	PWY-7237: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	3.20E-03	1.92E-02	4.24E-01	2.45E-01	8.25E-01
Isobutyric	PWY0-1586: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	6.42E-03	1.93E-02	5.11E-01	1.90E-01	8.06E-01
Isobutyric	PWY-5103: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	6.44E-03	1.93E-02	4.88E-01	1.90E-01	8.06E-01
Isobutyric	PWY-4702: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	6.44E-03	1.93E-02	5.02E-01	1.90E-01	8.06E-01
Butyric	PWY-6892: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	3.23E-03	1.94E-02	1.29E-01	2.44E-01	8.25E-01
Isobutyric	PWY-5306: NO_NAME|unclassified	6.49E-03	1.95E-02	4.07E-01	1.89E-01	8.06E-01
Butyric	PWY-5744: NO_NAME|unclassified	3.25E-03	1.95E-02	2.31E-01	2.44E-01	8.25E-01
Isobutyric	PWY-7383: NO_NAME	6.49E-03	1.95E-02	5.44E-01	1.89E-01	8.06E-01
Isovaleric	PWY-7383: NO_NAME	5.05E-03	1.95E-02	5.12E-01	2.10E-01	8.13E-01
Isobutyric	PYRIDNUCSYN-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	6.52E-03	1.96E-02	5.64E-01	1.89E-01	8.06E-01
Isovaleric	PYRIDNUCSYN-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.43E-03	1.96E-02	5.76E-01	2.40E-01	8.24E-01
Butyric	PWY-7664: NO_NAME|g_Veillonella.s_Veillonella_parvula	3.28E-03	1.97E-02	5.42E-01	2.43E-01	8.25E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Butyric	PWY-6163: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	3.29E-03	1.97E-02	1.44E-01	2.43E-01	8.25E-01
Butyric	PWY-6993: NO_NAME	6.58E-03	1.97E-02	5.21E-01	1.88E-01	8.06E-01
Acetic	PWYG-321: NO_NAME|unclassified	3.29E-03	1.97E-02	4.71E-01	2.43E-01	8.25E-01
Butyric	PWY-7664: NO_NAME|g_Haemophilus.s_Haemophilus_parainfluenzae	3.29E-03	1.97E-02	2.90E-01	2.43E-01	8.25E-01
Butyric	PWY-6993: NO_NAME|unclassified	6.62E-03	1.98E-02	5.21E-01	1.88E-01	8.05E-01
Isobutyric	PWY-5104: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	6.66E-03	2.00E-02	4.85E-01	1.87E-01	8.05E-01
Butyric	PWY-5304: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	3.33E-03	2.00E-02	1.29E-01	2.42E-01	8.24E-01
Isovaleric	PWY-6936: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.34E-03	2.01E-02	5.09E-01	2.42E-01	8.24E-01
Butyric	PWY-3001: NO_NAME|g_Blautia.s_Ruminococcus_obeum	3.39E-03	2.03E-02	2.44E-01	2.40E-01	8.24E-01
Butyric	PWY-2942: NO_NAME|g_Blautia.s_Ruminococcus_obeum	3.40E-03	2.04E-02	2.94E-01	2.40E-01	8.24E-01
Acetic	PWY-7664: NO_NAME|unclassified	3.40E-03	2.04E-02	4.49E-01	2.40E-01	8.24E-01
Isovaleric	COA-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.47E-03	2.08E-02	5.71E-01	2.39E-01	8.23E-01
Valeric	PANTOSYN-PWY: NO_NAME|g_Dorea.s_Dorea_longicatena	3.47E-03	2.08E-02	-4.78E-01	-8.23E-01	-2.39E-01
Butyric	PWY-6282: NO_NAME|g_Haemophilus.s_Haemophilus_parainfluenzae	3.48E-03	2.09E-02	2.90E-01	2.39E-01	8.23E-01
Isobutyric	PWY-7383: NO_NAME|unclassified	6.97E-03	2.09E-02	5.10E-01	1.84E-01	8.04E-01
Isovaleric	PWY-7383: NO_NAME|unclassified	4.88E-03	2.09E-02	4.74E-01	2.12E-01	8.14E-01
Valeric	PWY-5686: NO_NAME|g_Blautia.s_Ruminococcus_gnavus	3.50E-03	2.10E-02	-6.07E-01	-8.23E-01	-2.38E-01
Isobutyric	PWY-5676: NO_NAME|unclassified	7.01E-03	2.10E-02	5.81E-01	1.83E-01	8.04E-01
Isovaleric	PWY-5676: NO_NAME|unclassified	6.26E-03	2.10E-02	5.70E-01	1.92E-01	8.07E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Butyric	PWY0-862: NO_NAME|g_Veillonella.s_Veillonella_parvula	3.53E-03	2.12E-02	5.35E-01	2.37E-01	8.23E-01
Valeric	ARGININE-SYN4-PWY: NO_NAME|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	3.54E-03	2.12E-02	-4.92E-01	-8.23E-01	-2.37E-01
Acetic	PWY0-862: NO_NAME|unclassified	3.55E-03	2.13E-02	4.36E-01	2.37E-01	8.23E-01
Isovaleric	PWY-6385: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.55E-03	2.13E-02	5.57E-01	2.37E-01	8.23E-01
Isobutyric	PWY-6936: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	7.10E-03	2.13E-02	4.99E-01	1.82E-01	8.03E-01
Butyric	FASYN-INITIAL-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	3.56E-03	2.14E-02	2.95E-01	2.37E-01	8.23E-01
Isobutyric	PWY-1042: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	7.14E-03	2.14E-02	5.68E-01	1.82E-01	8.03E-01
Isovaleric	PEPTIDOGLYCANSYN-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.57E-03	2.14E-02	5.44E-01	2.36E-01	8.23E-01
Acetic	PWY-6126: NO_NAME|g_Dialister.s_Dialister_invisus	3.58E-03	2.15E-02	3.25E-01	2.36E-01	8.23E-01
Acetic	PWY-6282: NO_NAME|unclassified	3.60E-03	2.16E-02	4.22E-01	2.36E-01	8.22E-01
Isobutyric	PWY-6387: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	7.21E-03	2.16E-02	5.48E-01	1.81E-01	8.03E-01
Isovaleric	PWY-6387: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	3.62E-03	2.16E-02	5.61E-01	2.36E-01	8.22E-01
Butyric	VALSYN-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	3.61E-03	2.16E-02	3.06E-01	2.36E-01	8.22E-01
Isobutyric	PWY-6385: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	7.22E-03	2.17E-02	5.44E-01	1.81E-01	8.03E-01
Acetic	PWY-5973: NO_NAME|unclassified	3.61E-03	2.17E-02	4.36E-01	2.36E-01	8.22E-01
Isobutyric	HISTSYN-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	7.24E-03	2.17E-02	5.64E-01	1.80E-01	8.03E-01
Isovaleric	HISTSYN-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	4.06E-03	2.17E-02	5.74E-01	2.27E-01	8.19E-01
Valeric	ARO-PWY: NO_NAME|g_Eubacterium.s_Eubacterium_hallii	3.63E-03	2.18E-02	-4.19E-01	-8.22E-01	-2.35E-01
Butyric	PWY-7024: NO_NAME	3.64E-03	2.19E-02	2.31E-01	2.35E-01	8.22E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Butyric	PWY0-1586: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	3.65E-03	2.19E-02	1.44E-01	2.35E-01	8.22E-01
Butyric	BRANCHED-CHAIN-AA-SYN-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	3.67E-03	2.20E-02	2.98E-01	2.35E-01	8.22E-01
Isobutyric	PWY-7094: NO_NAME|unclassified	7.35E-03	2.21E-02	4.29E-01	1.79E-01	8.02E-01
Valeric	PWY-7094: NO_NAME|unclassified	3.91E-03	2.21E-02	6.81E-01	2.30E-01	8.20E-01
Acetic	PWY-841: NO_NAME|g_Dialister.s_Dialister_invisus	3.68E-03	2.21E-02	3.25E-01	2.34E-01	8.22E-01
Isobutyric	1CMET2-PWY: NO_NAME|g_Dorea.s_Dorea_longicatena	7.40E-03	2.22E-02	-5.64E-01	-8.02E-01	-1.79E-01
Valeric	1CMET2-PWY: NO_NAME|g_Dorea.s_Dorea_longicatena	4.65E-03	2.22E-02	-4.33E-01	-8.16E-01	-2.16E-01
Butyric	PWY-6151: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	3.71E-03	2.22E-02	1.44E-01	2.34E-01	8.22E-01
Valeric	COMPLETE-ARO-PWY: NO_NAME|g_Eubacterium.s_Eubacterium_hallii	3.73E-03	2.24E-02	-4.19E-01	-8.22E-01	-2.33E-01
Valeric	HEMESYN2-PWY: NO_NAME|g_Streptococcus.s_Streptococcus_salivarius	3.74E-03	2.24E-02	5.07E-01	2.33E-01	8.21E-01
Butyric	PWY-5097: NO_NAME|g_Blautia.s_Ruminococcus_obeum	3.74E-03	2.24E-02	3.24E-01	2.33E-01	8.21E-01
Acetic	PWY-5989: NO_NAME|unclassified	7.50E-03	2.25E-02	4.14E-01	1.78E-01	8.02E-01
Valeric	PWY-5989: NO_NAME|unclassified	7.03E-03	2.25E-02	6.30E-01	1.83E-01	8.04E-01
Butyric	PWY-5989: NO_NAME|g_Haemophilus.s_Haemophilus_parainfluenzae	3.76E-03	2.26E-02	2.90E-01	2.33E-01	8.21E-01
Isobutyric	PWY-6703: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	7.53E-03	2.26E-02	5.16E-01	1.77E-01	8.02E-01
Isovaleric	PWY-6703: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	4.38E-03	2.26E-02	5.23E-01	2.21E-01	8.17E-01
Valeric	PWY-6121: NO_NAME|unclassified	3.77E-03	2.26E-02	6.57E-01	2.32E-01	8.21E-01
Isobutyric	PEPTIDOGLYCANSYN-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	7.54E-03	2.26E-02	5.33E-01	1.77E-01	8.02E-01
Acetic	PWY-6609: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_3_1_46FAA	4.85E-03	2.31E-02	3.47E-01	2.13E-01	8.14E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Butyric	PWY-6609: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_3_1_46FAA	7.69E-03	2.31E-02	5.30E-01	1.76E-01	8.01E-01
Acetic	DENOVOPURINE2-PWY: NO_NAME|g_Dialister.s_Dialister_invisus	3.84E-03	2.31E-02	3.25E-01	2.31E-01	8.21E-01
Butyric	PWY-6385: NO_NAME|g_Blautia.s_Ruminococcus_obeum	3.85E-03	2.31E-02	2.07E-01	2.31E-01	8.21E-01
Valeric	PWY-7388: NO_NAME|g_Eubacterium.s_Eubacterium_hallii	3.88E-03	2.33E-02	-4.10E-01	-8.20E-01	-2.30E-01
Acetic	PWY-6125: NO_NAME|g_Dialister.s_Dialister_invisus	3.91E-03	2.34E-02	3.25E-01	2.30E-01	8.20E-01
Butyric	PWY-6386: NO_NAME|g_Blautia.s_Ruminococcus_obeum	3.91E-03	2.34E-02	1.71E-01	2.30E-01	8.20E-01
Butyric	BIOTIN-BIOSYNTHESIS-PWY: NO_NAME|g_Haemophilus.s_Haemophilus_parainfluenzae	3.92E-03	2.35E-02	3.31E-01	2.29E-01	8.20E-01
Valeric	PWY-6737: NO_NAME|g_Blautia.s_Ruminococcus_gnavus	3.92E-03	2.35E-02	-6.07E-01	-8.20E-01	-2.29E-01
Butyric	PWY-5103: NO_NAME|g_Blautia.s_Ruminococcus_obeum	3.93E-03	2.36E-02	2.95E-01	2.29E-01	8.20E-01
Acetic	PWY0-166: NO_NAME|g_Dialister.s_Dialister_invisus	3.94E-03	2.36E-02	3.25E-01	2.29E-01	8.20E-01
Butyric	PWY-6609: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	3.94E-03	2.36E-02	1.29E-01	2.29E-01	8.20E-01
Proprionic	PWY-7391: NO_NAME	3.95E-03	2.37E-02	4.48E-01	2.29E-01	8.20E-01
Butyric	PWY-6519: NO_NAME|g_Haemophilus.s_Haemophilus_parainfluenzae	3.96E-03	2.38E-02	3.31E-01	2.29E-01	8.20E-01
Isobutyric	PWY-5088: NO_NAME|unclassified	7.94E-03	2.38E-02	5.95E-01	1.73E-01	8.00E-01
Valeric	PWY-6435: NO_NAME|unclassified	3.98E-03	2.39E-02	6.09E-01	2.28E-01	8.20E-01
Butyric	PWY-7456: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	3.99E-03	2.39E-02	1.29E-01	2.28E-01	8.20E-01
Butyric	PWY0-1296: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	4.01E-03	2.41E-02	5.04E-01	2.28E-01	8.20E-01
Acetic	PWY-7210: NO_NAME|g_Dialister.s_Dialister_invisus	4.03E-03	2.42E-02	3.25E-01	2.27E-01	8.20E-01
Acetic	PWY-7184: NO_NAME|g_Dialister.s_Dialister_invisus	4.03E-03	2.42E-02	3.25E-01	2.27E-01	8.20E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Valeric	FUCCAT-PWY: NO_NAME|g_Blautia.s_Ruminococcus_gnavus	4.03E-03	2.42E-02	-5.42E-01	-8.19E-01	-2.27E-01
Acetic	PWY-7198: NO_NAME|g_Dialister.s_Dialister_invisus	4.03E-03	2.42E-02	3.25E-01	2.27E-01	8.19E-01
Acetic	PWY-7220: NO_NAME|g_Dialister.s_Dialister_invisus	4.07E-03	2.44E-02	3.25E-01	2.27E-01	8.19E-01
Acetic	PWY-7222: NO_NAME|g_Dialister.s_Dialister_invisus	4.07E-03	2.44E-02	3.25E-01	2.27E-01	8.19E-01
Acetic	PWY-6545: NO_NAME|g_Dialister.s_Dialister_invisus	4.08E-03	2.45E-02	3.25E-01	2.26E-01	8.19E-01
Valeric	PWY-6606: NO_NAME|unclassified	1.22E-02	2.45E-02	6.66E-01	1.36E-01	7.86E-01
Acetic	PWY-5022: NO_NAME|g_Dorea.s_Dorea_longicatena	6.41E-03	2.45E-02	4.05E-01	1.90E-01	8.06E-01
Butyric	PWY-5022: NO_NAME|g_Dorea.s_Dorea_longicatena	8.17E-03	2.45E-02	2.44E-01	1.70E-01	7.99E-01
Butyric	PWY-5744: NO_NAME	4.10E-03	2.46E-02	2.31E-01	2.26E-01	8.19E-01
Isobutyric	PWY-5677: NO_NAME	1.23E-02	2.47E-02	3.87E-01	1.35E-01	7.86E-01
Butyric	NONOXIPENT-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	4.11E-03	2.47E-02	1.44E-01	2.26E-01	8.19E-01
Isobutyric	COA-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	8.27E-03	2.48E-02	5.60E-01	1.69E-01	7.99E-01
Isobutyric	PWY-5177: NO_NAME|unclassified	8.30E-03	2.49E-02	4.84E-01	1.69E-01	7.99E-01
Valeric	PWY-5177: NO_NAME|unclassified	7.31E-03	2.49E-02	6.06E-01	1.80E-01	8.03E-01
Butyric	PWY-5695: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	4.15E-03	2.49E-02	1.29E-01	2.25E-01	8.19E-01
Butyric	GOLPDLCAT-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	4.17E-03	2.50E-02	3.44E-02	2.25E-01	8.19E-01
Isobutyric	PWY-6435: NO_NAME|unclassified	8.35E-03	2.50E-02	4.60E-01	1.69E-01	7.98E-01
Isobutyric	PWY-5265: NO_NAME|unclassified	1.05E-02	2.51E-02	5.11E-01	1.50E-01	7.91E-01
Isovaleric	PWY-5265: NO_NAME|unclassified	1.25E-02	2.51E-02	4.68E-01	1.34E-01	7.85E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Valeric	PWY-5265: NO_NAME|unclassified	9.45E-03	2.51E-02	6.40E-01	1.58E-01	7.94E-01
Isovaleric	PWY-6737: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	4.20E-03	2.52E-02	4.88E-01	2.24E-01	8.18E-01
Isovaleric	ARGORNPROST-PWY: NO_NAME|unclassified	1.28E-02	2.55E-02	4.34E-01	1.32E-01	7.84E-01
Isovaleric	PWY-7094: NO_NAME|unclassified	1.29E-02	2.57E-02	4.03E-01	1.31E-01	7.84E-01
Butyric	PWY-6121: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	4.32E-03	2.59E-02	1.29E-01	2.22E-01	8.18E-01
Isobutyric	PWY-7197: NO_NAME|g_Bilophila.s_Bilophila_wadsworthia	4.61E-03	2.59E-02	5.29E-01	2.17E-01	8.16E-01
Isovaleric	PWY-7197: NO_NAME|g_Bilophila.s_Bilophila_wadsworthia	8.64E-03	2.59E-02	5.10E-01	1.66E-01	7.97E-01
Butyric	FASYN-ELONG-PWY: NO_NAME|g_Haemophilus.s_Haemophilus_parainfluenzae	4.36E-03	2.62E-02	2.90E-01	2.21E-01	8.17E-01
Valeric	POLYISOPRENSYN-PWY: NO_NAME|g_Streptococcus.s_Streptococcus_salivarius	4.38E-03	2.63E-02	5.07E-01	2.21E-01	8.17E-01
Butyric	BRANCHED-CHAIN-AA-SYN-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	4.40E-03	2.64E-02	1.44E-01	2.20E-01	8.17E-01
Isobutyric	PWY-2942: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	8.81E-03	2.64E-02	5.67E-01	1.64E-01	7.97E-01
Isovaleric	PWY-2942: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	5.13E-03	2.64E-02	5.77E-01	2.08E-01	8.13E-01
Butyric	PWY-724: NO_NAME|g_Blautia.s_Ruminococcus_obeum	4.41E-03	2.65E-02	2.71E-01	2.20E-01	8.17E-01
Butyric	NONMEVIPP-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	4.41E-03	2.65E-02	1.44E-01	2.20E-01	8.17E-01
Butyric	PWY-7560: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	4.41E-03	2.65E-02	1.44E-01	2.20E-01	8.17E-01
Butyric	PWY-6700: NO_NAME|g_Streptococcus.s_Streptococcus_salivarius	4.41E-03	2.65E-02	4.68E-01	2.20E-01	8.17E-01
Butyric	THRESYN-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	4.42E-03	2.65E-02	2.60E-01	2.20E-01	8.17E-01
Butyric	PWY-6630: NO_NAME|g_Blautia.s_Ruminococcus_obeum	4.44E-03	2.66E-02	2.75E-01	2.20E-01	8.17E-01
Butyric	PWY-6123: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	4.45E-03	2.67E-02	1.29E-01	2.20E-01	8.17E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Butyric	PWY-7388: NO_NAME|g_Haemophilus.s_Haemophilus_parainfluenzae	4.47E-03	2.68E-02	2.35E-01	2.19E-01	8.17E-01
Butyric	PWY-5136: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	4.48E-03	2.69E-02	1.44E-01	2.19E-01	8.17E-01
Acetic	PWY-5971: NO_NAME|unclassified	4.50E-03	2.70E-02	4.38E-01	2.19E-01	8.16E-01
Butyric	PWY-621: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	4.51E-03	2.70E-02	1.44E-01	2.19E-01	8.16E-01
Butyric	PWY-5103: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	4.51E-03	2.71E-02	1.29E-01	2.18E-01	8.16E-01
Isovaleric	P185-PWY: NO_NAME|unclassified	4.52E-03	2.71E-02	4.16E-01	2.18E-01	8.16E-01
Isobutyric	PWY-6609: NO_NAME|unclassified	1.36E-02	2.72E-02	4.53E-01	1.27E-01	7.82E-01
Isovaleric	PWY-6609: NO_NAME|unclassified	1.11E-02	2.72E-02	4.30E-01	1.44E-01	7.89E-01
Valeric	PWY-6609: NO_NAME|unclassified	7.66E-03	2.72E-02	7.08E-01	1.76E-01	8.01E-01
Butyric	PWY-7315: NO_NAME|g_Blautia.s_Ruminococcus_obeum	4.54E-03	2.72E-02	5.47E-02	2.18E-01	8.16E-01
Butyric	PWY-7316: NO_NAME|g_Blautia.s_Ruminococcus_obeum	4.54E-03	2.72E-02	5.47E-02	2.18E-01	8.16E-01
Isovaleric	FUCCAT-PWY: NO_NAME|unclassified	4.54E-03	2.73E-02	4.43E-01	2.18E-01	8.16E-01
Butyric	PWY-5384: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	4.57E-03	2.74E-02	2.80E-01	2.18E-01	8.16E-01
Isobutyric	PWY-1622: NO_NAME	9.18E-03	2.75E-02	4.95E-01	1.61E-01	7.95E-01
Butyric	UDPNAGSYN-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	4.59E-03	2.76E-02	1.44E-01	2.17E-01	8.16E-01
Isovaleric	GLUDEG-II-PWY: NO_NAME|unclassified	4.61E-03	2.77E-02	6.49E-01	2.17E-01	8.16E-01
Butyric	PWY-5686: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	4.61E-03	2.77E-02	1.29E-01	2.17E-01	8.16E-01
Acetic	NONOXIPENT-PWY: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	4.64E-03	2.79E-02	4.72E-01	2.16E-01	8.16E-01
Butyric	P161-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	4.65E-03	2.79E-02	1.44E-01	2.16E-01	8.16E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isovaleric	1CMET2-PWY: NO_NAME|g_Dorea.s_Dorea_longicatena	1.40E-02	2.79E-02	-5.08E-01	-7.81E-01	-1.24E-01
Butyric	PWY-5367: NO_NAME|g_Blautia.s_Ruminococcus_obeum	4.66E-03	2.80E-02	3.97E-02	2.16E-01	8.16E-01
Isobutyric	PWY-7357: NO_NAME|g_Dorea.s_Dorea_longicatena	5.08E-03	2.80E-02	-5.75E-01	-8.13E-01	-2.09E-01
Isovaleric	PWY-7357: NO_NAME|g_Dorea.s_Dorea_longicatena	9.34E-03	2.80E-02	-5.24E-01	-7.95E-01	-1.59E-01
Valeric	PWY-7237: NO_NAME|g_Dorea.s_Dorea_longicatena	4.71E-03	2.82E-02	-3.27E-01	-8.15E-01	-2.15E-01
Isobutyric	PWY-6353: NO_NAME|unclassified	9.43E-03	2.83E-02	4.49E-01	1.58E-01	7.95E-01
Isovaleric	PWY-6353: NO_NAME|unclassified	5.97E-03	2.83E-02	4.62E-01	1.96E-01	8.09E-01
Butyric	THISYNARA-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	4.73E-03	2.84E-02	1.44E-01	2.15E-01	8.15E-01
Valeric	PWY-6628: NO_NAME|g_Eubacterium.s_Eubacterium_hallii	4.74E-03	2.85E-02	-3.85E-01	-8.15E-01	-2.15E-01
Butyric	PWY-7219: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	4.74E-03	2.85E-02	1.44E-01	2.15E-01	8.15E-01
Isobutyric	PWY-7197: NO_NAME|unclassified	9.23E-03	2.86E-02	4.87E-01	1.60E-01	7.95E-01
Isovaleric	PWY-7197: NO_NAME|unclassified	9.52E-03	2.86E-02	4.84E-01	1.58E-01	7.94E-01
Butyric	PWY-7388: NO_NAME|g_Blautia.s_Ruminococcus_obeum	4.78E-03	2.87E-02	3.16E-01	2.14E-01	8.15E-01
Isobutyric	PWY66-391: NO_NAME|unclassified	1.44E-02	2.87E-02	4.44E-01	1.22E-01	7.80E-01
Isovaleric	PWY66-391: NO_NAME|unclassified	9.79E-03	2.87E-02	4.09E-01	1.55E-01	7.93E-01
Valeric	PWY66-391: NO_NAME|unclassified	1.25E-02	2.87E-02	6.73E-01	1.34E-01	7.85E-01
Isobutyric	VALDEG-PWY: NO_NAME|unclassified	9.58E-03	2.88E-02	5.48E-01	1.57E-01	7.94E-01
Isovaleric	VALDEG-PWY: NO_NAME|unclassified	5.50E-03	2.88E-02	5.65E-01	2.03E-01	8.11E-01
Isobutyric	PWY66-399: NO_NAME|unclassified	9.60E-03	2.88E-02	5.43E-01	1.57E-01	7.94E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isovaleric	PWY66-399: NO_NAME|unclassified	8.39E-03	2.88E-02	5.13E-01	1.68E-01	7.98E-01
Acetic	PWY-6305: NO_NAME|unclassified	4.81E-03	2.88E-02	5.47E-01	2.13E-01	8.15E-01
Valeric	PWY0-162: NO_NAME|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	4.81E-03	2.88E-02	-4.00E-01	-8.15E-01	-2.13E-01
Valeric	PWY-6628: NO_NAME|unclassified	4.84E-03	2.90E-02	6.81E-01	2.13E-01	8.15E-01
Isovaleric	7ALPHADEHYDROX-PWY: NO_NAME	4.84E-03	2.90E-02	4.77E-01	2.13E-01	8.14E-01
Butyric	COMPLETE-ARO-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	4.84E-03	2.90E-02	1.44E-01	2.13E-01	8.14E-01
Isobutyric	P185-PWY: NO_NAME|unclassified	9.68E-03	2.90E-02	4.03E-01	1.56E-01	7.94E-01
Isobutyric	PWY3O-355: NO_NAME	7.13E-03	2.91E-02	5.30E-01	1.82E-01	8.03E-01
Isovaleric	PWY3O-355: NO_NAME	9.68E-03	2.91E-02	5.40E-01	1.56E-01	7.94E-01
Butyric	ARO-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	4.86E-03	2.92E-02	1.44E-01	2.13E-01	8.14E-01
Butyric	ILEUSYN-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	4.88E-03	2.93E-02	2.65E-01	2.12E-01	8.14E-01
Acetic	PWY-6527: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	4.89E-03	2.94E-02	4.84E-01	2.12E-01	8.14E-01
Valeric	PYRIDNUCSAL-PWY: NO_NAME|unclassified	4.90E-03	2.94E-02	6.64E-01	2.12E-01	8.14E-01
Isovaleric	7ALPHADEHYDROX-PWY: NO_NAME|unclassified	5.01E-03	3.01E-02	4.77E-01	2.10E-01	8.14E-01
Valeric	PWY-7224: NO_NAME|unclassified	5.01E-03	3.01E-02	6.14E-01	2.10E-01	8.13E-01
Isobutyric	PWY-6122: NO_NAME|g_Dorea.s_Dorea_longicatena	9.69E-03	3.02E-02	-5.49E-01	-7.94E-01	-1.56E-01
Isobutyric	PWY-6277: NO_NAME|g_Dorea.s_Dorea_longicatena	9.69E-03	3.02E-02	-5.49E-01	-7.94E-01	-1.56E-01
Isovaleric	PWY-6122: NO_NAME|g_Dorea.s_Dorea_longicatena	1.47E-02	3.02E-02	-5.20E-01	-7.79E-01	-1.19E-01
Isovaleric	PWY-6277: NO_NAME|g_Dorea.s_Dorea_longicatena	1.47E-02	3.02E-02	-5.20E-01	-7.79E-01	-1.19E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Valeric	PWY-6122: NO_NAME|g_Dorea.s_Dorea_longicatena	1.51E-02	3.02E-02	-3.75E-01	-7.79E-01	-1.17E-01
Valeric	PWY-6277: NO_NAME|g_Dorea.s_Dorea_longicatena	1.51E-02	3.02E-02	-3.75E-01	-7.79E-01	-1.17E-01
Butyric	PWY-7539: NO_NAME|g_Blautia.s_Ruminococcus_obeum	5.05E-03	3.03E-02	2.54E-01	2.10E-01	8.13E-01
Butyric	PWY-6124: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	5.06E-03	3.04E-02	1.44E-01	2.09E-01	8.13E-01
Valeric	PWY-6309: NO_NAME|unclassified	1.52E-02	3.04E-02	5.62E-01	1.17E-01	7.78E-01
Butyric	PWY-3841: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	5.10E-03	3.06E-02	1.44E-01	2.09E-01	8.13E-01
Butyric	GLUTORN-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	5.11E-03	3.07E-02	1.44E-01	2.09E-01	8.13E-01
Butyric	PWY-5676: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	5.11E-03	3.07E-02	1.44E-01	2.09E-01	8.13E-01
Valeric	PWY-6630: NO_NAME|unclassified	5.14E-03	3.08E-02	6.06E-01	2.08E-01	8.13E-01
Butyric	PWY-7111: NO_NAME|g_Blautia.s_Ruminococcus_obeum	5.15E-03	3.09E-02	2.21E-01	2.08E-01	8.13E-01
Butyric	PWY-5188: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	5.16E-03	3.10E-02	1.44E-01	2.08E-01	8.13E-01
Butyric	PWY-6628: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	5.18E-03	3.11E-02	1.44E-01	2.08E-01	8.13E-01
Proprionic	PWY-6609: NO_NAME	5.20E-03	3.12E-02	-4.88E-01	-8.12E-01	-2.07E-01
Valeric	LACTOSECAT-PWY: NO_NAME|g_Dorea.s_Dorea_longicatena	5.25E-03	3.15E-02	-4.23E-01	-8.12E-01	-2.07E-01
Isovaleric	P163-PWY: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_7_1_58FAA	5.25E-03	3.15E-02	6.02E-01	2.07E-01	8.12E-01
Butyric	PWY-724: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	5.28E-03	3.17E-02	1.44E-01	2.06E-01	8.12E-01
Valeric	PWYG-321: NO_NAME|unclassified	1.06E-02	3.18E-02	5.64E-01	1.48E-01	7.91E-01
Isovaleric	PWY-5177: NO_NAME|unclassified	1.61E-02	3.22E-02	4.36E-01	1.12E-01	7.76E-01
Valeric	PWY-7539: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	5.38E-03	3.23E-02	-5.26E-01	-8.12E-01	-2.05E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Valeric	PWY-5973: NO_NAME|g_Blautia.s_Ruminococcus_gnavus	5.39E-03	3.23E-02	-6.07E-01	-8.11E-01	-2.04E-01
Isovaleric	PWY-6662: NO_NAME|unclassified	5.42E-03	3.25E-02	4.30E-01	2.04E-01	8.11E-01
Butyric	GLYCOGENSYNTH-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	5.43E-03	3.26E-02	1.44E-01	2.04E-01	8.11E-01
Butyric	PWY-6168: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	5.44E-03	3.26E-02	1.44E-01	2.04E-01	8.11E-01
Acetic	PWY-7221: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	5.44E-03	3.26E-02	4.80E-01	2.04E-01	8.11E-01
Valeric	NONMEVIPP-PWY: NO_NAME|g_Eubacterium.s_Eubacterium_hallii	5.45E-03	3.27E-02	-3.45E-01	-8.11E-01	-2.04E-01
Valeric	PWY-7560: NO_NAME|g_Eubacterium.s_Eubacterium_hallii	5.45E-03	3.27E-02	-3.45E-01	-8.11E-01	-2.04E-01
Isobutyric	PWY-7397: NO_NAME	9.87E-03	3.28E-02	2.46E-01	1.54E-01	7.93E-01
Isovaleric	PWY-7397: NO_NAME	1.20E-02	3.28E-02	2.62E-01	1.37E-01	7.87E-01
Valeric	PWY-7397: NO_NAME	1.64E-02	3.28E-02	1.22E-01	1.10E-01	7.76E-01
Butyric	PWY-3001: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	5.49E-03	3.29E-02	1.44E-01	2.03E-01	8.11E-01
Butyric	1CMET2-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	5.49E-03	3.29E-02	1.44E-01	2.03E-01	8.11E-01
Butyric	SER-GLYSYN-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	5.50E-03	3.30E-02	2.99E-01	2.03E-01	8.11E-01
Butyric	PWY-5659: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	5.50E-03	3.30E-02	1.44E-01	2.03E-01	8.11E-01
Valeric	3-HYDROXYPHENYLACETATE-DEGRADATION-PWY: NO_NAME|unclassified	1.66E-02	3.31E-02	4.61E-01	1.09E-01	7.75E-01
Isobutyric	PWY-6588: NO_NAME|unclassified	1.11E-02	3.32E-02	5.79E-01	1.45E-01	7.89E-01
Isovaleric	PWY-6588: NO_NAME|unclassified	1.01E-02	3.32E-02	5.51E-01	1.52E-01	7.92E-01
Valeric	LACTOSECAT-PWY: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	5.53E-03	3.32E-02	-3.70E-01	-8.11E-01	-2.02E-01
Isobutyric	PANTOSYN-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.12E-02	3.35E-02	5.67E-01	1.44E-01	7.89E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isovaleric	PANTOSYN-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	7.02E-03	3.35E-02	5.77E-01	1.83E-01	8.04E-01
Butyric	PWY-5659: NO_NAME|g_Streptococcus.s_Streptococcus_salivarius	5.60E-03	3.36E-02	3.54E-01	2.01E-01	8.10E-01
Butyric	PWY-5103: NO_NAME|g_Veillonella.s_Veillonella_atypica	5.60E-03	3.36E-02	4.92E-01	2.01E-01	8.10E-01
Butyric	PANTO-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	5.67E-03	3.40E-02	1.29E-01	2.00E-01	8.10E-01
Butyric	HSERMETANA-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	5.69E-03	3.41E-02	1.81E-01	2.00E-01	8.10E-01
Butyric	PEPTIDOGLYCANSYN-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	5.70E-03	3.42E-02	1.44E-01	2.00E-01	8.10E-01
Valeric	DENOVOPURINE2-PWY: NO_NAME	5.72E-03	3.43E-02	4.42E-01	2.00E-01	8.10E-01
Isobutyric	PWY-4242: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.15E-02	3.44E-02	5.60E-01	1.42E-01	7.88E-01
Isovaleric	PWY-4242: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	7.98E-03	3.44E-02	5.71E-01	1.72E-01	8.00E-01
Butyric	PWY-7664: NO_NAME|unclassified	1.38E-02	3.44E-02	2.51E-01	1.25E-01	7.82E-01
Valeric	PWY-7664: NO_NAME|unclassified	1.72E-02	3.44E-02	5.27E-01	1.05E-01	7.74E-01
Valeric	PWY-6859: NO_NAME|g_Streptococcus.s_Streptococcus_salivarius	5.76E-03	3.46E-02	5.07E-01	1.99E-01	8.10E-01
Acetic	PWY-5109: NO_NAME|unclassified	1.15E-02	3.46E-02	4.90E-01	1.41E-01	7.88E-01
Valeric	PWY-5109: NO_NAME|unclassified	1.15E-02	3.46E-02	4.71E-01	1.41E-01	7.88E-01
Butyric	PWY-5097: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	5.78E-03	3.47E-02	1.44E-01	1.99E-01	8.09E-01
Isobutyric	PWY-6737: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.16E-02	3.48E-02	4.74E-01	1.41E-01	7.88E-01
Proprionic	PWY-6284: NO_NAME|unclassified	5.79E-03	3.48E-02	-5.74E-01	-8.09E-01	-1.99E-01
Isobutyric	P162-PWY: NO_NAME|unclassified	1.16E-02	3.49E-02	2.74E-01	1.40E-01	7.88E-01
Valeric	P162-PWY: NO_NAME|unclassified	7.46E-03	3.49E-02	6.17E-01	1.78E-01	8.02E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Valeric	PWY-6471: NO_NAME|unclassified	5.84E-03	3.51E-02	6.64E-01	1.98E-01	8.09E-01
Valeric	PWY-5136: NO_NAME|unclassified	5.84E-03	3.51E-02	6.91E-01	1.98E-01	8.09E-01
Butyric	PWY-6609: NO_NAME|g_Blautia.s_Ruminococcus_obeum	5.85E-03	3.51E-02	9.64E-02	1.98E-01	8.09E-01
Proprionic	PWY-6285: NO_NAME|unclassified	5.86E-03	3.51E-02	-5.69E-01	-8.09E-01	-1.98E-01
Proprionic	PWY-5156: NO_NAME|unclassified	5.90E-03	3.54E-02	-5.70E-01	-8.09E-01	-1.97E-01
Valeric	PWY-6147: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	5.90E-03	3.54E-02	-5.18E-01	-8.09E-01	-1.97E-01
Isobutyric	P163-PWY: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_7_1_58FAA	1.18E-02	3.54E-02	5.74E-01	1.39E-01	7.87E-01
Butyric	THRESYN-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	5.91E-03	3.55E-02	1.29E-01	1.97E-01	8.09E-01
Isovaleric	PWY-6435: NO_NAME|unclassified	1.78E-02	3.55E-02	4.19E-01	1.03E-01	7.73E-01
Butyric	PWY-2723: NO_NAME	5.93E-03	3.56E-02	5.46E-01	1.97E-01	8.09E-01
Valeric	PWY-5686: NO_NAME|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	5.96E-03	3.58E-02	-4.02E-01	-8.09E-01	-1.96E-01
Proprionic	PWY-5367: NO_NAME|unclassified	5.97E-03	3.58E-02	-5.70E-01	-8.09E-01	-1.96E-01
Valeric	PWY-6700: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	5.97E-03	3.58E-02	-4.75E-01	-8.09E-01	-1.96E-01
Butyric	FAO-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	5.97E-03	3.58E-02	1.44E-01	1.96E-01	8.09E-01
Valeric	PWY-7242: NO_NAME|unclassified	1.79E-02	3.58E-02	6.26E-01	1.02E-01	7.72E-01
Proprionic	PWY-6113: NO_NAME|unclassified	5.97E-03	3.58E-02	-5.69E-01	-8.08E-01	-1.96E-01
Butyric	P562-PWY: NO_NAME|unclassified	5.99E-03	3.59E-02	3.36E-01	1.96E-01	8.08E-01
Butyric	PWY-5971: NO_NAME|unclassified	1.82E-02	3.64E-02	2.84E-01	1.00E-01	7.72E-01
Valeric	PWY-5971: NO_NAME|unclassified	1.23E-02	3.64E-02	5.62E-01	1.36E-01	7.86E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Butyric	PWY-6527: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	6.07E-03	3.64E-02	1.44E-01	1.95E-01	8.08E-01
Butyric	PWY-6387: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	6.07E-03	3.64E-02	1.44E-01	1.95E-01	8.08E-01
Isobutyric	HISTSYN-PWY: NO_NAME	1.22E-02	3.65E-02	-3.95E-01	-7.86E-01	-1.36E-01
Isovaleric	HISTSYN-PWY: NO_NAME	9.03E-03	3.65E-02	-4.29E-01	-7.96E-01	-1.62E-01
Valeric	PWY-6386: NO_NAME|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	6.10E-03	3.66E-02	-3.96E-01	-8.08E-01	-1.94E-01
Butyric	PWY-7539: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	6.12E-03	3.67E-02	1.44E-01	1.94E-01	8.08E-01
Butyric	COMPLETE-ARO-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	6.13E-03	3.68E-02	1.58E-01	1.94E-01	8.08E-01
Acetic	ARGORNPROST-PWY: NO_NAME	1.23E-02	3.70E-02	4.65E-01	1.35E-01	7.86E-01
Butyric	ARGORNPROST-PWY: NO_NAME	1.13E-02	3.70E-02	3.77E-01	1.43E-01	7.89E-01
Valeric	URSIN-PWY: NO_NAME|unclassified	1.86E-02	3.72E-02	4.82E-01	9.83E-02	7.71E-01
Isobutyric	PWY-7288: NO_NAME|unclassified	1.48E-02	3.73E-02	4.44E-01	1.19E-01	7.79E-01
Isovaleric	PWY-7288: NO_NAME|unclassified	9.05E-03	3.73E-02	4.09E-01	1.62E-01	7.96E-01
Valeric	PWY-7288: NO_NAME|unclassified	1.86E-02	3.73E-02	6.73E-01	9.81E-02	7.71E-01
Butyric	PWY-5973: NO_NAME|unclassified	1.25E-02	3.74E-02	2.64E-01	1.34E-01	7.85E-01
Isobutyric	CENTFERM-PWY: NO_NAME|unclassified	1.13E-02	3.74E-02	4.70E-01	1.43E-01	7.89E-01
Valeric	CENTFERM-PWY: NO_NAME|unclassified	1.25E-02	3.74E-02	5.12E-01	1.34E-01	7.85E-01
Valeric	PWY-7224: NO_NAME	6.25E-03	3.75E-02	5.97E-01	1.93E-01	8.07E-01
Isobutyric	PWY-6590: NO_NAME|unclassified	1.09E-02	3.75E-02	4.90E-01	1.46E-01	7.90E-01
Valeric	PWY-6590: NO_NAME|unclassified	1.25E-02	3.75E-02	5.16E-01	1.34E-01	7.85E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isobutyric	PWY-6662: NO_NAME|unclassified	1.25E-02	3.76E-02	4.17E-01	1.34E-01	7.85E-01
Valeric	PWY0-1479: NO_NAME	6.27E-03	3.76E-02	4.99E-01	1.92E-01	8.07E-01
Butyric	ILEUSYN-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	6.28E-03	3.77E-02	1.44E-01	1.92E-01	8.07E-01
Butyric	VALSYN-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	6.28E-03	3.77E-02	1.44E-01	1.92E-01	8.07E-01
Valeric	PWY-5676: NO_NAME|unclassified	1.89E-02	3.77E-02	7.14E-01	9.70E-02	7.70E-01
Butyric	BRANCHED-CHAIN-AA-SYN-PWY: NO_NAME|g_Veillonella.s_Veillonella_parvula	6.31E-03	3.78E-02	3.74E-01	1.92E-01	8.07E-01
Isobutyric	HOMOSER-METSYN-PWY: NO_NAME|g_Dorea.s_Dorea_longicatena	1.13E-02	3.79E-02	-5.07E-01	-7.89E-01	-1.43E-01
Isovaleric	HOMOSER-METSYN-PWY: NO_NAME|g_Dorea.s_Dorea_longicatena	1.89E-02	3.79E-02	-4.59E-01	-7.70E-01	-9.67E-02
Valeric	HOMOSER-METSYN-PWY: NO_NAME|g_Dorea.s_Dorea_longicatena	1.31E-02	3.79E-02	-3.54E-01	-7.84E-01	-1.30E-01
Valeric	P163-PWY: NO_NAME|unclassified	1.89E-02	3.79E-02	5.40E-01	9.66E-02	7.70E-01
Isobutyric	PWY0-1296: NO_NAME|unclassified	1.27E-02	3.81E-02	4.52E-01	1.33E-01	7.85E-01
Isovaleric	PWY0-1296: NO_NAME|unclassified	9.39E-03	3.81E-02	4.34E-01	1.59E-01	7.95E-01
Isobutyric	PWY-1042: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_7_1_58FAA	1.27E-02	3.82E-02	5.57E-01	1.32E-01	7.85E-01
Isovaleric	PWY-1042: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_7_1_58FAA	8.36E-03	3.82E-02	5.84E-01	1.69E-01	7.98E-01
Butyric	PWY0-862: NO_NAME|unclassified	1.28E-02	3.84E-02	2.64E-01	1.32E-01	7.84E-01
Acetic	PWY3O-355: NO_NAME|unclassified	6.39E-03	3.84E-02	4.58E-01	1.91E-01	8.06E-01
Butyric	PWY-6628: NO_NAME|g_Blautia.s_Ruminococcus_obeum	6.40E-03	3.84E-02	6.87E-02	1.91E-01	8.06E-01
Butyric	PWY-6282: NO_NAME|unclassified	1.35E-02	3.84E-02	2.60E-01	1.27E-01	7.83E-01
Valeric	PWY-6282: NO_NAME|unclassified	1.92E-02	3.84E-02	5.21E-01	9.54E-02	7.70E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Valeric	PWY-6700: NO_NAME|g_Burkholderiales_noname.s_Burkholderiales_bacterium_1_1_47	6.43E-03	3.86E-02	-4.38E-01	-8.06E-01	-1.90E-01
Butyric	PWYG-321: NO_NAME|unclassified	1.93E-02	3.86E-02	2.73E-01	9.49E-02	7.70E-01
Acetic	PWY-5188: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	6.45E-03	3.87E-02	4.74E-01	1.90E-01	8.06E-01
Acetic	HISTSYN-PWY: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	6.47E-03	3.88E-02	4.83E-01	1.90E-01	8.06E-01
Butyric	PWY-5973: NO_NAME|g_Blautia.s_Ruminococcus_obeum	6.48E-03	3.89E-02	2.21E-01	1.90E-01	8.06E-01
Butyric	ARGSYN-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	6.50E-03	3.90E-02	3.48E-01	1.89E-01	8.06E-01
Butyric	FASYN-INITIAL-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	6.51E-03	3.90E-02	1.44E-01	1.89E-01	8.06E-01
Butyric	PWY-7400: NO_NAME|g_Blautia.s_Ruminococcus_obeum	6.51E-03	3.90E-02	3.48E-01	1.89E-01	8.06E-01
Acetic	PROPFERM-PWY: NO_NAME|unclassified	1.30E-02	3.91E-02	4.77E-01	1.30E-01	7.84E-01
Isobutyric	COA-PWY-1: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.31E-02	3.92E-02	5.60E-01	1.30E-01	7.84E-01
Isovaleric	COA-PWY-1: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	7.91E-03	3.92E-02	5.71E-01	1.73E-01	8.00E-01
Butyric	PANTOSYN-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	6.54E-03	3.93E-02	1.44E-01	1.89E-01	8.06E-01
Butyric	PWY-5104: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	6.56E-03	3.93E-02	1.44E-01	1.89E-01	8.06E-01
Butyric	COA-PWY-1: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	6.57E-03	3.94E-02	1.44E-01	1.88E-01	8.06E-01
Isobutyric	PWY-5188: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.32E-02	3.95E-02	5.50E-01	1.30E-01	7.83E-01
Isovaleric	PWY-5188: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.06E-02	3.95E-02	5.56E-01	1.48E-01	7.91E-01
Acetic	HSERMETANA-PWY: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	6.60E-03	3.96E-02	4.73E-01	1.88E-01	8.06E-01
Butyric	PWY-7385: NO_NAME|unclassified	6.65E-03	3.99E-02	5.36E-01	1.87E-01	8.05E-01
Valeric	PWY-5973: NO_NAME|unclassified	2.00E-02	4.00E-02	5.25E-01	9.17E-02	7.68E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isovaleric	PWY-5306: NO_NAME|unclassified	2.00E-02	4.00E-02	3.99E-01	9.17E-02	7.68E-01
Valeric	PWY0-862: NO_NAME|unclassified	2.01E-02	4.02E-02	5.25E-01	9.12E-02	7.68E-01
Butyric	PWY-5667: NO_NAME|g_Blautia.s_Ruminococcus_obeum	6.75E-03	4.05E-02	1.54E-01	1.86E-01	8.05E-01
Butyric	PWY0-1319: NO_NAME|g_Blautia.s_Ruminococcus_obeum	6.75E-03	4.05E-02	1.54E-01	1.86E-01	8.05E-01
Valeric	ILEUSYN-PWY: NO_NAME|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	6.76E-03	4.05E-02	-4.00E-01	-8.05E-01	-1.86E-01
Butyric	PWY-6147: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	6.77E-03	4.06E-02	1.44E-01	1.86E-01	8.05E-01
Isobutyric	PWY-5173: NO_NAME	1.95E-02	4.07E-02	5.31E-01	9.40E-02	7.69E-01
Isovaleric	PWY-5173: NO_NAME	2.04E-02	4.07E-02	4.91E-01	8.99E-02	7.67E-01
Valeric	PWY-5173: NO_NAME	9.32E-03	4.07E-02	5.58E-01	1.59E-01	7.95E-01
Isobutyric	PWY-7197: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.35E-02	4.09E-02	5.30E-01	1.27E-01	7.83E-01
Isovaleric	PWY-7197: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.36E-02	4.09E-02	5.18E-01	1.26E-01	7.82E-01
Butyric	PWY-6385: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	6.83E-03	4.10E-02	1.44E-01	1.85E-01	8.05E-01
Butyric	PWY-5104: NO_NAME|g_Veillonella.s_Veillonella_parvula	6.85E-03	4.11E-02	3.75E-01	1.85E-01	8.04E-01
Valeric	PWY-7111: NO_NAME|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	6.87E-03	4.12E-02	-4.10E-01	-8.04E-01	-1.85E-01
Butyric	DTDPRHAMSYN-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	6.89E-03	4.14E-02	9.93E-02	1.85E-01	8.04E-01
Valeric	METHGLYUT-PWY: NO_NAME	6.89E-03	4.14E-02	6.13E-01	1.84E-01	8.04E-01
Isobutyric	PWY-5989: NO_NAME|unclassified	2.08E-02	4.16E-02	3.97E-01	8.79E-02	7.67E-01
Butyric	PWY-5189: NO_NAME|unclassified	6.98E-03	4.19E-02	5.43E-01	1.83E-01	8.04E-01
Valeric	HISTSYN-PWY: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	6.99E-03	4.19E-02	-4.05E-01	-8.04E-01	-1.83E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Valeric	PWY-6595: NO_NAME|unclassified	2.10E-02	4.20E-02	5.45E-01	8.70E-02	7.66E-01
Valeric	PWY-7221: NO_NAME|g_Eubacterium.s_Eubacterium_hallii	7.02E-03	4.21E-02	-4.27E-01	-8.04E-01	-1.83E-01
Isobutyric	RIBOSYN2-PWY: NO_NAME	1.40E-02	4.21E-02	-5.05E-01	-7.81E-01	-1.24E-01
Isovaleric	RIBOSYN2-PWY: NO_NAME	1.09E-02	4.21E-02	-4.90E-01	-7.90E-01	-1.46E-01
Valeric	PWY-5789: NO_NAME	7.02E-03	4.21E-02	5.09E-01	1.83E-01	8.04E-01
Butyric	PWY-7219: NO_NAME|g_Veillonella.s_Veillonella_parvula	7.04E-03	4.22E-02	4.17E-01	1.83E-01	8.04E-01
Acetic	PWY-7111: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	7.05E-03	4.23E-02	3.94E-01	1.83E-01	8.04E-01
Butyric	PWY-6386: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	7.05E-03	4.23E-02	1.44E-01	1.83E-01	8.04E-01
Acetic	PWY-5347: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	7.07E-03	4.24E-02	4.84E-01	1.82E-01	8.03E-01
Butyric	PWY-6590: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	7.07E-03	4.24E-02	1.44E-01	1.82E-01	8.03E-01
Valeric	PWY-6122: NO_NAME|unclassified	7.10E-03	4.26E-02	6.09E-01	1.82E-01	8.03E-01
Valeric	PWY-6277: NO_NAME|unclassified	7.10E-03	4.26E-02	6.09E-01	1.82E-01	8.03E-01
Butyric	PWY0-1061: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	7.10E-03	4.26E-02	1.29E-01	1.82E-01	8.03E-01
Proprionic	PWY-6609: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_3_1_46FAA	2.13E-02	4.26E-02	5.42E-01	8.57E-02	7.66E-01
Isobutyric	7ALPHADEHYDROX-PWY: NO_NAME	1.42E-02	4.27E-02	4.69E-01	1.23E-01	7.81E-01
Acetic	PWY-6936: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	7.16E-03	4.30E-02	4.96E-01	1.81E-01	8.03E-01
Isobutyric	METH-ACETATE-PWY: NO_NAME|unclassified	1.44E-02	4.31E-02	5.09E-01	1.22E-01	7.80E-01
Isobutyric	FAO-PWY: NO_NAME|unclassified	1.45E-02	4.34E-02	4.04E-01	1.21E-01	7.80E-01
Valeric	PWY-6309: NO_NAME	2.18E-02	4.36E-02	5.29E-01	8.36E-02	7.65E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isobutyric	PWY-6737: NO_NAME|g_Dorea.s_Dorea_longicatena	8.96E-03	4.36E-02	-5.37E-01	-7.96E-01	-1.63E-01
Isovaleric	PWY-6737: NO_NAME|g_Dorea.s_Dorea_longicatena	1.79E-02	4.36E-02	-4.88E-01	-7.72E-01	-1.02E-01
Valeric	PWY-6737: NO_NAME|g_Dorea.s_Dorea_longicatena	2.18E-02	4.36E-02	-4.14E-01	-7.65E-01	-8.35E-02
Valeric	HSERMETANA-PWY: NO_NAME	7.28E-03	4.37E-02	-5.75E-01	-8.03E-01	-1.80E-01
Valeric	PWY-621: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	7.28E-03	4.37E-02	-3.37E-01	-8.03E-01	-1.80E-01
Valeric	PWY-6518: NO_NAME|unclassified	7.31E-03	4.39E-02	3.61E-01	1.80E-01	8.02E-01
Isovaleric	PWY-6590: NO_NAME|unclassified	2.20E-02	4.40E-02	4.51E-01	8.28E-02	7.64E-01
Butyric	ANAGLYCOLYSIS-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	7.35E-03	4.41E-02	1.44E-01	1.79E-01	8.02E-01
Isobutyric	7ALPHADEHYDROX-PWY: NO_NAME|unclassified	1.47E-02	4.42E-02	4.69E-01	1.20E-01	7.79E-01
Butyric	PWY-7282: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	7.40E-03	4.44E-02	1.44E-01	1.79E-01	8.02E-01
Isovaleric	PWY-6700: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	7.42E-03	4.45E-02	4.92E-01	1.78E-01	8.02E-01
Butyric	PWY-6122: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	7.43E-03	4.46E-02	1.29E-01	1.78E-01	8.02E-01
Butyric	PWY-6277: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	7.43E-03	4.46E-02	1.29E-01	1.78E-01	8.02E-01
Isobutyric	HEME-BIOSYNTHESIS-II: NO_NAME	2.09E-02	4.47E-02	4.30E-01	8.75E-02	7.66E-01
Isovaleric	HEME-BIOSYNTHESIS-II: NO_NAME	2.24E-02	4.47E-02	4.17E-01	8.12E-02	7.64E-01
Valeric	HEME-BIOSYNTHESIS-II: NO_NAME	1.89E-02	4.47E-02	4.23E-01	9.69E-02	7.70E-01
Butyric	PWY-6121: NO_NAME|g_Blautia.s_Ruminococcus_obeum	7.46E-03	4.48E-02	1.94E-02	1.78E-01	8.02E-01
Isobutyric	ARO-PWY: NO_NAME|g_Dorea.s_Dorea_longicatena	1.32E-02	4.48E-02	-5.65E-01	-7.83E-01	-1.29E-01
Isovaleric	ARO-PWY: NO_NAME|g_Dorea.s_Dorea_longicatena	2.24E-02	4.48E-02	-5.18E-01	-7.64E-01	-8.11E-02

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Valeric	ARO-PWY: NO_NAME|g_Dorea.s_Dorea_longicatena	2.05E-02	4.48E-02	-3.67E-01	-7.67E-01	-8.92E-02
Acetic	CITRULBIO-PWY: NO_NAME|g_Bacteroides.s_Bacteroides_eggerthii	1.95E-02	4.50E-02	5.52E-01	9.40E-02	7.69E-01
Butyric	CITRULBIO-PWY: NO_NAME|g_Bacteroides.s_Bacteroides_eggerthii	1.28E-02	4.50E-02	4.17E-01	1.32E-01	7.84E-01
Proprionic	CITRULBIO-PWY: NO_NAME|g_Bacteroides.s_Bacteroides_eggerthii	2.25E-02	4.50E-02	5.93E-01	8.07E-02	7.64E-01
Isovaleric	CENTFERM-PWY: NO_NAME|unclassified	2.26E-02	4.52E-02	4.23E-01	8.02E-02	7.63E-01
Butyric	PWY-6123: NO_NAME|g_Blautia.s_Ruminococcus_obeum	7.54E-03	4.52E-02	3.44E-02	1.77E-01	8.02E-01
Butyric	PWY-5100: NO_NAME|g_Veillonella.s_Veillonella_parvula	7.54E-03	4.52E-02	4.58E-01	1.77E-01	8.02E-01
Isobutyric	COMPLETE-ARO-PWY: NO_NAME|g_Dorea.s_Dorea_longicatena	1.32E-02	4.53E-02	-5.65E-01	-7.83E-01	-1.29E-01
Isovaleric	COMPLETE-ARO-PWY: NO_NAME|g_Dorea.s_Dorea_longicatena	2.27E-02	4.53E-02	-5.18E-01	-7.63E-01	-7.99E-02
Valeric	COMPLETE-ARO-PWY: NO_NAME|g_Dorea.s_Dorea_longicatena	2.10E-02	4.53E-02	-3.67E-01	-7.66E-01	-8.71E-02
Butyric	PWY-5695: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	7.59E-03	4.55E-02	4.94E-01	1.77E-01	8.01E-01
Butyric	COA-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	7.59E-03	4.56E-02	1.44E-01	1.77E-01	8.01E-01
Valeric	PWY0-1586: NO_NAME|unclassified	7.61E-03	4.56E-02	6.99E-01	1.76E-01	8.01E-01
Isobutyric	GLUDEG-II-PWY: NO_NAME|unclassified	1.53E-02	4.59E-02	6.32E-01	1.16E-01	7.78E-01
Valeric	COA-PWY: NO_NAME|g_Eubacterium.s_Eubacterium_hallii	7.67E-03	4.60E-02	-4.65E-01	-8.01E-01	-1.76E-01
Isovaleric	ARGININE-SYN4-PWY: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	7.70E-03	4.62E-02	4.74E-01	1.75E-01	8.01E-01
Butyric	PYRIDNUCSYN-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	7.71E-03	4.63E-02	1.29E-01	1.75E-01	8.01E-01
Valeric	P161-PWY: NO_NAME|g_Streptococcus.s_Streptococcus_salivarius	7.76E-03	4.66E-02	3.49E-01	1.75E-01	8.01E-01
Valeric	PWY0-1261: NO_NAME|unclassified	7.76E-03	4.66E-02	6.31E-01	1.75E-01	8.01E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isobutyric	PWY-6700: NO_NAME|g_Akkermansia.s_Akkermansia_muciniphila	1.56E-02	4.67E-02	4.88E-01	1.15E-01	7.77E-01
Valeric	BRANCHED-CHAIN-AA-SYN-PWY: NO_NAME|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	7.79E-03	4.67E-02	-4.00E-01	-8.01E-01	-1.74E-01
Isobutyric	THISYNARA-PWY: NO_NAME	1.57E-02	4.70E-02	-5.49E-01	-7.77E-01	-1.14E-01
Isovaleric	THISYNARA-PWY: NO_NAME	1.51E-02	4.70E-02	-5.62E-01	-7.78E-01	-1.17E-01
Valeric	PWY-7197: NO_NAME|unclassified	2.36E-02	4.72E-02	5.73E-01	7.60E-02	7.62E-01
Valeric	VALSYN-PWY: NO_NAME|g_Blautia.s_Ruminococcus_gnavus	7.90E-03	4.74E-02	-4.61E-01	-8.00E-01	-1.73E-01
Acetic	PWY-5097: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	7.90E-03	4.74E-02	4.89E-01	1.73E-01	8.00E-01
Isobutyric	PROPFERM-PWY: NO_NAME|unclassified	2.37E-02	4.74E-02	2.97E-01	7.57E-02	7.61E-01
Valeric	PWY-6396: NO_NAME	7.92E-03	4.75E-02	4.83E-01	1.73E-01	8.00E-01
Butyric	MET-SAM-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	7.92E-03	4.75E-02	1.29E-01	1.73E-01	8.00E-01
Valeric	PWY-2941: NO_NAME|unclassified	7.93E-03	4.76E-02	6.09E-01	1.73E-01	8.00E-01
Isobutyric	HEMESYN2-PWY: NO_NAME|g_Streptococcus.s_Streptococcus_salivarius	1.59E-02	4.76E-02	4.59E-01	1.13E-01	7.77E-01
Isobutyric	PWY-6641: NO_NAME	8.56E-03	4.77E-02	5.58E-01	1.66E-01	7.98E-01
Isovaleric	PWY-6641: NO_NAME	1.59E-02	4.77E-02	5.48E-01	1.13E-01	7.77E-01
Valeric	PWY-6123: NO_NAME|unclassified	7.96E-03	4.78E-02	5.99E-01	1.73E-01	8.00E-01
Butyric	PWY-5188: NO_NAME|unclassified	7.98E-03	4.79E-02	4.73E-01	1.72E-01	8.00E-01
Valeric	PWY-5103: NO_NAME|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	8.00E-03	4.80E-02	-4.00E-01	-8.00E-01	-1.72E-01
Valeric	PWY-2942: NO_NAME|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	8.04E-03	4.82E-02	-4.80E-01	-8.00E-01	-1.72E-01
Valeric	CALVIN-PWY: NO_NAME|g_Blautia.s_Ruminococcus_gnavus	8.04E-03	4.83E-02	-6.07E-01	-8.00E-01	-1.72E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Isovaleric	FAO-PWY: NO_NAME|unclassified	2.42E-02	4.83E-02	3.70E-01	7.39E-02	7.61E-01
Isobutyric	PWY-7187: NO_NAME|unclassified	2.03E-02	4.87E-02	4.39E-01	9.04E-02	7.68E-01
Isovaleric	PWY-7187: NO_NAME|unclassified	2.43E-02	4.87E-02	4.21E-01	7.32E-02	7.60E-01
Valeric	PWY-7187: NO_NAME|unclassified	2.15E-02	4.87E-02	5.74E-01	8.49E-02	7.65E-01
Valeric	ASPASN-PWY: NO_NAME|g_Streptococcus.s_Streptococcus_salivarius	8.12E-03	4.87E-02	3.02E-01	1.71E-01	7.99E-01
Valeric	NONMEVIPP-PWY: NO_NAME|g_Burkholderiales_noname.s_Burkholderiales_bacterium_1_1_47	8.12E-03	4.87E-02	-4.38E-01	-7.99E-01	-1.71E-01
Valeric	PWY-7560: NO_NAME|g_Burkholderiales_noname.s_Burkholderiales_bacterium_1_1_47	8.12E-03	4.87E-02	-4.38E-01	-7.99E-01	-1.71E-01
Acetic	PWY-6737: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	8.12E-03	4.87E-02	4.87E-01	1.71E-01	7.99E-01
Isobutyric	GALACTUROCAT-PWY: NO_NAME|unclassified	1.68E-02	4.88E-02	4.17E-01	1.08E-01	7.75E-01
Isovaleric	GALACTUROCAT-PWY: NO_NAME|unclassified	2.44E-02	4.88E-02	3.81E-01	7.29E-02	7.60E-01
Valeric	GALACTUROCAT-PWY: NO_NAME|unclassified	8.31E-03	4.88E-02	6.35E-01	1.69E-01	7.99E-01
Acetic	MET-SAM-PWY: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	8.15E-03	4.89E-02	4.84E-01	1.71E-01	7.99E-01
Butyric	ARO-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	8.17E-03	4.90E-02	8.93E-02	1.70E-01	7.99E-01
Valeric	PWY-6168: NO_NAME|g_Eubacterium.s_Eubacterium_hallii	8.20E-03	4.92E-02	-3.29E-01	-7.99E-01	-1.70E-01
Acetic	METSYN-PWY: NO_NAME|g_Ruminococcus.s_Ruminococcus_bromii	8.21E-03	4.93E-02	4.84E-01	1.70E-01	7.99E-01
Isobutyric	FUCCAT-PWY: NO_NAME|unclassified	1.64E-02	4.93E-02	4.69E-01	1.10E-01	7.75E-01
Butyric	GLYCOGENSYNTH-PWY: NO_NAME|g_Blautia.s_Ruminococcus_obeum	8.24E-03	4.94E-02	4.81E-02	1.70E-01	7.99E-01
Butyric	ILEUSYN-PWY: NO_NAME|g_Veillonella.s_Veillonella_atypica	8.25E-03	4.95E-02	4.92E-01	1.70E-01	7.99E-01
Valeric	PANTO-PWY: NO_NAME|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	8.29E-03	4.97E-02	-3.38E-01	-7.99E-01	-1.69E-01

Low BMI Control						
SCFA	**Pathway**	**p value**	**P correct BH**	**PPMC**	**LCI**	**UCI**
Butyric	METSYN-PWY: NO_NAME|g_Coprococcus.s_Coprococcus_eutactus	8.29E-03	4.97E-02	1.29E-01	1.69E-01	7.99E-01
Isobutyric	POLYISOPRENSYN-PWY: NO_NAME|g_Streptococcus.s_Streptococcus_salivarius	1.66E-02	4.98E-02	4.59E-01	1.09E-01	7.75E-01
Isobutyric	PWY-4984: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_7_1_58FAA	1.66E-02	4.98E-02	5.36E-01	1.09E-01	7.75E-01
Isovaleric	PWY-4984: NO_NAME|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_7_1_58FAA	8.87E-03	4.98E-02	5.51E-01	1.64E-01	7.96E-01

Appendix B

CHAPTER 6 ADDITIONAL MATERIAL

Supplementary Table 6.6 – Component C

Values presented as log fold change of measured faecal metabolites before and after treatment (LOG_2 (Timepoint2)- LOG_2 (Timepoint1)).

Group	ID	Allantoin	Carnitine	Citrate	Glycine	Hippurate	Leucine
Exercise	10	0.36	-0.82	0.21	0.18	0.69	-0.03
	34	-0.35	1.44	-0.66	0.45	-0.39	0.14
	45	-1.08	-0.77	-0.08	0.52	0.30	-0.16
	47	0.22	0.24	0.56	0.68	-0.20	0.09
	49	-0.83	-1.36	1.01	0.38	1.07	0.19
	50	-0.38	0.18	-0.05	0.16	-0.95	0.26
	54	0.37	1.00	0.18	0.51	-0.47	0.13
	57	0.71	2.03	-0.37	0.92	0.10	0.20
	58	0.69	-0.63	0.56	0.28	-0.34	0.08
	62	0.06	-1.15	0.17	0.39	-1.30	0.38
	64	-0.94	-0.24	-0.69	0.17	0.45	0.06
	67	-0.14	-0.65	0.40	-0.95	0.97	-0.02
	74	0.04	-0.86	0.64	0.06	0.83	0.16
	79	0.00	-0.63	0.52	0.31	-0.16	0.08
	81	-0.36	0.44	-0.04	-0.18	0.82	0.11
	84	0.72	-0.07	0.20	0.22	-0.29	0.11
	85	-0.37	-2.18	0.54	-0.12	-0.32	-0.62
	86	0.87	-0.76	-0.14	0.17	0.32	-0.01
	87	-0.26	-2.89	0.07	-0.08	-0.15	0.06
	89	0.85	-2.89	0.43	0.01	0.22	0.09
	91	0.63	0.48	0.89	1.24	0.43	0.39
	94	-0.53	-2.49	-0.26	-0.62	-0.44	-0.39

Group	ID	Allantoin	Carnitine	Citrate	Glycine	Hippurate	Leucine
Exercise & Protein	8	0.49	-0.07	-0.40	-1.17	-1.58	-0.09
	16	-0.53	-0.97	0.15	-0.66	-0.12	-0.52
	21	0.31	1.30	1.20	0.88	0.21	0.19
	22	0.00	-0.59	1.13	0.43	0.67	0.10
	25	-1.12	-0.98	-0.89	-0.50	0.03	0.10
	26	-0.04	-3.06	0.13	0.20	-1.30	0.30
	29	0.00	-1.01	-0.17	-0.96	0.06	-0.11
	32	0.24	-1.01	-0.79	-0.13	-0.41	0.30
	35	0.75	-0.83	0.01	-0.21	-0.22	0.12
	37	0.33	-0.68	-0.08	0.56	0.56	-0.03
	39	0.66	-0.31	0.35	-0.16	0.22	0.04
	42	0.24	3.36	0.14	-1.30	0.47	-0.58
	43	-1.52	0.86	-0.80	-0.47	0.46	0.12
	46	0.60	1.56	0.14	0.37	0.50	0.26
	56	0.26	0.39	-0.11	0.00	-0.50	-0.03
	68	0.01	-0.34	0.23	-0.43	0.24	-0.09
	70	1.02	-1.67	0.53	0.40	-1.90	0.03
	73	1.06	-1.92	-0.18	-0.53	0.55	-0.36
	78	-0.29	0.15	0.33	-0.44	-0.05	0.14
	82	-0.14	-0.51	-0.04	-0.07	-0.33	-0.19
	90	-0.22	0.76	0.52	0.59	0.06	0.04
	98	-0.60	-0.42	-0.49	0.26	-0.10	0.10

Group	ID	Allantoin	Carnitine	Citrate	Glycine	Hippurate	Leucine
Protein	1	-0.13	-0.19	-0.02	-0.16	-1.09	-0.19
	4	0.52	-0.64	-1.16	-0.25	-1.03	-0.30
	5	0.07	0.75	0.01	0.40	-0.55	0.32
	9	-0.58	-0.58	0.38	0.40	0.05	0.45
	12	-1.09	-0.12	0.02	-0.33	1.43	0.00
	14	-0.13	1.66	-0.07	-0.88	-0.43	-0.19
	15	-0.17	0.20	0.43	-0.35	2.45	-0.04
	18	0.88	2.05	0.47	0.66	-1.11	0.22
	19	-0.21	1.92	0.18	-0.72	-0.33	0.13
	23	-0.16	0.45	0.54	-1.25	-0.61	-0.08
	36	-0.30	-2.48	1.01	0.09	2.33	0.03
	96	-0.25	-1.82	-0.97	-0.57	0.54	0.15
	99	-0.12	-0.06	-0.89	-0.90	0.94	-0.22
	102	1.12	-1.32	0.11	-0.56	-0.85	-0.75
	104	0.42	-0.52	0.08	0.49	-0.73	-0.05
	105	-1.02	-0.23	-1.00	0.16	-0.21	-0.05
	106	0.29	-2.11	0.20	0.14	0.50	0.20
	108	-0.81	1.76	-0.37	0.31	-0.87	0.34
	109	-1.26	1.43	0.42	0.51	0.16	0.15
	110	-1.45	0.39	-0.36	-0.09	0.22	0.12
	111	0.77	0.15	-0.39	0.21	1.25	0.13
	112	1.67	-0.97	-0.15	0.03	-0.60	-0.09
	115	-0.75	2.59	0.97	0.70	-0.72	0.23
	117	-1.76	0.67	0.14	-0.53	-0.29	-0.13

Group	ID	PAG	Proline Betaine	Succinate	Trans Aconitate	TMAO	Valine
Exercise	10	-0.12	-1.26	0.18	-0.45	-0.11	0.01
	34	-0.98	1.67	-1.22	-0.87	-3.12	0.12
	45	-0.64	-0.38	0.50	0.30	-1.56	0.00
	47	-0.61	0.98	0.06	-0.28	0.84	-0.01
	49	-0.81	0.92	0.24	-0.06	1.55	0.38
	50	-0.55	0.33	0.63	0.99	-0.97	0.23
	54	-1.29	1.03	-0.16	-0.40	-0.44	0.34
	57	-0.90	0.63	-0.87	-0.05	0.31	0.23
	58	0.65	0.62	-0.36	0.30	0.28	-0.12
	62	0.96	-2.39	-1.16	0.01	-2.04	-0.02
	64	0.95	-0.37	-0.76	-0.80	0.22	0.10
	67	0.89	-0.50	-0.30	-0.40	-0.89	-0.12
	74	-0.22	0.40	0.82	-0.02	0.04	0.14
	79	-0.40	-0.60	-0.12	0.29	-0.62	0.11
	81	-1.69	-0.54	0.29	1.13	-2.81	0.03
	84	-0.71	-0.87	-0.35	-0.69	-0.04	0.00
	85	0.02	-0.31	0.14	0.06	-4.64	-0.68
	86	-0.55	-0.13	0.14	0.71	-1.47	0.01
	87	-0.56	0.69	0.14	-0.28	0.27	0.06
	89	-0.44	-2.91	0.03	-0.67	-3.78	0.24
	91	0.06	-1.32	-0.58	0.45	2.72	0.11
	94	0.17	-1.35	-0.69	1.62	-1.78	-0.60

Group	ID	PAG	Proline Betaine	Succinate	Trans Aconitate	TMAO	Valine
Exercise & Protein	8	2.48	-1.73	-0.34	0.12	2.35	0.01
	16	-1.08	0.51	0.81	-0.40	-0.13	-0.25
	21	-1.03	-0.53	-0.61	-0.13	1.53	-0.02
	22	-1.43	0.72	0.40	0.48	0.14	0.17
	25	-0.08	0.40	-0.50	-0.45	-0.54	0.24
	26	0.80	-0.91	-0.34	0.36	-0.35	0.17
	29	0.25	-2.49	1.02	-0.01	-2.78	-0.18
	32	-0.38	0.99	-0.40	0.82	-0.82	0.10
	35	-1.20	-1.57	-0.84	0.80	-0.30	0.14
	37	-0.02	-0.06	0.15	-0.14	-0.22	-0.31
	39	0.74	-0.24	-0.35	0.31	-0.13	-0.17
	42	-0.32	0.29	-0.39	0.08	-0.95	-0.75
	43	0.56	0.03	-0.27	-0.45	0.35	-0.15
	46	-1.00	0.71	-0.06	0.08	-1.20	0.06
	56	-0.99	1.78	-0.17	-1.51	-0.09	-0.06
	68	-0.40	-2.04	-0.27	0.31	-1.42	-0.14
	70	-0.06	-1.73	-0.18	0.41	-1.97	-0.20
	73	0.50	-1.12	-0.35	0.15	0.65	-0.62
	78	0.90	0.24	0.95	0.33	0.84	0.07
	82	0.36	1.92	-0.21	-0.53	0.29	0.13
	90	-0.08	0.90	-0.41	0.13	-0.49	-0.03
	98	1.56	-0.35	0.42	1.31	0.24	-0.01

Group	ID	PAG	Proline Betaine	Succinate	Trans Aconitate	TMAO	Valine
Protein	1	0.91	-0.04	1.08	-0.46	0.82	-0.12
	4	0.52	-1.21	-0.63	-0.06	0.09	-0.23
	5	-0.01	0.81	-0.15	-0.15	0.67	0.29
	9	0.63	-1.68	-0.50	-0.04	0.87	0.38
	12	0.72	-2.62	0.53	0.39	-0.40	-0.07
	14	0.76	3.28	0.74	-1.30	0.58	0.38
	15	-0.23	-0.19	1.21	-0.41	1.24	-0.13
	18	0.07	0.72	-0.38	-0.12	0.02	-0.12
	19	0.79	-0.25	0.20	0.10	0.34	-0.10
	23	0.71	0.95	0.59	-0.20	1.11	-0.20
	36	0.56	0.39	-0.41	0.01	-1.17	0.08
	96	0.63	0.19	-1.01	-0.18	1.16	-0.14
	99	-0.15	0.61	0.43	0.07	-0.36	-0.05
	102	-0.02	-1.53	0.08	-0.15	0.99	-0.47
	104	-0.13	-0.05	-0.24	-0.19	-2.72	0.17
	105	0.90	-0.22	-0.44	-0.23	2.24	0.32
	106	-0.02	-1.24	-1.22	-0.22	0.53	0.32
	108	-0.38	0.45	-0.44	-0.82	-0.25	0.27
	109	0.07	0.62	0.65	-0.66	0.87	0.24
	110	1.46	-0.35	0.95	-0.59	-0.22	0.18
	111	-0.57	-1.30	-0.72	0.35	-2.70	-0.02
	112	-0.09	-0.20	0.35	-0.04	1.61	0.07
	115	1.56	1.45	0.73	0.03	0.51	0.30
	117	-0.03	-0.09	1.04	-0.58	2.19	-0.09

Supplementary Table 6.6 – Component D

Values presented as log fold change of measured urinary metabolites before and after treatment (LOG_2 (Timepoint2)- LOG_2 (Timepoint1)).

Group	ID	Dimethyl-amine (DMA)	Gluta-mate	Methyl-amine (MA)	Phenyl-acetate	Serine	Trimethyl-amine (TMA)	Tyro-sine
Exercise	10	0.33	0.17	-0.43	-0.02	0.16	-0.41	0.56
	34	-0.10	-0.74	-0.08	0.25	-0.20	0.52	-0.34
	47	-0.23	0.10	0.82	0.68	0.16	0.19	0.34
	49	-0.24	-1.21	0.95	-0.60	-0.79	0.52	-1.64
	50	0.09	0.53	-1.27	-0.23	0.12	-0.45	0.23
	54	-0.02	1.19	-0.64	0.23	0.87	-0.49	1.73
	57	-0.49	0.10	0.45	-0.16	0.44	0.53	0.66
	58	0.03	-0.28	-0.18	1.43	-0.26	1.28	-0.35
	62	-0.03	-0.10	0.11	-0.10	-0.19	0.57	-0.16
	64	-0.36	-0.77	0.29	0.32	-0.17	0.57	-0.80
	67	-0.20	-0.14	-0.06	0.00	-0.17	-0.20	-0.02
	74	-0.03	0.80	0.85	0.99	0.16	-0.05	0.79
	79	0.16	0.74	0.30	0.68	0.07	0.43	0.05
	81	0.06	0.56	0.11	-0.48	0.44	-0.23	0.42
	84	-0.06	-0.11	-0.23	0.87	-0.03	0.27	-0.25
	85	-0.04	0.05	-0.17	0.18	0.05	-0.80	0.23
	86	-0.15	0.58	-0.10	-0.10	0.61	-0.10	0.84
	87	0.00	-0.44	-0.66	-0.16	-0.04	0.07	-0.45
	89	-0.05	-0.73	0.26	0.20	-0.48	0.59	-0.73
	91	-0.20	-0.07	-0.73	-0.15	-0.04	-3.03	-0.08
	94	-0.05	-0.38	-0.44	-0.91	-0.22	-2.04	-0.46

Group	ID	Dimethyl-amine (DMA)	Glutamate	Methyl-amine (MA)	Phenyl-acetate	Serine	Trimethyl-amine (TMA)	Tyrosine
Exercise & Protein	8	0.15	0.79	-0.84	1.00	0.69	0.65	1.04
	16	-0.07	-0.20	-1.39	0.16	-0.15	-0.50	-0.16
	21	-0.28	-0.29	0.21	-0.22	-0.29	0.57	-0.36
	22	-0.31	-0.16	1.07	-0.44	0.39	0.09	0.31
	25	0.30	-0.29	-0.76	0.89	0.25	0.63	-0.04
	26	-0.59	0.31	0.24	-0.48	0.24	0.02	0.48
	29	0.48	0.96	-0.95	0.24	0.29	-0.61	0.81
	32	-0.04	0.33	0.10	-0.13	0.03	0.35	-0.01
	35	0.00	-0.27	0.00	-0.44	0.09	0.68	-0.12
	37	-0.18	-0.12	0.54	0.15	-0.18	0.14	-0.48
	39	0.43	0.26	-0.30	-0.12	0.03	-1.63	-0.01
	42	-0.41	0.39	0.69	-0.04	0.43	0.28	0.67
	43	-0.12	0.04	-0.37	0.27	0.25	0.34	0.40
	46	0.05	0.17	-0.01	-0.09	0.06	-0.92	0.03
	56	0.04	0.00	0.71	0.05	-0.08	1.78	0.02
	68	-0.07	-0.24	0.17	0.19	-0.20	0.59	-0.51
	70	0.03	-0.13	0.23	-0.06	-0.14	0.03	-0.22
	73	0.50	0.12	0.08	1.09	-0.14	0.27	-0.25
	78	-0.36	-0.56	-0.70	-0.34	0.41	-1.19	0.63
	82	0.07	0.60	0.34	0.55	0.25	0.38	0.55
	90	0.01	-0.36	-0.28	0.01	-0.04	0.61	-0.34
	98	0.39	0.69	-1.76	0.31	0.07	-1.12	0.21

Group	ID	Dimethyl-amine (DMA)	Glutamate	Methyl-amine (MA)	Phenyl-acetate	Serine	Trimethyl-amine (TMA)	Tyrosine
Protein	1	0.19	-0.02	-0.05	0.08	0.01	-0.02	0.22
	5	-0.39	-0.84	0.58	-0.58	0.11	-0.19	-0.30
	9	0.08	0.00	-0.05	-0.48	0.55	-1.02	0.35
	12	0.03	-0.34	-0.97	0.62	0.18	-0.58	0.07
	14	0.89	-0.39	-0.64	0.23	0.03	1.73	-0.07
	15	-0.62	-0.61	1.26	-0.90	1.29	2.19	0.77
	18	-0.43	-1.66	0.94	-0.74	-0.04	1.78	-0.72
	19	-0.06	-0.28	0.04	-0.45	0.02	-0.07	-0.13
	23	0.25	-0.44	-0.25	-0.27	-0.36	0.83	-0.87
	36	0.43	0.50	0.47	1.12	-0.09	-0.10	0.53
	96	-0.46	-0.21	-0.12	-0.03	0.26	-0.33	0.18
	99	-0.34	0.19	0.07	0.58	0.38	0.28	0.03
	102	0.00	-0.51	0.31	-0.02	-0.28	0.96	-0.92
	104	-0.15	0.28	-0.25	0.37	0.62	-0.30	0.64
	105	0.11	-0.44	-0.22	0.24	0.13	-0.47	-0.13
	106	0.01	-1.00	-0.40	-0.47	-0.45	-0.16	-0.73
	108	0.03	-0.29	0.18	0.28	0.22	0.51	-0.28
	109	0.20	-0.40	-0.11	0.33	0.14	0.26	-0.24
	110	0.33	-1.14	2.58	0.88	-0.29	-0.07	-0.80
	111	0.33	0.74	-0.23	-0.43	0.35	-0.94	0.54
	112	-0.56	0.36	0.02	-0.23	0.40	-0.87	0.81
	115	0.06	-0.18	-0.14	0.04	-0.33	-0.26	-0.20
	117	-1.02	-0.58	0.60	-0.40	0.12	1.21	-0.28

Appendix C

CHAPTER 8 ADDITIONAL MATERIAL

Supplementary Table 8.2

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
GALACTUROCAT-PWY: D-galacturonate degradation	4.99E-08	2.01E-05	2.29E-03	9.18E-03	6.89E-03
GLCMANNANAUT-PWY: superpathway of N-acetylglucosamine, N-acetylmannosamine and N-acetylneuraminate degradation	4.99E-08	2.01E-05	3.43E-03	7.46E-03	4.02E-03
PWY-5030: L-histidine degradation III	4.99E-08	2.01E-05	5.96E-03	1.31E-03	4.65E-03
PWY-5177: glutaryl-CoA degradation	4.99E-08	2.01E-05	1.58E-03	8.18E-03	6.59E-03
PWY-6305: putrescine biosynthesis IV	4.99E-08	2.01E-05	1.79E-03	7.80E-03	6.00E-03
PWY-6507: 4-deoxy-L-threo-hex-4-enopyranuronate degradation	4.99E-08	2.01E-05	1.39E-03	6.61E-03	5.22E-03
SER-GLYSYN-PWY: superpathway of L-serine and glycine biosynthesis I	4.99E-08	2.01E-05	2.26E-03	8.69E-03	6.43E-03
DTDPRHAMSYN-PWY: dTDP-L-rhamnose biosynthesis I|g_Bacteroides.s_Bacteroides_ovatus	9.97E-08	2.44E-05	1.85E-03	1.43E-04	1.71E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_stercoris	9.97E-08	2.44E-05	1.50E-03	9.40E-05	1.41E-03
PWY-7357: thiamin formation from pyrithiamine and oxythiamine (yeast)| g_Faecalibacterium.s_Faecalibacterium_prausnitzii	9.97E-08	2.44E-05	3.04E-04	2.01E-03	1.70E-03
BRANCHED-CHAIN-AA-SYN-PWY: superpathway of branched amino acid biosynthesis| g_Faecalibacterium.s_Faecalibacterium_prausnitzii	1.99E-07	2.44E-05	4.70E-04	2.07E-03	1.60E-03
DTDPRHAMSYN-PWY: dTDP-L-rhamnose biosynthesis I|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	1.99E-07	2.44E-05	5.03E-04	2.70E-03	2.19E-03
ILEUSYN-PWY: L-isoleucine biosynthesis I (from threonine)|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	1.99E-07	2.44E-05	5.30E-04	2.33E-03	1.80E-03
PWY-5103: L-isoleucine biosynthesis III|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	1.99E-07	2.44E-05	4.67E-04	2.03E-03	1.57E-03
PWY-5177: glutaryl-CoA degradation|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	1.99E-07	2.44E-05	5.97E-04	2.50E-03	1.91E-03
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	1.99E-07	2.44E-05	5.53E-04	2.41E-03	1.86E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	1.99E-07	2.44E-05	4.40E-04	2.70E-03	2.26E-03
PWY-724: superpathway of L-lysine, L-threonine and L-methionine biosynthesis II|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	1.99E-07	2.44E-05	3.50E-04	1.89E-03	1.54E-03
UNINTEGRATED|g_Bacteroides.s_Bacteroides_stercoris	1.99E-07	2.44E-05	1.56E-03	1.48E-04	1.42E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
UNINTEGRATED|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	1.99E-07	2.44E-05	5.26E-04	2.48E-03	1.96E-03
VALSYN-PWY: L-valine biosynthesis|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	1.99E-07	2.44E-05	5.30E-04	2.33E-03	1.80E-03
ASPASN-PWY: superpathway of L-aspartate and L-asparagine biosynthesis|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	3.49E-07	2.44E-05	4.14E-04	2.14E-03	1.73E-03
GALACTUROCAT-PWY: D-galacturonate degradation I|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	3.49E-07	2.44E-05	6.16E-04	2.77E-03	2.16E-03
PANTO-PWY: phosphopantothenate biosynthesis I|g_Bacteroides.s_Bacteroides_vulgatus	3.49E-07	2.44E-05	1.72E-03	2.74E-04	1.44E-03
PEPTIDOGLYCANSYN-PWY: peptidoglycan biosynthesis I (meso-diaminopimelate containing)|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	3.49E-07	2.44E-05	3.23E-04	2.09E-03	1.76E-03
PWY-1042: glycolysis IV (plant cytosol)|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	3.49E-07	2.44E-05	6.25E-04	2.65E-03	2.03E-03
PWY-3001: superpathway of L-isoleucine biosynthesis I|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	3.49E-07	2.44E-05	4.57E-04	2.41E-03	1.95E-03
PWY-5097: L-lysine biosynthesis VI|g_Bacteroides.s_Bacteroides_vulgatus	3.49E-07	2.44E-05	1.75E-03	2.08E-04	1.54E-03
PWY-6385: peptidoglycan biosynthesis III (mycobacteria)|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	3.49E-07	2.44E-05	2.92E-04	2.17E-03	1.88E-03
PWY-6386: UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	3.49E-07	2.44E-05	2.72E-04	2.00E-03	1.73E-03
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	3.49E-07	2.44E-05	3.54E-04	2.30E-03	1.95E-03
PWY-6507: 4-deoxy-L-threo-hex-4-enopyranuronate degradation|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	3.49E-07	2.44E-05	5.14E-04	2.24E-03	1.72E-03
PWY-6737: starch degradation V|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	3.49E-07	2.44E-05	4.84E-04	2.77E-03	2.28E-03
SER-GLYSYN-PWY: superpathway of L-serine and glycine biosynthesis I|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	3.49E-07	2.44E-05	2.85E-04	1.93E-03	1.65E-03
THISYNARA-PWY: superpathway of thiamin diphosphate biosynthesis III (eukaryotes)|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	3.49E-07	2.44E-05	3.37E-04	1.93E-03	1.59E-03
THRESYN-PWY: superpathway of L-threonine biosynthesis|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	3.49E-07	2.44E-05	4.53E-04	2.54E-03	2.08E-03
ARO-PWY: chorismate biosynthesis I	3.49E-07	2.44E-05	4.59E-03	8.72E-03	4.13E-03
PWY-7242: D-fructuronate degradation	3.49E-07	2.44E-05	2.46E-03	6.70E-03	4.25E-03
GLCMANNANAUT-PWY: superpathway of N-acetylglucosamine, N-acetylmannosamine and N-acetylneuraminate degradation|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	5.98E-07	2.44E-05	4.00E-04	2.17E-03	1.77E-03
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	5.98E-07	2.44E-05	3.13E-04	1.95E-03	1.64E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-5686: UMP biosynthesis|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	5.98E-07	2.44E-05	5.76E-04	2.51E-03	1.93E-03
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	5.98E-07	2.44E-05	5.85E-04	2.82E-03	2.23E-03
PWY-6123: inosine-5'-phosphate biosynthesis I|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	5.98E-07	2.44E-05	5.48E-04	2.86E-03	2.32E-03
PWY-6317: galactose degradation I (Leloir pathway)|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	5.98E-07	2.44E-05	3.53E-04	1.85E-03	1.50E-03
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Bacteroides.s_Bacteroides_vulgatus	5.98E-07	2.44E-05	1.84E-03	2.60E-04	1.58E-03
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	5.98E-07	2.44E-05	3.13E-04	1.95E-03	1.64E-03
PWY66-422: D-galactose degradation V (Leloir pathway)|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	5.98E-07	2.44E-05	3.53E-04	1.85E-03	1.50E-03
VALSYN-PWY: L-valine biosynthesis|g_Bacteroides.s_Bacteroides_vulgatus	5.98E-07	2.44E-05	1.84E-03	2.60E-04	1.58E-03
COMPLETE-ARO-PWY: superpathway of aromatic amino acid biosynthesis	5.98E-07	2.44E-05	4.96E-03	9.29E-03	4.34E-03
PWY-5097: L-lysine biosynthesis VI|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	9.47E-07	2.44E-05	4.04E-04	2.01E-03	1.60E-03
PWY-6124: inosine-5'-phosphate biosynthesis II|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	9.47E-07	2.44E-05	5.92E-04	2.71E-03	2.12E-03
PWY-6151: S-adenosyl-L-methionine cycle I|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	9.47E-07	2.44E-05	4.35E-04	2.82E-03	2.38E-03
PWY-6737: starch degradation V	9.47E-07	2.44E-05	3.44E-03	7.34E-03	3.90E-03
1CMET2-PWY: N10-formyl-tetrahydrofolate biosynthesis|g_Bacteroides.s_Bacteroides_coprocola	1.49E-06	2.44E-05	0.00E+00	2.60E-03	2.60E-03
1CMET2-PWY: N10-formyl-tetrahydrofolate biosynthesis|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.60E-03	0.00E+00	1.60E-03
ARGININE-SYN4-PWY: L-ornithine de novo biosynthesis|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.94E-03	0.00E+00	1.94E-03
ARGININE-SYN4-PWY: L-ornithine de novo biosynthesis|g_Bacteroides.s_Bacteroides_massiliensis	1.49E-06	2.44E-05	0.00E+00	2.00E-03	2.00E-03
ASPASN-PWY: superpathway of L-aspartate and L-asparagine biosynthesis|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.78E-03	0.00E+00	1.78E-03
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Acidaminococcus.s_Acidaminococcus_sp_BV3L6	1.49E-06	2.44E-05	1.84E-03	0.00E+00	1.84E-03
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Dialister.s_Dialister_invisus	1.49E-06	2.44E-05	1.40E-03	0.00E+00	1.40E-03
COA-PWY: coenzyme A biosynthesis I|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.49E-06	2.44E-05	0.00E+00	2.51E-03	2.51E-03
HISDEG-PWY: L-histidine degradation I|g_Acidaminococcus.s_Acidaminococcus_sp_BV3L6	1.49E-06	2.44E-05	1.93E-03	0.00E+00	1.93E-03
HISTSYN-PWY: L-histidine biosynthesis|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.82E-03	0.00E+00	1.82E-03
NONOXIPENT-PWY: pentose phosphate pathway (non-oxidative branch)|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.49E-06	2.44E-05	0.00E+00	2.05E-03	2.05E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PANTO-PWY: phosphopantothenate biosynthesis I|g_Acidaminococcus.s_Acidaminococcus_sp_BV3L6	1.49E-06	2.44E-05	1.84E-03	0.00E+00	1.84E-03
PANTO-PWY: phosphopantothenate biosynthesis I|g_Bacteroides.s_Bacteroides_coprocola	1.49E-06	2.44E-05	0.00E+00	3.44E-03	3.44E-03
PANTO-PWY: phosphopantothenate biosynthesis I|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	2.05E-03	0.00E+00	2.05E-03
PANTO-PWY: phosphopantothenate biosynthesis I|g_Bacteroides.s_Bacteroides_fragilis	1.49E-06	2.44E-05	7.87E-04	0.00E+00	7.87E-04
PANTO-PWY: phosphopantothenate biosynthesis I|g_Bacteroides.s_Bacteroides_massiliensis	1.49E-06	2.44E-05	0.00E+00	2.53E-03	2.53E-03
PANTO-PWY: phosphopantothenate biosynthesis I|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.49E-06	2.44E-05	0.00E+00	1.99E-03	1.99E-03
PWY-1042: glycolysis IV (plant cyto-sol)|g_Bacteroides.s_Bacteroides_coprocola	1.49E-06	2.44E-05	0.00E+00	2.80E-03	2.80E-03
PWY-1042: glycolysis IV (plant cyto-sol)|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.84E-03	0.00E+00	1.84E-03
PWY-1269: CMP-3-deoxy-D-manno-octulosonate biosynthesis I|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.83E-03	0.00E+00	1.83E-03
PWY-2942: L-lysine biosynthesis III|g_Bacteroides.s_Bacteroides_coprocola	1.49E-06	2.44E-05	0.00E+00	2.94E-03	2.94E-03
PWY-2942: L-lysine biosynthesis III|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.96E-03	0.00E+00	1.96E-03
PWY-2942: L-lysine biosynthesis III|g_Bacteroides.s_Bacteroides_massiliensis	1.49E-06	2.44E-05	0.00E+00	2.18E-03	2.18E-03
PWY-2942: L-lysine biosynthesis III|g_Dialister.s_Dialister_invisus	1.49E-06	2.44E-05	1.71E-03	0.00E+00	1.71E-03
PWY-2942: L-lysine biosynthesis III|g_Prevotella.s_Prevotella_copri	1.49E-06	2.44E-05	0.00E+00	3.86E-03	3.86E-03
PWY-3841: folate transformations II|g_Bacteroides.s_Bacteroides_coprocola	1.49E-06	2.44E-05	0.00E+00	2.92E-03	2.92E-03
PWY-3841: folate transformations II|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.75E-03	0.00E+00	1.75E-03
PWY-5097: L-lysine biosynthesis VI|g_Bacteroides.s_Bacteroides_coprocola	1.49E-06	2.44E-05	0.00E+00	2.88E-03	2.88E-03
PWY-5097: L-lysine biosynthesis VI|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.92E-03	0.00E+00	1.92E-03
PWY-5097: L-lysine biosynthesis VI|g_Bacteroides.s_Bacteroides_massiliensis	1.49E-06	2.44E-05	0.00E+00	2.22E-03	2.22E-03
PWY-5097: L-lysine biosynthesis VI|g_Dialister.s_Dialister_invisus	1.49E-06	2.44E-05	1.70E-03	0.00E+00	1.70E-03
PWY-5097: L-lysine biosynthesis VI|g_Prevotella.s_Prevotella_copri	1.49E-06	2.44E-05	0.00E+00	3.84E-03	3.84E-03
PWY-5188: tetrapyrrole biosynthesis I (from gluta-mate)|g_Dialister.s_Dialister_invisus	1.49E-06	2.44E-05	1.38E-03	0.00E+00	1.38E-03
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.56E-03	0.00E+00	1.56E-03
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Dialister.s_Dialister_invisus	1.49E-06	2.44E-05	1.47E-03	0.00E+00	1.47E-03
PWY-5686: UMP biosynthe-sis|g_Acidaminococcus.s_Acidaminococcus_sp_BV3L6	1.49E-06	2.44E-05	1.98E-03	0.00E+00	1.98E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-5686: UMP biosynthesis|g_Alistipes.s_Alistipes_putredinis	1.49E-06	2.44E-05	0.00E+00	2.94E-03	2.94E-03
PWY-5686: UMP biosynthesis|g_Bacteroides.s_Bacteroides_coprocola	1.49E-06	2.44E-05	0.00E+00	3.17E-03	3.17E-03
PWY-5686: UMP biosynthesis|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.94E-03	0.00E+00	1.94E-03
PWY-5686: UMP biosynthesis|g_Dialister.s_Dialister_invisus	1.49E-06	2.44E-05	1.47E-03	0.00E+00	1.47E-03
PWY-5686: UMP biosynthesis|g_Prevotella.s_Prevotella_copri	1.49E-06	2.44E-05	0.00E+00	3.78E-03	3.78E-03
PWY-5695: urate biosynthesis/inosine 5'-phosphate degradation|g_Bacteroides.s_Bacteroides_coprocola	1.49E-06	2.44E-05	0.00E+00	2.77E-03	2.77E-03
PWY-5695: urate biosynthesis/inosine 5'-phosphate degradation|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.81E-03	0.00E+00	1.81E-03
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I|g_Acidaminococcus.s_Acidaminococcus_sp_BV3L6	1.49E-06	2.44E-05	2.00E-03	0.00E+00	2.00E-03
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I|g_Dialister.s_Dialister_invisus	1.49E-06	2.44E-05	1.54E-03	0.00E+00	1.54E-03
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Acidaminococcus.s_Acidaminococcus_sp_BV3L6	1.49E-06	2.44E-05	1.89E-03	0.00E+00	1.89E-03
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Dialister.s_Dialister_invisus	1.49E-06	2.44E-05	1.48E-03	0.00E+00	1.48E-03
PWY-6124: inosine-5'-phosphate biosynthesis II|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.77E-03	0.00E+00	1.77E-03
PWY-6151: S-adenosyl-L-methionine cycle I|g_Prevotella.s_Prevotella_copri	1.49E-06	2.44E-05	0.00E+00	3.50E-03	3.50E-03
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Acidaminococcus.s_Acidaminococcus_sp_BV3L6	1.49E-06	2.44E-05	1.89E-03	0.00E+00	1.89E-03
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Dialister.s_Dialister_invisus	1.49E-06	2.44E-05	1.48E-03	0.00E+00	1.48E-03
PWY-6386: UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)|g_Acidaminococcus.s_Acidaminococcus_sp_BV3L6	1.49E-06	2.44E-05	2.05E-03	0.00E+00	2.05E-03
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)|g_Acidaminococcus.s_Acidaminococcus_sp_BV3L6	1.49E-06	2.44E-05	2.10E-03	0.00E+00	2.10E-03
PWY-6609: adenine and adenosine salvage III|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.52E-03	0.00E+00	1.52E-03
PWY-6700: queuosine biosynthesis|g_Acidaminococcus.s_Acidaminococcus_sp_BV3L6	1.49E-06	2.44E-05	1.83E-03	0.00E+00	1.83E-03
PWY-6700: queuosine biosynthesis|g_Bacteroides.s_Bacteroides_coprocola	1.49E-06	2.44E-05	0.00E+00	2.58E-03	2.58E-03
PWY-6700: queuosine biosynthesis|g_Prevotella.s_Prevotella_copri	1.49E-06	2.44E-05	0.00E+00	3.92E-03	3.92E-03
PWY-6703: preQ0 biosynthesis|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.32E-03	0.00E+00	1.32E-03
PWY-6703: preQ0 biosynthesis|g_Dialister.s_Dialister_invisus	1.49E-06	2.44E-05	1.41E-03	0.00E+00	1.41E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Acidaminococcus.s_Acidaminococcus_sp_BV3L6	1.49E-06	2.44E-05	1.96E-03	0.00E+00	1.96E-03
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Parabacteroides.s_Parabacteroides_merdae	1.49E-06	2.44E-05	0.00E+00	1.51E-03	1.51E-03
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.49E-06	2.44E-05	0.00E+00	1.95E-03	1.95E-03
PWY-7197: pyrimidine deoxyribonucleotide phosphorylation|g_Dialister.s_Dialister_invisus	1.49E-06	2.44E-05	1.18E-03	0.00E+00	1.18E-03
PWY-7208: superpathway of pyrimidine nucleobases salvage|g_Dialister.s_Dialister_invisus	1.49E-06	2.44E-05	1.24E-03	0.00E+00	1.24E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Acidaminococcus.s_Acidaminococcus_sp_BV3L6	1.49E-06	2.44E-05	1.81E-03	0.00E+00	1.81E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Alistipes.s_Alistipes_putredinis	1.49E-06	2.44E-05	0.00E+00	3.39E-03	3.39E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_coprocola	1.49E-06	2.44E-05	0.00E+00	3.10E-03	3.10E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.62E-03	0.00E+00	1.62E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_massiliensis	1.49E-06	2.44E-05	0.00E+00	2.12E-03	2.12E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Bifidobacterium.s_Bifidobacterium_bifidum	1.49E-06	2.44E-05	7.34E-04	0.00E+00	7.34E-04
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Dialister.s_Dialister_invisus	1.49E-06	2.44E-05	1.60E-03	0.00E+00	1.60E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Eubacterium.s_Eubacterium_ventriosum	1.49E-06	2.44E-05	1.35E-03	0.00E+00	1.35E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Parabacteroides.s_Parabacteroides_merdae	1.49E-06	2.44E-05	0.00E+00	2.40E-03	2.40E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.49E-06	2.44E-05	0.00E+00	1.70E-03	1.70E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Prevotella.s_Prevotella_copri	1.49E-06	2.44E-05	0.00E+00	3.71E-03	3.71E-03
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.23E-03	0.00E+00	1.23E-03
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_massiliensis	1.49E-06	2.44E-05	0.00E+00	2.68E-03	2.68E-03
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Bifidobacterium.s_Bifidobacterium_bifidum	1.49E-06	2.44E-05	6.98E-04	0.00E+00	6.98E-04
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Dialister.s_Dialister_invisus	1.49E-06	2.44E-05	1.42E-03	0.00E+00	1.42E-03
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Prevotella.s_Prevotella_copri	1.49E-06	2.44E-05	0.00E+00	3.72E-03	3.72E-03
PWY-7228: superpathway of guanosine nucleotides de novo biosynthesis I|g_Dialister.s_Dialister_invisus	1.49E-06	2.44E-05	1.24E-03	0.00E+00	1.24E-03
PWY-7663: gondoate biosynthesis (anaerobic)|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.96E-03	0.00E+00	1.96E-03
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.56E-03	0.00E+00	1.56E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Dialister.s_Dialister_invisus	1.49E-06	2.44E-05	1.47E-03	0.00E+00	1.47E-03
PWY0-845: superpathway of pyridoxal 5'-phosphate biosynthesis and salvage|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.54E-03	0.00E+00	1.54E-03
PYRIDOXSYN-PWY: pyridoxal 5'-phosphate biosynthesis I|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.52E-03	0.00E+00	1.52E-03
RHAMCAT-PWY: L-rhamnose degradation I|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.66E-03	0.00E+00	1.66E-03
UNINTEGRAT-ED|g_Acidaminococcus.s_Acidaminococcus_sp_BV3L6	1.49E-06	2.44E-05	1.99E-03	0.00E+00	1.99E-03
UNINTEGRATED|g_Alistipes.s_Alistipes_putredinis	1.49E-06	2.44E-05	0.00E+00	2.81E-03	2.81E-03
UNINTEGRATED|g_Bacteroides.s_Bacteroides_coprocola	1.49E-06	2.44E-05	0.00E+00	2.93E-03	2.93E-03
UNINTEGRATED|g_Bacteroides.s_Bacteroides_faecis	1.49E-06	2.44E-05	1.95E-03	0.00E+00	1.95E-03
UNINTEGRATED|g_Bacteroides.s_Bacteroides_fragilis	1.49E-06	2.44E-05	1.57E-03	0.00E+00	1.57E-03
UNINTEGRATED|g_Bacteroides.s_Bacteroides_massiliensis	1.49E-06	2.44E-05	0.00E+00	3.01E-03	3.01E-03
UNINTEGRATED|g_Bifidobacterium.s_Bifidobacterium_bifidum	1.49E-06	2.44E-05	6.64E-04	0.00E+00	6.64E-04
UNINTEGRATED|g_Blautia.s_Ruminococcus_gnavus	1.49E-06	2.44E-05	1.70E-03	0.00E+00	1.70E-03
UNINTEGRATED|g_Dialister.s_Dialister_invisus	1.49E-06	2.44E-05	1.51E-03	0.00E+00	1.51E-03
UNINTEGRATED|g_Eubacterium.s_Eubacterium_ventriosum	1.49E-06	2.44E-05	1.39E-03	0.00E+00	1.39E-03
UNINTEGRATED|g_Parabacteroides.s_Parabacteroides_distasonis	1.49E-06	2.44E-05	0.00E+00	2.46E-03	2.46E-03
UNINTEGRATED|g_Parabacteroides.s_Parabacteroides_merdae	1.49E-06	2.44E-05	0.00E+00	2.30E-03	2.30E-03
UNINTEGRATED|g_Paraprevotella.s_Paraprevotella_clara	1.49E-06	2.44E-05	0.00E+00	2.73E-03	2.73E-03
UNINTEGRAT-ED|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.49E-06	2.44E-05	0.00E+00	1.97E-03	1.97E-03
UNINTEGRATED|g_Prevotella.s_Prevotella_copri	1.49E-06	2.44E-05	0.00E+00	3.77E-03	3.77E-03
UNINTEGRATED|g_Ruminococcus.s_Ruminococcus_callidus	1.49E-06	2.44E-05	0.00E+00	2.52E-03	2.52E-03
VALSYN-PWY: L-valine biosynthe-sis|g_Acidaminococcus.s_Acidaminococcus_sp_BV3L6	1.49E-06	2.44E-05	1.96E-03	0.00E+00	1.96E-03
VALSYN-PWY: L-valine biosynthe-sis|g_Parabacteroides.s_Parabacteroides_merdae	1.49E-06	2.44E-05	0.00E+00	1.51E-03	1.51E-03
PWY-5005: biotin biosynthesis II	1.49E-06	2.44E-05	4.69E-03	0.00E+00	4.69E-03
PEPTIDOGLYCANSYN-PWY: peptidoglycan biosynthesis I (meso-diaminopimelate containing)|g_Bacteroides.s_Bacteroides_vulgatus	1.50E-06	2.44E-05	1.78E-03	2.63E-04	1.51E-03
PWY-3841: folate transformations II|g_Bacteroides.s_Bacteroides_vulgatus	1.50E-06	2.44E-05	1.70E-03	2.91E-04	1.41E-03
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	1.50E-06	2.44E-05	6.79E-04	2.76E-03	2.08E-03
PWY-6151: S-adenosyl-L-methionine cycle I|g_Bacteroides.s_Bacteroides_vulgatus	1.50E-06	2.44E-05	1.78E-03	2.50E-04	1.53E-03
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	1.50E-06	2.44E-05	6.79E-04	2.76E-03	2.08E-03
PWY-7220: adenosine deoxyribonucleotides de novo biosynthesis II|g_Bacteroides.s_Bacteroides_vulgatus	1.50E-06	2.44E-05	1.74E-03	2.66E-04	1.47E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	1.50E-06	2.44E-05	5.30E-04	2.87E-03	2.34E-03
PWY-7222: guanosine deoxyribonucleotides de novo biosynthesis II|g_Bacteroides.s_Bacteroides_vulgatus	1.50E-06	2.44E-05	1.74E-03	2.66E-04	1.47E-03
PWY-7228: superpathway of guanosine nucleotides de novo biosynthesis I|g_Bacteroides.s_Bacteroides_vulgatus	1.50E-06	2.44E-05	1.74E-03	2.78E-04	1.46E-03
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II	1.50E-06	2.44E-05	4.47E-03	7.25E-03	2.78E-03
PWY-6163: chorismate biosynthesis from 3-dehydroquinate	1.50E-06	2.44E-05	4.70E-03	8.89E-03	4.18E-03
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis	1.50E-06	2.44E-05	4.47E-03	7.25E-03	2.78E-03
PWY0-162: superpathway of pyrimidine ribonucleotides de novo biosynthesis	1.50E-06	2.44E-05	4.38E-03	8.67E-03	4.30E-03
PYRIDOXSYN-PWY: pyridoxal 5'-phosphate biosynthesis I	1.50E-06	2.44E-05	5.59E-03	1.67E-03	3.92E-03
PWY-4242: pantothenate and coenzyme A biosynthesis III|g_Bacteroides.s_Bacteroides_xylanisolvens	2.08E-06	3.37E-05	1.61E-03	2.70E-05	1.58E-03
GALACT-GLUCUROCAT-PWY: superpathway of hexuronide and hexuronate degradation|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	2.24E-06	3.48E-05	5.99E-04	2.44E-03	1.84E-03
NONOXIPENT-PWY: pentose phosphate pathway (non-oxidative branch)|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	2.24E-06	3.48E-05	3.13E-04	2.10E-03	1.79E-03
PWY-5659: GDP-mannose biosynthesis|g_Bacteroides.s_Bacteroides_ovatus	2.24E-06	3.48E-05	2.02E-03	3.09E-04	1.71E-03
PWY-5695: urate biosynthesis/inosine 5'-phosphate degradation|g_Bacteroides.s_Bacteroides_vulgatus	2.24E-06	3.48E-05	1.78E-03	2.65E-04	1.51E-03
PWY-6126: superpathway of adenosine nucleotides de novo biosynthesis II|g_Bacteroides.s_Bacteroides_vulgatus	2.24E-06	3.48E-05	1.77E-03	2.64E-04	1.51E-03
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)|g_Bacteroides.s_Bacteroides_vulgatus	2.24E-06	3.48E-05	1.78E-03	2.68E-04	1.51E-03
PWY-7229: superpathway of adenosine nucleotides de novo biosynthesis I|g_Bacteroides.s_Bacteroides_vulgatus	2.24E-06	3.48E-05	1.78E-03	2.72E-04	1.50E-03
NONOXIPENT-PWY: pentose phosphate pathway (non-oxidative branch)	2.24E-06	3.48E-05	5.12E-03	8.68E-03	3.56E-03
PWY-2942: L-lysine biosynthesis III|g_Bacteroides.s_Bacteroides_vulgatus	3.34E-06	5.04E-05	1.78E-03	2.82E-04	1.50E-03
PWY-6151: S-adenosyl-L-methionine cycle I|g_Bacteroides.s_Bacteroides_uniformis	3.34E-06	5.04E-05	1.47E-03	2.89E-04	1.18E-03
PWY-6609: adenine and adenosine salvage III|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	3.34E-06	5.04E-05	4.92E-04	2.35E-03	1.86E-03
PWY-7242: D-fructuronate degradation|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	3.34E-06	5.04E-05	5.99E-04	2.43E-03	1.84E-03
PWY0-1296: purine ribonucleosides degradation|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	3.34E-06	5.04E-05	5.40E-04	2.30E-03	1.76E-03
1CMET2-PWY: N10-formyl-tetrahydrofolate biosynthesis|g_Bacteroides.s_Bacteroides_stercoris	3.40E-06	5.11E-05	1.52E-03	1.69E-05	1.51E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY66-422: D-galactose degradation V (Leloir pathway)|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	3.45E-06	5.14E-05	1.15E-03	5.04E-05	1.10E-03
1CMET2-PWY: N10-formyl-tetrahydrofolate biosynthesis|g_Bacteroides.s_Bacteroides_ovatus	4.28E-06	6.32E-05	1.70E-03	6.58E-05	1.64E-03
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Bacteroides.s_Bacteroides_xylanisolvens	4.28E-06	6.32E-05	1.35E-03	4.34E-05	1.31E-03
GLUCUROCAT-PWY: superpathway of β-D-glucuronide and D-glucuronate degradation|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	4.84E-06	6.41E-05	5.85E-04	2.33E-03	1.74E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_ovatus	4.84E-06	6.41E-05	1.79E-03	3.08E-04	1.48E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_vulgatus	4.84E-06	6.41E-05	1.77E-03	3.48E-04	1.42E-03
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_vulgatus	4.84E-06	6.41E-05	1.68E-03	2.91E-04	1.39E-03
PWY-6317: galactose degradation I (Leloir pathway)|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	5.45E-06	6.41E-05	1.05E-03	5.46E-05	9.92E-04
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Bacteroides.s_Bacteroides_coprocola	5.52E-06	6.41E-05	0.00E+00	2.51E-03	2.51E-03
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Bacteroides.s_Bacteroides_faecis	5.52E-06	6.41E-05	1.67E-03	0.00E+00	1.67E-03
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	5.52E-06	6.41E-05	0.00E+00	2.27E-03	2.27E-03
DTDPRHAMSYN-PWY: dTDP-L-rhamnose biosynthesis I|g_Bacteroides.s_Bacteroides_fragilis	5.52E-06	6.41E-05	1.09E-03	0.00E+00	1.09E-03
NONMEVIPP-PWY: methylerythritol phosphate pathway I|g_Alistipes.s_Alistipes_finegoldii	5.52E-06	6.41E-05	6.41E-04	0.00E+00	6.41E-04
PANTO-PWY: phosphopantothenate biosynthesis I|g_Parabacteroides.s_Parabacteroides_merdae	5.52E-06	6.41E-05	0.00E+00	1.62E-03	1.62E-03
PWY-2942: L-lysine biosynthesis III|g_Bacteroides.s_Bacteroides_fragilis	5.52E-06	6.41E-05	8.38E-04	0.00E+00	8.38E-04
PWY-2942: L-lysine biosynthesis III|g_Parabacteroides.s_Parabacteroides_merdae	5.52E-06	6.41E-05	0.00E+00	1.97E-03	1.97E-03
PWY-3841: folate transformations II|g_Bacteroides.s_Bacteroides_massiliensis	5.52E-06	6.41E-05	0.00E+00	2.52E-03	2.52E-03
PWY-3841: folate transformations II|g_Parabacteroides.s_Parabacteroides_merdae	5.52E-06	6.41E-05	0.00E+00	1.87E-03	1.87E-03
PWY-5097: L-lysine biosynthesis VI|g_Parabacteroides.s_Parabacteroides_merdae	5.52E-06	6.41E-05	0.00E+00	2.23E-03	2.23E-03
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Bacteroides.s_Bacteroides_coprocola	5.52E-06	6.41E-05	0.00E+00	2.45E-03	2.45E-03
PWY-5686: UMP biosynthesis|g_Bacteroides.s_Bacteroides_fragilis	5.52E-06	6.41E-05	7.44E-04	0.00E+00	7.44E-04
PWY-5686: UMP biosynthesis|g_Parabacteroides.s_Parabacteroides_merdae	5.52E-06	6.41E-05	0.00E+00	1.77E-03	1.77E-03
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I|g_Alistipes.s_Alistipes_putredinis	5.52E-06	6.41E-05	0.00E+00	2.65E-03	2.65E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Alistipes.s_Alistipes_putredinis	5.52E-06	6.41E-05	0.00E+00	2.63E-03	2.63E-03
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Eubacterium.s_Eubacterium_ventriosum	5.52E-06	6.41E-05	1.10E-03	0.00E+00	1.10E-03
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	5.52E-06	6.41E-05	0.00E+00	1.87E-03	1.87E-03
PWY-6123: inosine-5'-phosphate biosynthesis I|g_Bacteroides.s_Bacteroides_faecis	5.52E-06	6.41E-05	1.61E-03	0.00E+00	1.61E-03
PWY-6151: S-adenosyl-L-methionine cycle I|g_Bacteroides.s_Bacteroides_massiliensis	5.52E-06	6.41E-05	0.00E+00	3.27E-03	3.27E-03
PWY-6168: flavin biosynthesis III (fungi)|g_Bacteroides.s_Bacteroides_faecis	5.52E-06	6.41E-05	1.64E-03	0.00E+00	1.64E-03
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Alistipes.s_Alistipes_putredinis	5.52E-06	6.41E-05	0.00E+00	2.63E-03	2.63E-03
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Eubacterium.s_Eubacterium_ventriosum	5.52E-06	6.41E-05	1.10E-03	0.00E+00	1.10E-03
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	5.52E-06	6.41E-05	0.00E+00	1.87E-03	1.87E-03
PWY-6386: UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)|g_Bifidobacterium.s_Bifidobacterium_bifidum	5.52E-06	6.41E-05	6.30E-04	0.00E+00	6.30E-04
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)|g_Bifidobacterium.s_Bifidobacterium_bifidum	5.52E-06	6.41E-05	6.09E-04	0.00E+00	6.09E-04
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	5.52E-06	6.41E-05	0.00E+00	1.80E-03	1.80E-03
PWY-6527: stachyose degradation|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	5.52E-06	6.41E-05	0.00E+00	1.64E-03	1.64E-03
PWY-6700: queuosine biosynthesis|g_Bacteroides.s_Bacteroides_faecis	5.52E-06	6.41E-05	1.41E-03	0.00E+00	1.41E-03
PWY-6700: queuosine biosynthesis|g_Bacteroides.s_Bacteroides_massiliensis	5.52E-06	6.41E-05	0.00E+00	2.14E-03	2.14E-03
PWY-6700: queuosine biosynthesis|g_Dialister.s_Dialister_invisus	5.52E-06	6.41E-05	1.39E-03	0.00E+00	1.39E-03
PWY-6703: preQ0 biosynthesis|g_Bacteroides.s_Bacteroides_massiliensis	5.52E-06	6.41E-05	0.00E+00	1.93E-03	1.93E-03
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Bacteroides.s_Bacteroides_massiliensis	5.52E-06	6.41E-05	0.00E+00	1.79E-03	1.79E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	5.52E-06	6.41E-05	0.00E+00	2.37E-03	2.37E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_fragilis	5.52E-06	6.41E-05	6.76E-04	0.00E+00	6.76E-04
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Paraprevotella.s_Paraprevotella_clara	5.52E-06	6.41E-05	0.00E+00	3.16E-03	3.16E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Ruminococcus.s_Ruminococcus_callidus	5.52E-06	6.41E-05	0.00E+00	2.51E-03	2.51E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Parabacteroides.s_Parabacteroides_distasonis	5.52E-06	6.41E-05	0.00E+00	2.59E-03	2.59E-03
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Bacteroides.s_Bacteroides_coprocola	5.52E-06	6.41E-05	0.00E+00	2.45E-03	2.45E-03
SER-GLYSYN-PWY: superpathway of L-serine and glycine biosynthesis I|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	5.52E-06	6.41E-05	0.00E+00	1.98E-03	1.98E-03
UNINTEGRATED|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	5.52E-06	6.41E-05	0.00E+00	1.75E-03	1.75E-03
UNINTEGRATED|g_Alistipes.s_Alistipes_shahii	5.52E-06	6.41E-05	0.00E+00	1.85E-03	1.85E-03
UNINTEGRATED|g_Coprococcus.s_Coprococcus_comes	5.52E-06	6.41E-05	0.00E+00	1.23E-03	1.23E-03
VALSYN-PWY: L-valine biosynthesis|g_Bacteroides.s_Bacteroides_massiliensis	5.52E-06	6.41E-05	0.00E+00	1.79E-03	1.79E-03
VALSYN-PWY: L-valine biosynthesis|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	5.52E-06	6.41E-05	0.00E+00	1.71E-03	1.71E-03
GLUCONEO-PWY: gluconeogenesis I	5.52E-06	6.41E-05	0.00E+00	5.53E-03	5.53E-03
PWY-821: superpathway of sulfur amino acid biosynthesis (Saccharomyces cerevisiae)	5.52E-06	6.41E-05	3.48E-03	0.00E+00	3.48E-03
PYRIDOXSYN-PWY: pyridoxal 5'-phosphate biosynthesis I|g_Bacteroides.s_Bacteroides_stercoris	5.68E-06	6.57E-05	1.47E-03	4.86E-05	1.43E-03
PWY-3841: folate transformations II|g_Bacteroides.s_Bacteroides_stercoris	6.05E-06	6.86E-05	1.54E-03	4.04E-05	1.50E-03
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Bacteroides.s_Bacteroides_stercoris	6.05E-06	6.86E-05	1.46E-03	3.94E-05	1.42E-03
PWY-5686: UMP biosynthesis|g_Bacteroides.s_Bacteroides_stercoris	6.05E-06	6.86E-05	1.48E-03	4.63E-05	1.44E-03
PWY-7663: gondoate biosynthesis (anaerobic)|g_Bacteroides.s_Bacteroides_ovatus	6.05E-06	6.86E-05	2.01E-03	4.56E-05	1.96E-03
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Bacteroides.s_Bacteroides_stercoris	6.05E-06	6.86E-05	1.46E-03	3.94E-05	1.42E-03
PANTO-PWY: phosphopantothenate biosynthesis I|g_Bacteroides.s_Bacteroides_stercoris	6.32E-06	7.00E-05	1.53E-03	4.45E-05	1.48E-03
PWY-5097: L-lysine biosynthesis VI|g_Bacteroides.s_Bacteroides_stercoris	6.32E-06	7.00E-05	1.50E-03	4.22E-05	1.46E-03
PWY-6151: S-adenosyl-L-methionine cycle I|g_Bacteroides.s_Bacteroides_stercoris	6.32E-06	7.00E-05	1.39E-03	4.22E-05	1.35E-03
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Bacteroides.s_Bacteroides_xylanisolvens	6.32E-06	7.00E-05	1.33E-03	3.33E-05	1.30E-03
PWY-7663: gondoate biosynthesis (anaerobic)|g_Bacteroides.s_Bacteroides_stercoris	6.32E-06	7.00E-05	1.42E-03	4.51E-05	1.38E-03
VALSYN-PWY: L-valine biosynthesis|g_Bacteroides.s_Bacteroides_xylanisolvens	6.32E-06	7.00E-05	1.33E-03	3.33E-05	1.30E-03
PEPTIDOGLYCANSYN-PWY: peptidoglycan biosynthesis I (meso-diaminopimelate containing)|g_Bacteroides.s_Bacteroides_ovatus	6.51E-06	7.09E-05	1.56E-03	9.90E-05	1.46E-03
PWY-3841: folate transformations II|g_Bacteroides.s_Bacteroides_ovatus	6.51E-06	7.09E-05	1.87E-03	1.21E-04	1.75E-03
PWY-5686: UMP biosynthesis|g_Bacteroides.s_Bacteroides_ovatus	6.51E-06	7.09E-05	1.86E-03	1.31E-04	1.73E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY66-400: glycolysis VI (metazoan)|g_Bacteroides.s_Bacteroides_ovatus	6.51E-06	7.09E-05	1.82E-03	1.33E-04	1.68E-03
PWY-2942: L-lysine biosynthesis III|g_Bacteroides.s_Bacteroides_stercoris	6.63E-06	7.18E-05	1.50E-03	6.31E-05	1.43E-03
HISDEG-PWY: L-histidine degradation I	6.68E-06	7.18E-05	6.13E-03	7.56E-04	5.37E-03
PWY-6531: mannitol cycle	6.68E-06	7.18E-05	3.84E-03	7.79E-04	3.06E-03
PWY-7456: mannan degradation|g_Bacteroides.s_Bacteroides_ovatus	6.69E-06	7.18E-05	1.40E-03	5.33E-05	1.35E-03
GALACT-GLUCUROCAT-PWY: superpathway of hexuronide and hexuronate degradation	6.93E-06	7.32E-05	3.51E-03	8.34E-03	4.83E-03
PWY-1269: CMP-3-deoxy-D-manno-octulosonate biosynthesis I	6.93E-06	7.32E-05	5.19E-03	1.64E-03	3.55E-03
PWY-5101: L-isoleucine biosynthesis II	6.93E-06	7.32E-05	5.44E-03	1.93E-03	3.51E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis	6.93E-06	7.32E-05	5.18E-03	1.05E-02	5.32E-03
1CMET2-PWY: N10-formyl-tetrahydrofolate biosynthesis|g_Bacteroides.s_Bacteroides_uniformis	7.07E-06	7.42E-05	1.37E-03	1.17E-04	1.26E-03
PWY-6703: preQ0 biosynthesis|g_Bacteroides.s_Bacteroides_stercoris	7.07E-06	7.42E-05	1.58E-03	5.33E-05	1.53E-03
PYRIDOXSYN-PWY: pyridoxal 5'-phosphate biosynthesis I|g_Bacteroides.s_Bacteroides_uniformis	7.28E-06	7.61E-05	1.78E-03	1.45E-04	1.64E-03
PANTOSYN-PWY: pantothenate and coenzyme A biosynthesis I|g_Bacteroides.s_Bacteroides_ovatus	7.85E-06	8.14E-05	1.92E-03	1.24E-04	1.79E-03
PWY-1269: CMP-3-deoxy-D-manno-octulosonate biosynthesis I|g_Bacteroides.s_Bacteroides_ovatus	7.85E-06	8.14E-05	1.78E-03	1.03E-04	1.68E-03
ARGININE-SYN4-PWY: L-ornithine de novo biosynthesis|g_Bacteroides.s_Bacteroides_ovatus	8.08E-06	8.14E-05	1.82E-03	1.07E-04	1.71E-03
ARGININE-SYN4-PWY: L-ornithine de novo biosynthesis|g_Bacteroides.s_Bacteroides_stercoris	8.08E-06	8.14E-05	1.48E-03	7.63E-05	1.40E-03
PWY-6385: peptidoglycan biosynthesis III (mycobacteria)|g_Bacteroides.s_Bacteroides_ovatus	8.08E-06	8.14E-05	1.94E-03	1.39E-04	1.80E-03
PWY-6386: UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)|g_Bacteroides.s_Bacteroides_ovatus	8.08E-06	8.14E-05	1.54E-03	1.07E-04	1.43E-03
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)|g_Bacteroides.s_Bacteroides_ovatus	8.08E-06	8.14E-05	1.52E-03	9.56E-05	1.43E-03
PWY0-845: superpathway of pyridoxal 5'-phosphate biosynthesis and salvage|g_Bacteroides.s_Bacteroides_ovatus	8.08E-06	8.14E-05	2.17E-03	1.59E-04	2.01E-03
PYRIDOXSYN-PWY: pyridoxal 5'-phosphate biosynthesis I|g_Bacteroides.s_Bacteroides_ovatus	8.08E-06	8.14E-05	2.01E-03	1.67E-04	1.84E-03
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Bacteroides.s_Bacteroides_ovatus	8.22E-06	8.14E-05	1.79E-03	1.63E-04	1.63E-03
COA-PWY: coenzyme A biosynthesis I|g_Bacteroides.s_Bacteroides_ovatus	8.22E-06	8.14E-05	2.09E-03	1.90E-04	1.90E-03
PWY-5695: urate biosynthesis/inosine 5'-phosphate degradation|g_Bacteroides.s_Bacteroides_stercoris	8.22E-06	8.14E-05	1.53E-03	7.76E-05	1.45E-03
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Bacteroides.s_Bacteroides_stercoris	8.22E-06	8.14E-05	1.51E-03	7.65E-05	1.43E-03
THISYNARA-PWY: superpathway of thiamin diphosphate biosynthesis III (eukaryotes)|g_Bacteroides.s_Bacteroides_ovatus	8.22E-06	8.14E-05	2.10E-03	1.63E-04	1.94E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
VALSYN-PWY: L-valine biosynthesis|g_Bacteroides.s_Bacteroides_stercoris	8.22E-06	8.14E-05	1.51E-03	7.65E-05	1.43E-03
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Bacteroides.s_Bacteroides_vulgatus	8.29E-06	8.15E-05	1.88E-03	1.96E-04	1.68E-03
RHAMCAT-PWY: L-rhamnose degradation I|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	8.29E-06	8.15E-05	2.18E-04	2.31E-03	2.09E-03
NONMEVIPP-PWY: methylerythritol phosphate pathway I|g_Bacteroides.s_Bacteroides_ovatus	8.78E-06	8.54E-05	1.73E-03	1.10E-04	1.62E-03
PEPTIDOGLYCANSYN-PWY: peptidoglycan biosynthesis I (meso-diaminopimelate containing)|g_Bacteroides.s_Bacteroides_uniformis	8.78E-06	8.54E-05	1.40E-03	1.48E-04	1.25E-03
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)|g_Bacteroides.s_Bacteroides_uniformis	8.78E-06	8.54E-05	1.46E-03	1.45E-04	1.31E-03
PWY0-845: superpathway of pyridoxal 5'-phosphate biosynthesis and salvage|g_Bacteroides.s_Bacteroides_uniformis	9.06E-06	8.78E-05	1.69E-03	1.44E-04	1.55E-03
PWY-6700: queuosine biosynthesis|g_Bacteroides.s_Bacteroides_ovatus	9.33E-06	8.98E-05	1.93E-03	1.24E-04	1.80E-03
PWY-4984: urea cycle	9.33E-06	8.98E-05	5.50E-04	7.66E-03	7.11E-03
UNINTEGRATED|g_Bacteroides.s_Bacteroides_vulgatus	9.72E-06	9.24E-05	1.78E-03	3.83E-04	1.40E-03
ARGININE-SYN4-PWY: L-ornithine de novo biosynthesis	9.72E-06	9.24E-05	5.42E-03	1.75E-03	3.67E-03
PWY-5686: UMP biosynthesis	9.72E-06	9.24E-05	4.90E-03	1.06E-02	5.68E-03
ARGININE-SYN4-PWY: L-ornithine de novo biosynthesis|g_Bacteroides.s_Bacteroides_uniformis	9.73E-06	9.24E-05	1.45E-03	1.49E-04	1.30E-03
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Bifidobacterium.s_Bifidobacterium_longum	9.87E-06	9.31E-05	1.02E-03	8.69E-05	9.30E-04
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Bifidobacterium.s_Bifidobacterium_longum	9.87E-06	9.31E-05	1.02E-03	8.69E-05	9.30E-04
ANAGLYCOLYSIS-PWY: glycolysis III (from glucose)|g_Bacteroides.s_Bacteroides_ovatus	1.02E-05	9.44E-05	1.94E-03	2.05E-04	1.74E-03
PWY-1042: glycolysis IV (plant cytosol)|g_Bacteroides.s_Bacteroides_ovatus	1.02E-05	9.44E-05	1.75E-03	1.66E-04	1.59E-03
PWY-2942: L-lysine biosynthesis III|g_Bacteroides.s_Bacteroides_ovatus	1.02E-05	9.44E-05	1.99E-03	1.74E-04	1.82E-03
PWY-7282: 4-amino-2-methyl-5-phosphomethylpyrimidine biosynthesis (yeast)|g_Bacteroides.s_Bacteroides_ovatus	1.02E-05	9.44E-05	2.04E-03	1.60E-04	1.88E-03
PWY-7323: superpathway of GDP-mannose-derived O-antigen building blocks biosynthesis|g_Bacteroides.s_Bacteroides_ovatus	1.02E-05	9.44E-05	1.89E-03	1.48E-04	1.74E-03
PWY-6703: preQ0 biosynthesis|g_Bacteroides.s_Bacteroides_vulgatus	1.03E-05	9.49E-05	1.72E-03	1.83E-04	1.53E-03
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I|g_Bifidobacterium.s_Bifidobacterium_longum	1.04E-05	9.62E-05	9.49E-04	7.10E-05	8.78E-04
GLYCOGENSYNTH-PWY: glycogen biosynthesis I (from ADP-D-Glucose)|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	1.07E-05	9.81E-05	1.06E-03	7.87E-05	9.84E-04
PWY-6386: UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)|g_Bacteroides.s_Bacteroides_uniformis	1.09E-05	9.93E-05	1.38E-03	1.40E-04	1.24E-03
PWY-6609: adenine and adenosine salvage III|g_Bacteroides.s_Bacteroides_caccae	1.09E-05	9.93E-05	8.80E-04	8.03E-05	8.00E-04

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Bacteroides.s_Bacteroides_caccae	1.15E-05	1.04E-04	8.26E-04	7.41E-05	7.51E-04
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Bacteroides.s_Bacteroides_uniformis	1.15E-05	1.04E-04	1.47E-03	1.39E-04	1.33E-03
PWY-5686: UMP biosynthesis|g_Bacteroides.s_Bacteroides_uniformis	1.15E-05	1.04E-04	1.59E-03	1.76E-04	1.41E-03
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Bacteroides.s_Bacteroides_uniformis	1.15E-05	1.04E-04	1.47E-03	1.39E-04	1.33E-03
PWY-6123: inosine-5'-phosphate biosynthesis I|g_Bacteroides.s_Bacteroides_ovatus	1.24E-05	1.10E-04	1.71E-03	2.00E-04	1.51E-03
PWY-6700: queuosine biosynthesis|g_Bacteroides.s_Bacteroides_stercoris	1.24E-05	1.10E-04	1.69E-03	1.20E-04	1.57E-03
1CMET2-PWY: N10-formyl-tetrahydrofolate biosynthesis|g_Bacteroides.s_Bacteroides_caccae	1.24E-05	1.10E-04	8.24E-04	7.10E-05	7.53E-04
PWY-4242: pantothenate and coenzyme A biosynthesis III|g_Bacteroides.s_Bacteroides_ovatus	1.24E-05	1.10E-04	1.47E-03	1.22E-04	1.35E-03
PWY0-845: superpathway of pyridoxal 5'-phosphate biosynthesis and salvage|g_Bacteroides.s_Bacteroides_caccae	1.24E-05	1.10E-04	8.02E-04	6.79E-05	7.34E-04
PYRIDOXSYN-PWY: pyridoxal 5'-phosphate biosynthesis I|g_Bacteroides.s_Bacteroides_caccae	1.24E-05	1.10E-04	8.26E-04	7.65E-05	7.49E-04
ASPASN-PWY: superpathway of L-aspartate and L-asparagine biosynthesis|g_Bacteroides.s_Bacteroides_ovatus	1.26E-05	1.10E-04	1.76E-03	1.77E-04	1.58E-03
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_ovatus	1.26E-05	1.10E-04	1.92E-03	1.74E-04	1.74E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Bifidobacterium.s_Bifidobacterium_longum	1.30E-05	1.14E-04	8.99E-04	6.47E-05	8.35E-04
PWY-621: sucrose degradation III (sucrose invertase)|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	1.36E-05	1.18E-04	7.02E-04	2.07E-03	1.36E-03
UNINTEGRATED|g_Bacteroides.s_Bacteroides_caccae	1.36E-05	1.18E-04	1.04E-03	2.91E-04	7.45E-04
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Alistipes.s_Alistipes_finegoldii	1.39E-05	1.19E-04	5.44E-04	7.18E-06	5.36E-04
HOMOSER-METSYN-PWY: L-methionine biosynthesis I|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	1.39E-05	1.19E-04	8.31E-04	2.11E-05	8.10E-04
PWY-5659: GDP-mannose biosynthesis|g_Alistipes.s_Alistipes_finegoldii	1.39E-05	1.19E-04	5.79E-04	4.51E-06	5.74E-04
PWY-7199: pyrimidine deoxyribonucleosides salvage|g_Alistipes.s_Alistipes_finegoldii	1.39E-05	1.19E-04	5.35E-04	1.06E-05	5.24E-04
PWY-7371: 1,4-dihydroxy-6-naphthoate biosynthesis II|g_Alistipes.s_Alistipes_finegoldii	1.39E-05	1.19E-04	6.24E-04	8.69E-06	6.15E-04
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Bacteroides.s_Bacteroides_uniformis	1.53E-05	1.29E-04	1.36E-03	1.81E-04	1.18E-03
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Bacteroides.s_Bacteroides_ovatus	1.53E-05	1.29E-04	1.89E-03	2.12E-04	1.68E-03
PWY-7663: gondoate biosynthesis (anaerobic)|g_Bacteroides.s_Bacteroides_caccae	1.53E-05	1.29E-04	8.26E-04	9.52E-05	7.31E-04
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Bacteroides.s_Bacteroides_ovatus	1.53E-05	1.29E-04	1.89E-03	2.12E-04	1.68E-03
PWY-6703: preQ0 biosynthesis|g_Bacteroides.s_Bacteroides_caccae	1.54E-05	1.30E-04	8.86E-04	1.13E-04	7.73E-04

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-6609: adenine and adenosine salvage III|g_Bacteroides.s_Bacteroides_ovatus	1.55E-05	1.30E-04	1.99E-03	2.44E-04	1.75E-03
PWY-7323: superpathway of GDP-mannose-derived O-antigen building blocks biosynthesis	1.55E-05	1.30E-04	4.71E-03	8.90E-04	3.82E-03
PWY-6125: superpathway of guanosine nucleotides de novo biosynthesis II|g_Bacteroides.s_Bacteroides_vulgatus	1.56E-05	1.31E-04	1.72E-03	2.03E-04	1.52E-03
ARGSYNBSUB-PWY: L-arginine biosynthesis II (acetyl cycle)|g_Bifidobacterium.s_Bifidobacterium_longum	1.61E-05	1.32E-04	9.46E-04	1.00E-04	8.46E-04
UDPNAGSYN-PWY: UDP-N-acetyl-D-glucosamine biosynthesis I|g_Bifidobacterium.s_Bifidobacterium_longum	1.61E-05	1.32E-04	7.57E-04	6.46E-05	6.92E-04
PWY-6737: starch degradation V|g_Alistipes.s_Alistipes_finegoldii	1.64E-05	1.32E-04	5.95E-04	8.14E-06	5.86E-04
PWY-6124: inosine-5'-phosphate biosynthesis II|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	1.65E-05	1.32E-04	9.01E-04	6.97E-05	8.31E-04
PWY-6700: queuosine biosynthesis|g_Bacteroides.s_Bacteroides_uniformis	1.66E-05	1.32E-04	1.31E-03	1.24E-04	1.19E-03
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_caccae	1.66E-05	1.32E-04	8.32E-04	9.02E-05	7.41E-04
COA-PWY: coenzyme A biosynthesis I|g_Bacteroides.s_Bacteroides_xylanisolvens	1.72E-05	1.32E-04	1.64E-03	2.05E-04	1.44E-03
PWY-5686: UMP biosynthesis|g_Alistipes.s_Alistipes_finegoldii	1.73E-05	1.32E-04	6.67E-04	8.75E-06	6.59E-04
PWY-6151: S-adenosyl-L-methionine cycle I|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	1.73E-05	1.32E-04	8.52E-04	4.17E-05	8.10E-04
PWY-3841: folate transformations II|g_Bacteroides.s_Bacteroides_uniformis	1.76E-05	1.32E-04	1.27E-03	1.40E-04	1.13E-03
PWY-5695: urate biosynthesis/inosine 5'-phosphate degradation|g_Bacteroides.s_Bacteroides_caccae	1.83E-05	1.32E-04	8.57E-04	1.13E-04	7.44E-04
PWY0-1586: peptidoglycan maturation (meso-diaminopimelate containing)|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	1.86E-05	1.32E-04	1.95E-04	2.08E-03	1.89E-03
CITRULBIO-PWY: L-citrulline biosynthesis	1.88E-05	1.32E-04	1.20E-03	7.67E-03	6.47E-03
PHOSLIPSYN-PWY: superpathway of phospholipid biosynthesis I (bacteria)	1.88E-05	1.32E-04	6.38E-03	1.16E-03	5.22E-03
PWY4FS-7: phosphatidylglycerol biosynthesis I (plastidic)	1.88E-05	1.32E-04	6.33E-03	1.18E-03	5.15E-03
PWY4FS-8: phosphatidylglycerol biosynthesis II (non-plastidic)	1.88E-05	1.32E-04	6.33E-03	1.18E-03	5.15E-03
ARGININE-SYN4-PWY: L-ornithine de novo biosynthesis|g_Bacteroides.s_Bacteroides_fragilis	1.88E-05	1.32E-04	7.11E-04	0.00E+00	7.11E-04
HISDEG-PWY: L-histidine degradation I|g_Alistipes.s_Alistipes_shahii	1.88E-05	1.32E-04	0.00E+00	1.47E-03	1.47E-03
NONMEVIPP-PWY: methylerythritol phosphate pathway I|g_Alistipes.s_Alistipes_putredinis	1.88E-05	1.32E-04	0.00E+00	2.90E-03	2.90E-03
NONMEVIPP-PWY: methylerythritol phosphate pathway I|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.88E-05	1.32E-04	0.00E+00	1.70E-03	1.70E-03
PANTO-PWY: phosphopantothenate biosynthesis I|g_Alistipes.s_Alistipes_putredinis	1.88E-05	1.32E-04	0.00E+00	2.05E-03	2.05E-03
PANTO-PWY: phosphopantothenate biosynthesis I|g_Eubacterium.s_Eubacterium_ventriosum	1.88E-05	1.32E-04	1.12E-03	0.00E+00	1.12E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PANTO-PWY: phosphopantothenate biosynthesis I|g_Paraprevotella.s_Paraprevotella_clara	1.88E-05	1.32E-04	0.00E+00	2.22E-03	2.22E-03
PEPTIDOGLYCANSYN-PWY: peptidoglycan biosynthesis I (meso-diaminopimelate containing)|g_Alistipes.s_Alistipes_putredinis	1.88E-05	1.32E-04	0.00E+00	2.38E-03	2.38E-03
PEPTIDOGLYCANSYN-PWY: peptidoglycan biosynthesis I (meso-diaminopimelate containing)|g_Parabacteroides.s_Parabacteroides_merdae	1.88E-05	1.32E-04	0.00E+00	1.87E-03	1.87E-03
PWY-2942: L-lysine biosynthesis III|g_Parabacteroides.s_Parabacteroides_distasonis	1.88E-05	1.32E-04	0.00E+00	1.86E-03	1.86E-03
PWY-5097: L-lysine biosynthesis VI|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.88E-05	1.32E-04	0.00E+00	1.75E-03	1.75E-03
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Parabacteroides.s_Parabacteroides_merdae	1.88E-05	1.32E-04	0.00E+00	1.88E-03	1.88E-03
PWY-5686: UMP biosynthesis|g_Bacteroides.s_Bacteroides_massiliensis	1.88E-05	1.32E-04	0.00E+00	1.68E-03	1.68E-03
PWY-5686: UMP biosynthesis|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.88E-05	1.32E-04	0.00E+00	1.98E-03	1.98E-03
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.88E-05	1.32E-04	0.00E+00	1.88E-03	1.88E-03
PWY-6386: UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)|g_Alistipes.s_Alistipes_putredinis	1.88E-05	1.32E-04	0.00E+00	2.44E-03	2.44E-03
PWY-6386: UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)|g_Parabacteroides.s_Parabacteroides_merdae	1.88E-05	1.32E-04	0.00E+00	2.10E-03	2.10E-03
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)|g_Alistipes.s_Alistipes_putredinis	1.88E-05	1.32E-04	0.00E+00	2.45E-03	2.45E-03
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)|g_Parabacteroides.s_Parabacteroides_merdae	1.88E-05	1.32E-04	0.00E+00	2.13E-03	2.13E-03
PWY-6700: queuosine biosynthesis|g_Parabacteroides.s_Parabacteroides_distasonis	1.88E-05	1.32E-04	0.00E+00	2.39E-03	2.39E-03
PWY-6700: queuosine biosynthesis|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.88E-05	1.32E-04	0.00E+00	1.60E-03	1.60E-03
PWY-6737: starch degradation V|g_Catenibacterium.s_Catenibacterium_mitsuokai	1.88E-05	1.32E-04	0.00E+00	5.76E-04	5.76E-04
PWY-6737: starch degradation V|g_Coprococcus.s_Coprococcus_comes	1.88E-05	1.32E-04	0.00E+00	1.46E-03	1.46E-03
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Catenibacterium.s_Catenibacterium_mitsuokai	1.88E-05	1.32E-04	0.00E+00	5.13E-04	5.13E-04
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Parabacteroides.s_Parabacteroides_distasonis	1.88E-05	1.32E-04	0.00E+00	1.68E-03	1.68E-03
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Paraprevotella.s_Paraprevotella_clara	1.88E-05	1.32E-04	0.00E+00	1.55E-03	1.55E-03
PWY-7208: superpathway of pyrimidine nucleobases salvage|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.88E-05	1.32E-04	0.00E+00	2.03E-03	2.03E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Erysipelotrichaceae_noname.s_Eubacterium_biforme	1.88E-05	1.32E-04	0.00E+00	7.73E-04	7.73E-04
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Alistipes.s_Alistipes_putredinis	1.88E-05	1.32E-04	0.00E+00	1.80E-03	1.80E-03
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.88E-05	1.32E-04	0.00E+00	2.11E-03	2.11E-03
PWY-7228: superpathway of guanosine nucleotides de novo biosynthesis I|g_Bacteroides.s_Bacteroides_massiliensis	1.88E-05	1.32E-04	0.00E+00	1.87E-03	1.87E-03
PWY-7237: myo-, chiro- and scillo-inositol degradation|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	1.88E-05	1.32E-04	1.00E-03	0.00E+00	1.00E-03
PWY-724: superpathway of L-lysine, L-threonine and L-methionine biosynthesis II|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.88E-05	1.32E-04	0.00E+00	1.79E-03	1.79E-03
PWY-7357: thiamin formation from pyrithiamine and oxythiamine (yeast)|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.88E-05	1.32E-04	0.00E+00	1.70E-03	1.70E-03
PWY-7560: methylerythritol phosphate pathway II|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.88E-05	1.32E-04	0.00E+00	1.63E-03	1.63E-03
PWY-7663: gondoate biosynthesis (anaerobic)|g_Bacteroides.s_Bacteroides_fragilis	1.88E-05	1.32E-04	7.24E-04	0.00E+00	7.24E-04
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Parabacteroides.s_Parabacteroides_merdae	1.88E-05	1.32E-04	0.00E+00	1.88E-03	1.88E-03
RHAMCAT-PWY: L-rhamnose degradation I|g_Eubacterium.s_Eubacterium_ventriosum	1.88E-05	1.32E-04	1.30E-03	0.00E+00	1.30E-03
THISYNARA-PWY: superpathway of thiamin diphosphate biosynthesis III (eukaryotes)|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.88E-05	1.32E-04	0.00E+00	1.73E-03	1.73E-03
THRESYN-PWY: superpathway of L-threonine biosynthesis|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.88E-05	1.32E-04	0.00E+00	1.73E-03	1.73E-03
TRPSYN-PWY: L-tryptophan biosynthesis|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.88E-05	1.32E-04	0.00E+00	1.53E-03	1.53E-03
UNINTEGRATED|g_Catenibacterium.s_Catenibacterium_mitsuokai	1.88E-05	1.32E-04	0.00E+00	6.25E-04	6.25E-04
UNINTEGRATED|g_Erysipelotrichaceae_noname.s_Eubacterium_biforme	1.88E-05	1.32E-04	0.00E+00	7.88E-04	7.88E-04
UNINTEGRATED|g_Ruminococcus.s_Ruminococcus_lactaris	1.88E-05	1.32E-04	0.00E+00	1.11E-03	1.11E-03
UNINTEGRATED|g_Veillonella.s_Veillonella_parvula	1.88E-05	1.32E-04	5.98E-04	0.00E+00	5.98E-04
VALSYN-PWY: L-valine biosynthesis|g_Catenibacterium.s_Catenibacterium_mitsuokai	1.88E-05	1.32E-04	0.00E+00	5.13E-04	5.13E-04
VALSYN-PWY: L-valine biosynthesis|g_Parabacteroides.s_Parabacteroides_distasonis	1.88E-05	1.32E-04	0.00E+00	1.68E-03	1.68E-03
VALSYN-PWY: L-valine biosynthesis|g_Paraprevotella.s_Paraprevotella_clara	1.88E-05	1.32E-04	0.00E+00	1.55E-03	1.55E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
FUC-RHAMCAT-PWY: superpathway of fucose and rhamnose degradation	1.88E-05	1.32E-04	3.23E-03	0.00E+00	3.23E-03
P162-PWY: L-glutamate degradation V (via hydroxyglutarate)	1.88E-05	1.32E-04	2.54E-03	0.00E+00	2.54E-03
PWY-6124: inosine-5'-phosphate biosynthesis II|g_Bacteroides.s_Bacteroides_ovatus	1.92E-05	1.34E-04	1.70E-03	2.52E-04	1.45E-03
PANTO-PWY: phosphopantothenate biosynthesis I|g_Alistipes.s_Alistipes_finegoldii	2.03E-05	1.41E-04	6.67E-04	1.10E-05	6.56E-04
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I|g_Alistipes.s_Alistipes_finegoldii	2.03E-05	1.41E-04	5.44E-04	9.80E-06	5.34E-04
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Alistipes.s_Alistipes_finegoldii	2.03E-05	1.41E-04	5.54E-04	9.62E-06	5.44E-04
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Alistipes.s_Alistipes_finegoldii	2.03E-05	1.41E-04	5.54E-04	9.62E-06	5.44E-04
PWY-6700: queuosine biosynthesis|g_Bacteroides.s_Bacteroides_caccae	2.12E-05	1.46E-04	8.06E-04	7.69E-05	7.29E-04
UNINTEGRATED|g_Bifidobacterium.s_Bifidobacterium_longum	2.12E-05	1.46E-04	1.11E-03	1.47E-04	9.63E-04
PWY-5686: UMP biosynthesis|g_Bacteroides.s_Bacteroides_caccae	2.25E-05	1.54E-04	8.13E-04	1.08E-04	7.05E-04
PWY-5695: urate biosynthesis/inosine 5'-phosphate degradation|g_Bacteroides.s_Bacteroides_ovatus	2.31E-05	1.58E-04	1.85E-03	2.17E-04	1.63E-03
ILEUSYN-PWY: L-isoleucine biosynthesis I (from threonine)|g_Bifidobacterium.s_Bifidobacterium_longum	2.32E-05	1.58E-04	1.08E-03	1.16E-04	9.65E-04
ARGININE-SYN4-PWY: L-ornithine de novo biosynthesis|g_Bacteroides.s_Bacteroides_caccae	2.34E-05	1.59E-04	9.41E-04	1.17E-04	8.24E-04
NONMEVIPP-PWY: methylerythritol phosphate pathway I|g_Bacteroides.s_Bacteroides_vulgatus	2.34E-05	1.59E-04	1.80E-03	2.31E-04	1.57E-03
COBALSYN-PWY: adenosylcobalamin salvage from cobinamide I|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	2.36E-05	1.60E-04	3.66E-04	2.18E-03	1.82E-03
PANTO-PWY: phosphopantothenate biosynthesis I|g_Bacteroides.s_Bacteroides_ovatus	2.36E-05	1.60E-04	2.11E-03	2.29E-04	1.88E-03
ARGSYN-PWY: L-arginine biosynthesis I (via L-ornithine)|g_Bifidobacterium.s_Bifidobacterium_longum	2.46E-05	1.66E-04	9.62E-04	1.03E-04	8.60E-04
PWY-7400: L-arginine biosynthesis IV (archaebacteria)|g_Bifidobacterium.s_Bifidobacterium_longum	2.46E-05	1.66E-04	9.64E-04	1.02E-04	8.62E-04
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Alistipes.s_Alistipes_finegoldii	2.52E-05	1.69E-04	6.19E-04	1.13E-05	6.07E-04
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Alistipes.s_Alistipes_finegoldii	2.52E-05	1.69E-04	5.39E-04	1.29E-05	5.26E-04
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Alistipes.s_Alistipes_finegoldii	2.52E-05	1.69E-04	6.19E-04	1.13E-05	6.07E-04
RHAMCAT-PWY: L-rhamnose degradation I|g_Bacteroides.s_Bacteroides_ovatus	2.53E-05	1.69E-04	1.93E-03	3.92E-04	1.54E-03
HSERMETANA-PWY: L-methionine biosynthesis III	2.53E-05	1.69E-04	2.22E-03	7.66E-03	5.45E-03
PWY-3841: folate transformations II|g_Bacteroides.s_Bacteroides_caccae	2.65E-05	1.75E-04	8.69E-04	1.26E-04	7.42E-04
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Bacteroides.s_Bacteroides_caccae	2.65E-05	1.75E-04	8.84E-04	1.05E-04	7.79E-04
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Bacteroides.s_Bacteroides_caccae	2.65E-05	1.75E-04	8.84E-04	1.05E-04	7.79E-04

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Bifidobacterium.s_Bifidobacterium_longum	2.85E-05	1.88E-04	1.07E-03	1.28E-04	9.43E-04
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	2.89E-05	1.90E-04	3.89E-04	2.18E-03	1.79E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Alistipes.s_Alistipes_finegoldii	3.12E-05	2.05E-04	5.84E-04	1.61E-05	5.68E-04
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	3.13E-05	2.05E-04	1.00E-03	8.95E-05	9.15E-04
VALSYN-PWY: L-valine biosynthesis|g_Bifidobacterium.s_Bifidobacterium_longum	3.21E-05	2.09E-04	1.08E-03	1.41E-04	9.40E-04
P161-PWY: acetylene degradation	3.33E-05	2.17E-04	2.20E-03	1.24E-04	2.07E-03
PWY-5484: glycolysis II (from fructose 6-phosphate)	3.40E-05	2.20E-04	5.77E-03	2.08E-03	3.68E-03
UNMAPPED	3.40E-05	2.20E-04	5.48E-03	1.69E-02	1.14E-02
PWY-5097: L-lysine biosynthesis VI|g_Bacteroides.s_Bacteroides_caccae	3.46E-05	2.24E-04	8.55E-04	1.42E-04	7.13E-04
PWY-1042: glycolysis IV (plant cytosol)|g_Bacteroides.s_Bacteroides_caccae	4.12E-05	2.66E-04	8.67E-04	1.18E-04	7.49E-04
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_dorei	4.22E-05	2.71E-04	8.42E-05	1.97E-03	1.88E-03
PWY-6936: seleno-amino acid biosynthesis|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	4.23E-05	2.71E-04	8.57E-04	1.62E-05	8.41E-04
PYRIDNUCSYN-PWY: NAD biosynthesis I (from aspartate)|g_Alistipes.s_Alistipes_finegoldii	4.23E-05	2.71E-04	6.16E-04	8.53E-06	6.07E-04
COA-PWY: coenzyme A biosynthesis I|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	4.28E-05	2.74E-04	3.40E-04	2.19E-03	1.85E-03
UNINTEGRATED|g_Dorea.s_Dorea_longicatena	4.53E-05	2.88E-04	5.78E-04	1.46E-03	8.79E-04
PWY-5097: L-lysine biosynthesis VI	4.53E-05	2.88E-04	5.17E-03	1.07E-02	5.50E-03
PWY0-845: superpathway of pyridoxal 5'-phosphate biosynthesis and salvage	4.53E-05	2.88E-04	5.78E-03	2.49E-03	3.29E-03
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Bifidobacterium.s_Bifidobacterium_longum	4.58E-05	2.90E-04	1.04E-03	1.18E-04	9.20E-04
P461-PWY: hexitol fermentation to lactate, formate, ethanol and acetate|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	4.74E-05	3.00E-04	9.95E-04	1.41E-04	8.54E-04
PWY-5097: L-lysine biosynthesis VI|g_Bacteroides.s_Bacteroides_ovatus	5.02E-05	3.16E-04	1.64E-03	2.13E-04	1.43E-03
PWY-6703: preQ0 biosynthesis|g_Bacteroides.s_Bacteroides_ovatus	5.02E-05	3.16E-04	1.39E-03	2.26E-04	1.17E-03
PANTO-PWY: phosphopantothenate biosynthesis I|g_Bacteroides.s_Bacteroides_caccae	5.22E-05	3.25E-04	8.60E-04	1.74E-04	6.86E-04
ARO-PWY: chorismate biosynthesis I|g_Alistipes.s_Alistipes_finegoldii	5.24E-05	3.25E-04	5.75E-04	6.34E-06	5.69E-04
PWY-6163: chorismate biosynthesis from 3-dehydroquinate|g_Alistipes.s_Alistipes_finegoldii	5.24E-05	3.25E-04	5.57E-04	5.60E-06	5.51E-04
UNINTEGRATED|g_Alistipes.s_Alistipes_finegoldii	5.25E-05	3.25E-04	6.58E-04	3.03E-05	6.28E-04
PWY-6700: queuosine biosynthesis|g_Bacteroides.s_Bacteroides_vulgatus	5.26E-05	3.25E-04	1.73E-03	2.44E-04	1.49E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-4242: pantothenate and coenzyme A biosynthesis III|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	5.61E-05	3.25E-04	8.22E-05	2.05E-03	1.97E-03
DTDPRHAMSYN-PWY: dTDP-L-rhamnose biosynthesis I|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	5.82E-05	3.25E-04	7.41E-04	9.12E-05	6.50E-04
PWY-7371: 1,4-dihydroxy-6-naphthoate biosynthesis II	5.84E-05	3.25E-04	1.63E-03	1.14E-04	1.52E-03
1CMET2-PWY: N10-formyl-tetrahydrofolate biosynthesis|g_Bacteroides.s_Bacteroides_massiliensis	5.96E-05	3.25E-04	0.00E+00	1.77E-03	1.77E-03
ARO-PWY: chorismate biosynthesis I|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	5.96E-05	3.25E-04	0.00E+00	1.76E-03	1.76E-03
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Alistipes.s_Alistipes_putredinis	5.96E-05	3.25E-04	0.00E+00	1.90E-03	1.90E-03
COA-PWY: coenzyme A biosynthesis I|g_Eubacterium.s_Eubacterium_ventriosum	5.96E-05	3.25E-04	9.11E-04	0.00E+00	9.11E-04
COBALSYN-PWY: adenosylcobalamin salvage from cobinamide I|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	5.96E-05	3.25E-04	0.00E+00	1.97E-03	1.97E-03
COMPLETE-ARO-PWY: superpathway of aromatic amino acid biosynthesis|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	5.96E-05	3.25E-04	0.00E+00	1.77E-03	1.77E-03
DTDPRHAMSYN-PWY: dTDP-L-rhamnose biosynthesis I|g_Alistipes.s_Alistipes_shahii	5.96E-05	3.25E-04	0.00E+00	1.95E-03	1.95E-03
PANTO-PWY: phosphopantothenate biosynthesis I|g_Bacteroides.s_Bacteroides_dorei	5.96E-05	3.25E-04	0.00E+00	1.92E-03	1.92E-03
PANTO-PWY: phosphopantothenate biosynthesis I|g_Parabacteroides.s_Parabacteroides_distasonis	5.96E-05	3.25E-04	0.00E+00	1.31E-03	1.31E-03
PANTOSYN-PWY: pantothenate and coenzyme A biosynthesis I|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	5.96E-05	3.25E-04	0.00E+00	2.01E-03	2.01E-03
PEPTIDOGLYCANSYN-PWY: peptidoglycan biosynthesis I (meso-diaminopimelate containing)|g_Bacteroides.s_Bacteroides_fragilis	5.96E-05	3.25E-04	8.75E-04	0.00E+00	8.75E-04
PEPTIDOGLYCANSYN-PWY: peptidoglycan biosynthesis I (meso-diaminopimelate containing)|g_Eubacterium.s_Eubacterium_ventriosum	5.96E-05	3.25E-04	1.13E-03	0.00E+00	1.13E-03
PWY-2942: L-lysine biosynthesis III|g_Paraprevotella.s_Paraprevotella_clara	5.96E-05	3.25E-04	0.00E+00	1.85E-03	1.85E-03
PWY-3841: folate transformations II|g_Bacteroides.s_Bacteroides_fragilis	5.96E-05	3.25E-04	8.36E-04	0.00E+00	8.36E-04
PWY-3841: folate transformations II|g_Parabacteroides.s_Parabacteroides_distasonis	5.96E-05	3.25E-04	0.00E+00	1.54E-03	1.54E-03
PWY-4041: γ-glutamyl cycle|g_Escherichia.s_Escherichia_coli	5.96E-05	3.25E-04	5.92E-04	0.00E+00	5.92E-04
PWY-5097: L-lysine biosynthesis VI|g_Bacteroides.s_Bacteroides_fragilis	5.96E-05	3.25E-04	7.49E-04	0.00E+00	7.49E-04
PWY-5097: L-lysine biosynthesis VI|g_Parabacteroides.s_Parabacteroides_distasonis	5.96E-05	3.25E-04	0.00E+00	1.39E-03	1.39E-03
PWY-5097: L-lysine biosynthesis VI|g_Paraprevotella.s_Paraprevotella_clara	5.96E-05	3.25E-04	0.00E+00	1.87E-03	1.87E-03
PWY-5188: tetrapyrrole biosynthesis I (from glutamate)|g_Ruminococcus.s_Ruminococcus_callidus	5.96E-05	3.25E-04	0.00E+00	1.99E-03	1.99E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-5659: GDP-mannose biosynthesis|g_Alistipes.s_Alistipes_shahii	5.96E-05	3.25E-04	0.00E+00	1.59E-03	1.59E-03
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Paraprevotella.s_Paraprevotella_clara	5.96E-05	3.25E-04	0.00E+00	1.94E-03	1.94E-03
PWY-5695: urate biosynthesis/inosine 5'-phosphate degradation|g_Bacteroides.s_Bacteroides_fragilis	5.96E-05	3.25E-04	4.52E-04	0.00E+00	4.52E-04
PWY-5695: urate biosynthesis/inosine 5'-phosphate degradation|g_Parabacteroides.s_Parabacteroides_distasonis	5.96E-05	3.25E-04	0.00E+00	1.02E-03	1.02E-03
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	5.96E-05	3.25E-04	0.00E+00	1.79E-03	1.79E-03
PWY-6124: inosine-5'-phosphate biosynthesis II|g_Bacteroides.s_Bacteroides_fragilis	5.96E-05	3.25E-04	9.45E-04	0.00E+00	9.45E-04
PWY-6151: S-adenosyl-L-methionine cycle I|g_Erysipelotrichaceae_noname.s_Eubacterium_biforme	5.96E-05	3.25E-04	0.00E+00	7.09E-04	7.09E-04
PWY-6151: S-adenosyl-L-methionine cycle I|g_Eubacterium.s_Eubacterium_ventriosum	5.96E-05	3.25E-04	1.05E-03	0.00E+00	1.05E-03
PWY-6163: chorismate biosynthesis from 3-dehydroquinate|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	5.96E-05	3.25E-04	0.00E+00	1.77E-03	1.77E-03
PWY-6270: isoprene biosynthesis I|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	5.96E-05	3.25E-04	0.00E+00	1.61E-03	1.61E-03
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	5.96E-05	3.25E-04	0.00E+00	1.79E-03	1.79E-03
PWY-6386: UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)|g_Paraprevotella.s_Paraprevotella_clara	5.96E-05	3.25E-04	0.00E+00	1.61E-03	1.61E-03
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)|g_Bacteroides.s_Bacteroides_fragilis	5.96E-05	3.25E-04	9.60E-04	0.00E+00	9.60E-04
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)|g_Eubacterium.s_Eubacterium_ventriosum	5.96E-05	3.25E-04	1.19E-03	0.00E+00	1.19E-03
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)|g_Paraprevotella.s_Paraprevotella_clara	5.96E-05	3.25E-04	0.00E+00	1.57E-03	1.57E-03
PWY-6609: adenine and adenosine salvage III|g_Bacteroides.s_Bacteroides_fragilis	5.96E-05	3.25E-04	6.49E-04	0.00E+00	6.49E-04
PWY-6703: preQ0 biosynthesis|g_Parabacteroides.s_Parabacteroides_distasonis	5.96E-05	3.25E-04	0.00E+00	2.22E-03	2.22E-03
PWY-6703: preQ0 biosynthesis|g_Parabacteroides.s_Parabacteroides_merdae	5.96E-05	3.25E-04	0.00E+00	1.09E-03	1.09E-03
PWY-6737: starch degradation V|g_Blautia.s_Ruminococcus_gnavus	5.96E-05	3.25E-04	1.02E-03	0.00E+00	1.02E-03
PWY-6737: starch degradation V|g_Eubacterium.s_Eubacterium_ventriosum	5.96E-05	3.25E-04	9.22E-04	0.00E+00	9.22E-04
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Coprococcus.s_Coprococcus_comes	5.96E-05	3.25E-04	0.00E+00	9.71E-04	9.71E-04
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Eubacterium.s_Eubacterium_ventriosum	5.96E-05	3.25E-04	9.15E-04	0.00E+00	9.15E-04
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Blautia.s_Ruminococcus_gnavus	5.96E-05	3.25E-04	7.62E-04	0.00E+00	7.62E-04

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Catenibacterium.s_Catenibacterium_mitsuokai	5.96E-05	3.25E-04	0.00E+00	6.03E-04	6.03E-04
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Coprococcus.s_Coprococcus_comes	5.96E-05	3.25E-04	0.00E+00	9.57E-04	9.57E-04
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Fusobacterium.s_Fusobacterium_varium	5.96E-05	3.25E-04	5.08E-04	0.00E+00	5.08E-04
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Parabacteroides.s_Parabacteroides_distasonis	5.96E-05	3.25E-04	0.00E+00	1.25E-03	1.25E-03
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_fragilis	5.96E-05	3.25E-04	7.17E-04	0.00E+00	7.17E-04
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Paraprevotella.s_Paraprevotella_clara	5.96E-05	3.25E-04	0.00E+00	2.33E-03	2.33E-03
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Ruminococcus.s_Ruminococcus_callidus	5.96E-05	3.25E-04	0.00E+00	1.58E-03	1.58E-03
PWY0-1296: purine ribonucleosides degradation|g_Blautia.s_Ruminococcus_gnavus	5.96E-05	3.25E-04	8.09E-04	0.00E+00	8.09E-04
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Paraprevotella.s_Paraprevotella_clara	5.96E-05	3.25E-04	0.00E+00	1.94E-03	1.94E-03
UNINTEGRATED|g_Barnesiella.s_Barnesiella_intestinihominis	5.96E-05	3.25E-04	0.00E+00	1.42E-03	1.42E-03
UNINTEGRATED|g_Fusobacterium.s_Fusobacterium_varium	5.96E-05	3.25E-04	4.69E-04	0.00E+00	4.69E-04
GLYCOLYSIS-TCA-GLYOX-BYPASS: superpathway of glycolysis, pyruvate dehydrogenase, TCA, and glyoxylate bypass	5.96E-05	3.25E-04	1.67E-03	0.00E+00	1.67E-03
PWY-4041: γ-glutamyl cycle	5.96E-05	3.25E-04	1.74E-03	0.00E+00	1.74E-03
TCA-GLYOX-BYPASS: superpathway of glyoxylate bypass and TCA	5.96E-05	3.25E-04	1.60E-03	0.00E+00	1.60E-03
PWY-6527: stachyose degradation|g_Dorea.s_Dorea_longicatena	5.97E-05	3.25E-04	3.65E-04	1.14E-03	7.78E-04
UNINTEGRATED|g_Bacteroides.s_Bacteroides_ovatus	5.97E-05	3.25E-04	2.07E-03	6.48E-04	1.42E-03
ANAEROFRUCAT-PWY: homolactic fermentation	5.97E-05	3.25E-04	6.62E-03	3.48E-03	3.14E-03
GLYCOLYSIS: glycolysis I (from glucose 6-phosphate)	5.97E-05	3.25E-04	5.79E-03	2.17E-03	3.61E-03
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I	5.97E-05	3.25E-04	5.48E-03	7.98E-03	2.50E-03
PWY-6317: galactose degradation I (Leloir pathway)	5.97E-05	3.25E-04	3.73E-03	6.42E-03	2.68E-03
PWY-6703: preQ0 biosynthesis|g_Bacteroides.s_Bacteroides_uniformis	6.40E-05	3.47E-04	1.50E-03	2.04E-04	1.30E-03
NONOXIPENT-PWY: pentose phosphate pathway (non-oxidative branch)|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	7.12E-05	3.85E-04	9.81E-04	1.36E-04	8.44E-04
PWY-6737: starch degradation V|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	7.12E-05	3.85E-04	1.16E-03	1.42E-04	1.02E-03
DTDPRHAMSYN-PWY: dTDP-L-rhamnose biosynthesis I|g_Bacteroides.s_Bacteroides_caccae	7.42E-05	4.00E-04	9.60E-04	1.69E-04	7.92E-04
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Bacteroides.s_Bacteroides_caccae	7.75E-05	4.16E-04	8.69E-04	1.91E-04	6.78E-04
RHAMCAT-PWY: L-rhamnose degradation I|g_Bacteroides.s_Bacteroides_caccae	7.75E-05	4.16E-04	8.17E-04	1.75E-04	6.42E-04
VALSYN-PWY: L-valine biosynthesis|g_Bacteroides.s_Bacteroides_caccae	7.75E-05	4.16E-04	8.69E-04	1.91E-04	6.78E-04

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
GLUTORN-PWY: L-ornithine biosynthesis|g_Bifidobacterium.s_Bifidobacterium_longum	7.82E-05	4.18E-04	8.06E-04	1.15E-04	6.92E-04
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_caccae	7.82E-05	4.18E-04	9.63E-04	3.00E-04	6.63E-04
DENOVOPURINE2-PWY: superpathway of purine nucleotides de novo biosynthesis II	7.90E-05	4.21E-04	7.24E-04	6.41E-03	5.69E-03
PWY-7539: 6-hydroxymethyl-dihydropterin diphosphate biosynthesis III (Chlamydia)|g_Bacteroides.s_Bacteroides_xylanisolvens	8.29E-05	4.41E-04	1.05E-03	1.32E-04	9.21E-04
UNINTEGRATED|g_Bacteroides.s_Bacteroides_thetaiotaomicron	8.36E-05	4.44E-04	2.91E-04	2.52E-03	2.23E-03
PWY-1042: glycolysis IV (plant cytosol)|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	8.77E-05	4.65E-04	1.22E-03	1.48E-04	1.08E-03
PWY-2942: L-lysine biosynthesis III|g_Bacteroides.s_Bacteroides_caccae	9.32E-05	4.93E-04	8.28E-04	1.62E-04	6.67E-04
COBALSYN-PWY: adenosylcobalamin salvage from cobinamide I	1.01E-04	5.35E-04	3.52E-03	6.54E-03	3.01E-03
PWY-7221: guanosine ribonucleotides de novo biosynthesis	1.01E-04	5.35E-04	5.10E-03	1.05E-02	5.35E-03
PEPTIDOGLYCANSYN-PWY: peptidoglycan biosynthesis I (meso-diaminopimelate containing)|g_Alistipes.s_Alistipes_finegoldii	1.05E-04	5.48E-04	5.70E-04	1.17E-05	5.59E-04
PWY-6386: UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)|g_Alistipes.s_Alistipes_finegoldii	1.05E-04	5.48E-04	5.77E-04	1.18E-05	5.65E-04
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)|g_Alistipes.s_Alistipes_finegoldii	1.05E-04	5.48E-04	5.61E-04	1.09E-05	5.50E-04
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Roseburia.s_Roseburia_inulinivorans	1.05E-04	5.48E-04	5.10E-04	1.50E-05	4.95E-04
PWY-621: sucrose degradation III (sucrose invertase)|g_Eubacterium.s_Eubacterium_rectale	1.05E-04	5.50E-04	8.26E-04	6.77E-05	7.59E-04
UNINTEGRATED|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	1.06E-04	5.52E-04	1.13E-03	1.92E-04	9.36E-04
PWY-7187: pyrimidine deoxyribonucleotides de novo biosynthesis II	1.16E-04	6.02E-04	8.19E-04	5.97E-03	5.15E-03
PWY0-1586: peptidoglycan maturation (meso-diaminopimelate containing)|g_Escherichia.s_Escherichia_coli	1.20E-04	6.23E-04	5.39E-04	2.67E-06	5.37E-04
UNINTEGRATED|g_Escherichia.s_Escherichia_coli	1.20E-04	6.23E-04	4.55E-04	1.77E-06	4.53E-04
GLYOXYLATE-BYPASS: glyoxylate cycle	1.20E-04	6.23E-04	1.79E-03	3.73E-05	1.75E-03
P461-PWY: hexitol fermentation to lactate, formate, ethanol and acetate	1.23E-04	6.38E-04	3.80E-03	7.51E-04	3.05E-03
P441-PWY: superpathway of N-acetylneuraminate degradation	1.27E-04	6.55E-04	3.53E-03	7.65E-04	2.77E-03
METSYN-PWY: L-homoserine and L-methionine biosynthesis	1.31E-04	6.72E-04	2.11E-03	4.63E-03	2.51E-03
PWY66-422: D-galactose degradation V (Leloir pathway)	1.31E-04	6.72E-04	3.75E-03	6.45E-03	2.70E-03
PWY-1269: CMP-3-deoxy-D-manno-octulosonate biosynthesis I|g_Bacteroides.s_Bacteroides_caccae	1.47E-04	7.51E-04	7.42E-04	1.66E-04	5.76E-04
PWY4LZ-257: superpathway of fermentation (Chlamydomonas reinhardtii)	1.48E-04	7.56E-04	2.96E-03	2.67E-04	2.69E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
BRANCHED-CHAIN-AA-SYN-PWY: superpathway of branched amino acid biosynthesis|g_Bifidobacterium.s_Bifidobacterium_longum	1.50E-04	7.63E-04	8.90E-04	9.35E-05	7.96E-04
PWY-5103: L-isoleucine biosynthesis III|g_Bifidobacterium.s_Bifidobacterium_longum	1.50E-04	7.63E-04	8.80E-04	9.19E-05	7.88E-04
SER-GLYSYN-PWY: superpathway of L-serine and glycine biosynthesis I|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	1.56E-04	7.94E-04	7.97E-04	6.66E-05	7.30E-04
PWY-5100: pyruvate fermentation to acetate and lactate II|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	1.56E-04	7.94E-04	6.40E-04	1.62E-04	4.78E-04
CALVIN-PWY: Calvin-Benson-Bassham cycle|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	1.59E-04	8.05E-04	1.35E-03	2.04E-04	1.14E-03
PWY-2942: L-lysine biosynthesis III	1.67E-04	8.11E-04	5.20E-03	1.06E-02	5.38E-03
PWY-5347: superpathway of L-methionine biosynthesis (transsulfuration)	1.67E-04	8.11E-04	2.41E-03	4.96E-03	2.55E-03
ASPASN-PWY: superpathway of L-aspartate and L-asparagine biosynthesis|g_Escherichia.s_Escherichia_coli	1.75E-04	8.11E-04	3.93E-04	0.00E+00	3.93E-04
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Bifidobacterium.s_Bifidobacterium_bifidum	1.75E-04	8.11E-04	5.11E-04	0.00E+00	5.11E-04
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Eubacterium.s_Eubacterium_ventriosum	1.75E-04	8.11E-04	9.58E-04	0.00E+00	9.58E-04
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Paraprevotella.s_Paraprevotella_clara	1.75E-04	8.11E-04	0.00E+00	1.39E-03	1.39E-03
GALACTUROCAT-PWY: D-galacturonate degradation I|g_Alistipes.s_Alistipes_shahii	1.75E-04	8.11E-04	0.00E+00	9.30E-04	9.30E-04
GLYCOGENSYNTH-PWY: glycogen biosynthesis I (from ADP-D-Glucose)|g_Erysipelotrichaceae_noname.s_Eubacterium_biforme	1.75E-04	8.11E-04	0.00E+00	7.22E-04	7.22E-04
GLYOXYLATE-BYPASS: glyoxylate cycle|g_Escherichia.s_Escherichia_coli	1.75E-04	8.11E-04	4.78E-04	0.00E+00	4.78E-04
HISDEG-PWY: L-histidine degradation I|g_Fusobacterium.s_Fusobacterium_varium	1.75E-04	8.11E-04	3.89E-04	0.00E+00	3.89E-04
PEPTIDOGLYCANSYN-PWY: peptidoglycan biosynthesis I (meso-diaminopimelate containing)|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.75E-04	8.11E-04	0.00E+00	1.46E-03	1.46E-03
PWY-2942: L-lysine biosynthesis III|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	1.75E-04	8.11E-04	0.00E+00	1.94E-03	1.94E-03
PWY-3841: folate transformations II|g_Bacteroides.s_Bacteroides_dorei	1.75E-04	8.11E-04	0.00E+00	1.34E-03	1.34E-03
PWY-4242: pantothenate and coenzyme A biosynthesis III|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.75E-04	8.11E-04	0.00E+00	1.33E-03	1.33E-03
PWY-5100: pyruvate fermentation to acetate and lactate II|g_Eubacterium.s_Eubacterium_ventriosum	1.75E-04	8.11E-04	8.92E-04	0.00E+00	8.92E-04
PWY-5173: superpathway of acetyl-CoA biosynthesis|g_Escherichia.s_Escherichia_coli	1.75E-04	8.11E-04	6.19E-04	0.00E+00	6.19E-04
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Coprococcus.s_Coprococcus_comes	1.75E-04	8.11E-04	0.00E+00	1.13E-03	1.13E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Eubacterium.s_Eubacterium_ventriosum	1.75E-04	8.11E-04	1.24E-03	0.00E+00	1.24E-03
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Roseburia.s_Roseburia_inulinivorans	1.75E-04	8.11E-04	5.51E-04	0.00E+00	5.51E-04
PWY-5686: UMP biosynthesis|g_Parabacteroides.s_Parabacteroides_distasonis	1.75E-04	8.11E-04	0.00E+00	1.43E-03	1.43E-03
PWY-5695: urate biosynthesis/inosine 5'-phosphate degradation|g_Escherichia.s_Escherichia_coli	1.75E-04	8.11E-04	4.08E-04	0.00E+00	4.08E-04
PWY-5695: urate biosynthesis/inosine 5'-phosphate degradation|g_Roseburia.s_Roseburia_inulinivorans	1.75E-04	8.11E-04	8.08E-04	0.00E+00	8.08E-04
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I|g_Eubacterium.s_Eubacterium_ventriosum	1.75E-04	8.11E-04	9.75E-04	0.00E+00	9.75E-04
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Erysipelotrichaceae_noname.s_Eubacterium_biforme	1.75E-04	8.11E-04	0.00E+00	7.65E-04	7.65E-04
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Fusobacterium.s_Fusobacterium_varium	1.75E-04	8.11E-04	3.89E-04	0.00E+00	3.89E-04
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Ruminococcus.s_Ruminococcus_callidus	1.75E-04	8.11E-04	0.00E+00	1.65E-03	1.65E-03
PWY-6151: S-adenosyl-L-methionine cycle I|g_Paraprevotella.s_Paraprevotella_clara	1.75E-04	8.11E-04	0.00E+00	1.34E-03	1.34E-03
PWY-6151: S-adenosyl-L-methionine cycle I|g_Ruminococcus.s_Ruminococcus_callidus	1.75E-04	8.11E-04	0.00E+00	1.34E-03	1.34E-03
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Erysipelotrichaceae_noname.s_Eubacterium_biforme	1.75E-04	8.11E-04	0.00E+00	7.65E-04	7.65E-04
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Fusobacterium.s_Fusobacterium_varium	1.75E-04	8.11E-04	3.89E-04	0.00E+00	3.89E-04
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Ruminococcus.s_Ruminococcus_callidus	1.75E-04	8.11E-04	0.00E+00	1.65E-03	1.65E-03
PWY-6385: peptidoglycan biosynthesis III (mycobacteria)|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.75E-04	8.11E-04	0.00E+00	1.96E-03	1.96E-03
PWY-6386: UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.75E-04	8.11E-04	0.00E+00	1.48E-03	1.48E-03
PWY-6700: queuosine biosynthesis|g_Bacteroides.s_Bacteroides_fragilis	1.75E-04	8.11E-04	8.56E-04	0.00E+00	8.56E-04
PWY-6700: queuosine biosynthesis|g_Paraprevotella.s_Paraprevotella_clara	1.75E-04	8.11E-04	0.00E+00	1.05E-03	1.05E-03
PWY-6703: preQ0 biosynthesis|g_Bacteroides.s_Bacteroides_dorei	1.75E-04	8.11E-04	0.00E+00	2.71E-03	2.71E-03
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Blautia.s_Ruminococcus_gnavus	1.75E-04	8.11E-04	8.82E-04	0.00E+00	8.82E-04
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Escherichia.s_Escherichia_coli	1.75E-04	8.11E-04	4.98E-04	0.00E+00	4.98E-04
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Ruminococcus.s_Ruminococcus_lactaris	1.75E-04	8.11E-04	0.00E+00	6.80E-04	6.80E-04
PWY-7197: pyrimidine deoxyribonucleotide phosphorylation|g_Phascolarctobacterium.s_Phascolarctobacterium_succinatutens	1.75E-04	8.11E-04	0.00E+00	1.89E-03	1.89E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Veillonella.s_Veillonella_parvula	1.75E-04	8.11E-04	6.39E-04	0.00E+00	6.39E-04
PWY0-1061: superpathway of L-alanine biosynthesis|g_Escherichia.s_Escherichia_coli	1.75E-04	8.11E-04	5.28E-04	0.00E+00	5.28E-04
PWY0-1296: purine ribonucleosides degradation|g_Eubacterium.s_Eubacterium_ventriosum	1.75E-04	8.11E-04	9.98E-04	0.00E+00	9.98E-04
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Coprococcus.s_Coprococcus_comes	1.75E-04	8.11E-04	0.00E+00	1.13E-03	1.13E-03
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Eubacterium.s_Eubacterium_ventriosum	1.75E-04	8.11E-04	1.24E-03	0.00E+00	1.24E-03
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Roseburia.s_Roseburia_inulinivorans	1.75E-04	8.11E-04	5.51E-04	0.00E+00	5.51E-04
PWY66-422: D-galactose degradation V (Leloir pathway)|g_Erysipelotrichaceae_noname.s_Eubacterium_biforme	1.75E-04	8.11E-04	0.00E+00	6.25E-04	6.25E-04
PYRIDNUCSYN-PWY: NAD biosynthesis I (from aspartate)|g_Alistipes.s_Alistipes_shahii	1.75E-04	8.11E-04	0.00E+00	1.22E-03	1.22E-03
THRESYN-PWY: superpathway of L-threonine biosynthesis|g_Erysipelotrichaceae_noname.s_Eubacterium_biforme	1.75E-04	8.11E-04	0.00E+00	7.06E-04	7.06E-04
UDPNAGSYN-PWY: UDP-N-acetyl-D-glucosamine biosynthesis I|g_Escherichia.s_Escherichia_coli	1.75E-04	8.11E-04	5.54E-04	0.00E+00	5.54E-04
UNINTEGRATED|g_Alistipes.s_Alistipes_indistinctus	1.75E-04	8.11E-04	0.00E+00	1.15E-03	1.15E-03
UNINTEGRATED|g_Clostridium.s_Clostridium_bolteae	1.75E-04	8.11E-04	8.22E-04	0.00E+00	8.22E-04
VALSYN-PWY: L-valine biosynthesis|g_Coprococcus.s_Coprococcus_comes	1.75E-04	8.11E-04	0.00E+00	7.99E-04	7.99E-04
PWY-6630: superpathway of L-tyrosine biosynthesis	1.75E-04	8.11E-04	1.72E-03	0.00E+00	1.72E-03
PWY-2942: L-lysine biosynthesis III|g_Bacteroides.s_Bacteroides_uniformis	1.94E-04	8.94E-04	1.49E-03	3.48E-04	1.15E-03
HEXITOLDEGSUPER-PWY: superpathway of hexitol degradation (bacteria)	1.94E-04	8.94E-04	4.54E-03	1.53E-03	3.01E-03
COLANSYN-PWY: colanic acid building blocks biosynthesis	1.96E-04	9.03E-04	6.77E-03	2.33E-03	4.44E-03
PWY0-1298: superpathway of pyrimidine deoxyribonucleosides degradation	2.28E-04	1.05E-03	4.02E-03	1.18E-03	2.84E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_uniformis	2.37E-04	1.09E-03	1.47E-03	4.02E-04	1.07E-03
HOMOSER-METSYN-PWY: L-methionine biosynthesis I	2.69E-04	1.23E-03	1.92E-03	4.29E-03	2.37E-03
PANTO-PWY: phosphopantothenate biosynthesis I|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	2.76E-04	1.26E-03	9.35E-04	1.83E-04	7.52E-04
PENTOSE-P-PWY: pentose phosphate pathway	2.81E-04	1.28E-03	5.18E-03	2.09E-03	3.09E-03
PWY-5690: TCA cycle II (plants and fungi)	2.90E-04	1.32E-03	9.51E-04	6.18E-03	5.23E-03
PWY0-1297: superpathway of purine deoxyribonucleosides degradation	3.21E-04	1.46E-03	3.61E-03	8.18E-04	2.80E-03
UNINTEGRATED|g_Eubacterium.s_Eubacterium_eligens	3.22E-04	1.46E-03	3.34E-04	3.74E-03	3.40E-03
PWY-5695: urate biosynthesis/inosine 5'-phosphate degradation|g_Bacteroides.s_Bacteroides_dorei	3.24E-04	1.47E-03	1.93E-05	1.08E-03	1.06E-03
PWY-6151: S-adenosyl-L-methionine cycle I|g_Bifidobacterium.s_Bifidobacterium_longum	3.30E-04	1.49E-03	6.50E-04	5.22E-05	5.98E-04

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-5097: L-lysine biosynthesis VI|g_Bacteroides.s_Bacteroides_uniformis	3.36E-04	1.52E-03	1.48E-03	3.69E-04	1.11E-03
PWY-5695: urate biosynthesis/inosine 5'-phosphate degradation|g_Faecalibacterium.s_Faecalibacterium_prausnitzii	3.38E-04	1.52E-03	5.43E-04	2.12E-03	1.58E-03
PWY-6151: S-adenosyl-L-methionine cycle I	3.38E-04	1.52E-03	5.16E-03	9.73E-03	4.57E-03
PYRIDNUCSYN-PWY: NAD biosynthesis I (from aspartate)	3.38E-04	1.52E-03	3.88E-03	7.60E-03	3.71E-03
PANTO-PWY: phosphopantothenate biosynthesis I|g_Bacteroides.s_Bacteroides_uniformis	3.38E-04	1.52E-03	1.50E-03	3.71E-04	1.13E-03
VALSYN-PWY: L-valine biosynthesis|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	3.39E-04	1.52E-03	1.09E-03	1.70E-04	9.16E-04
TCA: TCA cycle I (prokaryotic)	3.63E-04	1.62E-03	1.46E-03	6.46E-03	5.00E-03
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_dorei	3.92E-04	1.75E-03	8.14E-06	1.14E-03	1.13E-03
FERMENTATION-PWY: mixed acid fermentation	3.92E-04	1.75E-03	1.72E-03	2.79E-05	1.69E-03
PWY-6628: superpathway of L-phenylalanine biosynthesis	3.92E-04	1.75E-03	1.91E-03	2.53E-05	1.88E-03
PWY-5695: urate biosynthesis/inosine 5'-phosphate degradation|g_Bacteroides.s_Bacteroides_xylanisolvens	4.01E-04	1.78E-03	9.53E-04	8.19E-05	8.71E-04
PWY-6609: adenine and adenosine salvage III|g_Bacteroides.s_Bacteroides_xylanisolvens	4.01E-04	1.78E-03	1.06E-03	9.83E-05	9.59E-04
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	4.08E-04	1.81E-03	1.21E-03	2.31E-04	9.77E-04
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	4.08E-04	1.81E-03	1.21E-03	2.31E-04	9.77E-04
GLUCUROCAT-PWY: superpathway of β-D-glucuronide and D-glucuronate degradation	4.21E-04	1.86E-03	4.28E-03	7.88E-03	3.60E-03
PWY-6527: stachyose degradation	4.21E-04	1.86E-03	3.93E-03	6.27E-03	2.34E-03
PWY66-400: glycolysis VI (metazoan)	4.21E-04	1.86E-03	3.77E-03	1.48E-03	2.28E-03
PWY0-1296: purine ribonucleosides degradation|g_Roseburia.s_Roseburia_inulinivorans	4.74E-04	1.99E-03	9.49E-04	6.52E-05	8.84E-04
RIBOSYN2-PWY: flavin biosynthesis I (bacteria and plants)	4.76E-04	1.99E-03	4.18E-03	1.66E-03	2.52E-03
1CMET2-PWY: N10-formyl-tetrahydrofolate biosynthesis|g_Parabacteroides.s_Parabacteroides_distasonis	4.83E-04	1.99E-03	0.00E+00	1.03E-03	1.03E-03
1CMET2-PWY: N10-formyl-tetrahydrofolate biosynthesis|g_Parabacteroides.s_Parabacteroides_merdae	4.83E-04	1.99E-03	0.00E+00	7.96E-04	7.96E-04
ARO-PWY: chorismate biosynthesis I|g_Roseburia.s_Roseburia_inulinivorans	4.83E-04	1.99E-03	5.72E-04	0.00E+00	5.72E-04
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Parabacteroides.s_Parabacteroides_merdae	4.83E-04	1.99E-03	0.00E+00	1.27E-03	1.27E-03
DTDPRHAMSYN-PWY: dTDP-L-rhamnose biosynthesis I|g_Erysipelotrichaceae_noname.s_Eubacterium_biforme	4.83E-04	1.99E-03	0.00E+00	5.16E-04	5.16E-04
PANTO-PWY: phosphopantothenate biosynthesis I|g_Escherichia.s_Escherichia_coli	4.83E-04	1.99E-03	5.43E-04	0.00E+00	5.43E-04
PEPTIDOGLYCANSYN-PWY: peptidoglycan biosynthesis I (meso-diaminopimelate containing)|g_Parabacteroides.s_Parabacteroides_distasonis	4.83E-04	1.99E-03	0.00E+00	1.13E-03	1.13E-03

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PEPTIDOGLYCANSYN-PWY: peptidoglycan biosynthesis I (meso-diaminopimelate containing)|g_Paraprevotella.s_Paraprevotella_clara	4.83E-04	1.99E-03	0.00E+00	1.43E-03	1.43E-03
PWY-1269: CMP-3-deoxy-D-manno-octulosonate biosynthesis I|g_Alistipes.s_Alistipes_shahii	4.83E-04	1.99E-03	0.00E+00	1.23E-03	1.23E-03
PWY-5100: pyruvate fermentation to acetate and lactate II|g_Erysipelotrichaceae_noname.s_Eubacterium_biforme	4.83E-04	1.99E-03	0.00E+00	6.47E-04	6.47E-04
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Parabacteroides.s_Parabacteroides_distasonis	4.83E-04	1.99E-03	0.00E+00	1.39E-03	1.39E-03
PWY-5695: urate biosynthesis/inosine 5'-phosphate degradation|g_Alistipes.s_Alistipes_shahii	4.83E-04	1.99E-03	0.00E+00	8.34E-04	8.34E-04
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I|g_Blautia.s_Ruminococcus_gnavus	4.83E-04	1.99E-03	9.17E-04	0.00E+00	9.17E-04
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Blautia.s_Ruminococcus_gnavus	4.83E-04	1.99E-03	9.15E-04	0.00E+00	9.15E-04
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Catenibacterium.s_Catenibacterium_mitsuokai	4.83E-04	1.99E-03	0.00E+00	4.78E-04	4.78E-04
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Coprococcus.s_Coprococcus_comes	4.83E-04	1.99E-03	0.00E+00	8.22E-04	8.22E-04
PWY-6151: S-adenosyl-L-methionine cycle I|g_Collinsella.s_Collinsella_aerofaciens	4.83E-04	1.99E-03	0.00E+00	4.21E-04	4.21E-04
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Blautia.s_Ruminococcus_gnavus	4.83E-04	1.99E-03	9.15E-04	0.00E+00	9.15E-04
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Catenibacterium.s_Catenibacterium_mitsuokai	4.83E-04	1.99E-03	0.00E+00	4.78E-04	4.78E-04
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Coprococcus.s_Coprococcus_comes	4.83E-04	1.99E-03	0.00E+00	8.22E-04	8.22E-04
PWY-6385: peptidoglycan biosynthesis III (mycobacteria)|g_Parabacteroides.s_Parabacteroides_distasonis	4.83E-04	1.99E-03	0.00E+00	1.09E-03	1.09E-03
PWY-6386: UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)|g_Parabacteroides.s_Parabacteroides_distasonis	4.83E-04	1.99E-03	0.00E+00	1.11E-03	1.11E-03
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)|g_Bacteroides.s_Bacteroides_dorei	4.83E-04	1.99E-03	0.00E+00	8.20E-04	8.20E-04
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)|g_Parabacteroides.s_Parabacteroides_distasonis	4.83E-04	1.99E-03	0.00E+00	1.22E-03	1.22E-03
PWY-6608: guanosine nucleotides degradation III|g_Escherichia.s_Escherichia_coli	4.83E-04	1.99E-03	3.74E-04	0.00E+00	3.74E-04
PWY-6609: adenine and adenosine salvage III|g_Escherichia.s_Escherichia_coli	4.83E-04	1.99E-03	5.28E-04	0.00E+00	5.28E-04
PWY-6700: queuosine biosynthesis|g_Ruminococcus.s_Ruminococcus_callidus	4.83E-04	1.99E-03	0.00E+00	1.72E-03	1.72E-03
PWY-6703: preQ0 biosynthesis|g_Ruminococcus.s_Ruminococcus_callidus	4.83E-04	1.99E-03	0.00E+00	1.56E-03	1.56E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Barnesiella.s_Barnesiella_intestinihominis	4.83E-04	1.99E-03	0.00E+00	1.83E-03	1.83E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Escherichia.s_Escherichia_coli	4.83E-04	1.99E-03	6.11E-04	0.00E+00	6.11E-04

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-7220: adenosine deoxyribonucleotides de novo biosynthesis II|g_Bacteroides.s_Bacteroides_dorei	4.83E-04	1.99E-03	0.00E+00	1.03E-03	1.03E-03
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Alistipes.s_Alistipes_shahii	4.83E-04	1.99E-03	0.00E+00	1.02E-03	1.02E-03
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Eubacterium.s_Eubacterium_ventriosum	4.83E-04	1.99E-03	1.10E-03	0.00E+00	1.10E-03
PWY-7222: guanosine deoxyribonucleotides de novo biosynthesis II|g_Bacteroides.s_Bacteroides_dorei	4.83E-04	1.99E-03	0.00E+00	1.03E-03	1.03E-03
PWY-7228: superpathway of guanosine nucleotides de novo biosynthesis I|g_Bacteroides.s_Bacteroides_dorei	4.83E-04	1.99E-03	0.00E+00	1.10E-03	1.10E-03
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Parabacteroides.s_Parabacteroides_distasonis	4.83E-04	1.99E-03	0.00E+00	1.39E-03	1.39E-03
SALVADEHYPOX-PWY: adenosine nucleotides degradation II|g_Escherichia.s_Escherichia_coli	4.83E-04	1.99E-03	4.11E-04	0.00E+00	4.11E-04
TRPSYN-PWY: L-tryptophan biosynthesis|g_Bacteroides.s_Bacteroides_fragilis	4.83E-04	1.99E-03	9.21E-04	0.00E+00	9.21E-04
UNINTEGRATED|g_Coprococcus.s_Coprococcus_catus	4.83E-04	1.99E-03	0.00E+00	1.15E-03	1.15E-03
VALSYN-PWY: L-valine biosynthesis|g_Escherichia.s_Escherichia_coli	4.83E-04	1.99E-03	6.39E-04	0.00E+00	6.39E-04
KETOGLUCONMET-PWY: ketogluconate metabolism	4.83E-04	1.99E-03	1.50E-03	0.00E+00	1.50E-03
PWY-6629: superpathway of L-tryptophan biosynthesis	4.83E-04	1.99E-03	1.37E-03	0.00E+00	1.37E-03
PWY-5686: UMP biosynthesis|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	4.89E-04	2.01E-03	8.71E-04	1.30E-04	7.41E-04
PWY-7560: methylerythritol phosphate pathway II	5.23E-04	2.15E-03	6.65E-03	3.16E-03	3.49E-03
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_xylanisolvens	5.65E-04	2.32E-03	1.81E-03	5.86E-04	1.23E-03
PWY-6123: inosine-5'-phosphate biosynthesis I|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	5.70E-04	2.33E-03	7.58E-04	5.63E-05	7.02E-04
VALSYN-PWY: L-valine biosynthesis|g_Roseburia.s_Roseburia_inulinivorans	5.70E-04	2.33E-03	5.29E-04	2.89E-05	5.00E-04
PANTO-PWY: phosphopantothenate biosynthesis I|g_Eubacterium.s_Eubacterium_ramulus	5.80E-04	2.37E-03	5.28E-04	2.54E-05	5.03E-04
NONMEVIPP-PWY: methylerythritol phosphate pathway I	6.46E-04	2.63E-03	5.68E-03	9.63E-03	3.94E-03
PWY-6270: isoprene biosynthesis I	6.46E-04	2.63E-03	6.64E-03	3.26E-03	3.38E-03
PWY66-409: superpathway of purine nucleotide salvage	6.48E-04	2.64E-03	2.82E-03	7.00E-04	2.12E-03
PWY-5173: superpathway of acetyl-CoA biosynthesis	7.02E-04	2.85E-03	1.86E-03	1.00E-04	1.76E-03
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	7.07E-04	2.87E-03	7.40E-04	7.65E-05	6.64E-04
PWY-6147: 6-hydroxymethyl-dihydropterin diphosphate biosynthesis I|g_Bacteroides.s_Bacteroides_xylanisolvens	7.24E-04	2.94E-03	9.85E-04	2.40E-04	7.45E-04
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Bifidobacterium.s_Bifidobacterium_longum	7.41E-04	3.00E-03	7.02E-04	1.25E-04	5.77E-04
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_thetaiotaomicron	8.34E-04	3.37E-03	2.82E-04	2.09E-03	1.81E-03
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	8.40E-04	3.39E-03	8.12E-04	1.33E-04	6.79E-04

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthe-sis|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	8.40E-04	3.39E-03	8.12E-04	1.33E-04	6.79E-04
PWY-6123: inosine-5'-phosphate biosynthesis I|g_Bifidobacterium.s_Bifidobacterium_longum	9.17E-04	3.69E-03	8.23E-04	8.15E-05	7.42E-04
UNINTEGRATED|g_Bacteroides.s_Bacteroides_dorei	9.71E-04	3.90E-03	6.77E-04	1.89E-03	1.21E-03
NONMEVIPP-PWY: methylerythritol phosphate pathway I|g_Roseburia.s_Roseburia_intestinalis	9.82E-04	3.94E-03	1.01E-03	1.39E-04	8.74E-04
PWY-6163: chorismate biosynthesis from 3-dehydroqui-nate|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	1.00E-03	4.01E-03	6.77E-04	1.08E-04	5.70E-04
PWY-1042: glycolysis IV (plant cyto-sol)|g_Anaerostipes.s_Anaerostipes_hadrus	1.02E-03	4.05E-03	9.56E-04	9.02E-05	8.66E-04
PWY-5791: 1,4-dihydroxy-2-naphthoate biosynthesis II (plants)	1.02E-03	4.05E-03	2.32E-03	2.51E-04	2.07E-03
PWY-5837: 1,4-dihydroxy-2-naphthoate biosynthesis I	1.02E-03	4.05E-03	2.32E-03	2.51E-04	2.07E-03
PWY-5897: superpathway of menaquinol-11 biosynthesis	1.02E-03	4.05E-03	2.33E-03	2.59E-04	2.07E-03
PWY-5898: superpathway of menaquinol-12 biosynthesis	1.02E-03	4.05E-03	2.33E-03	2.59E-04	2.07E-03
PWY-5899: superpathway of menaquinol-13 biosynthesis	1.02E-03	4.05E-03	2.33E-03	2.59E-04	2.07E-03
PWY0-1415: superpathway of heme biosynthesis from uroporphyrinogen-III	1.02E-03	4.05E-03	1.74E-03	9.11E-05	1.65E-03
PWY-7210: pyrimidine deoxyribonucleotides biosynthesis from CTP	1.09E-03	4.31E-03	3.60E-03	1.24E-03	2.36E-03
PWY-5659: GDP-mannose biosynthe-sis|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	1.17E-03	4.49E-03	6.88E-04	9.12E-05	5.97E-04
UNINTEGRATED|g_Bacteroides.s_Bacteroides_xylanisolvens	1.18E-03	4.49E-03	1.52E-03	6.55E-04	8.66E-04
PWY-7219: adenosine ribonucleotides de novo biosynthe-sis|g_Eubacterium.s_Eubacterium_ramulus	1.18E-03	4.49E-03	4.64E-04	9.16E-06	4.55E-04
PWY-6892: thiazole biosynthesis I (E. coli)	1.18E-03	4.49E-03	1.65E-03	3.54E-05	1.61E-03
THISYN-PWY: superpathway of thiamin diphosphate biosynthesis I	1.18E-03	4.49E-03	1.79E-03	4.48E-05	1.75E-03
POLYAMINSYN3-PWY: superpathway of polyamine biosynthesis II	1.19E-03	4.49E-03	1.03E-03	7.56E-03	6.52E-03
COA-PWY-1: coenzyme A biosynthesis II (mammali-an)|g_Bacteroides.s_Bacteroides_dorei	1.25E-03	4.49E-03	0.00E+00	8.69E-04	8.69E-04
COA-PWY-1: coenzyme A biosynthesis II (mammali-an)|g_Bacteroides.s_Bacteroides_fragilis	1.25E-03	4.49E-03	4.96E-04	0.00E+00	4.96E-04
COA-PWY-1: coenzyme A biosynthesis II (mammali-an)|g_Parabacteroides.s_Parabacteroides_distasonis	1.25E-03	4.49E-03	0.00E+00	8.52E-04	8.52E-04
COA-PWY-1: coenzyme A biosynthesis II (mammali-an)|g_Veillonella.s_Veillonella_parvula	1.25E-03	4.49E-03	6.33E-04	0.00E+00	6.33E-04
COA-PWY: coenzyme A biosynthesis I|g_Veillonella.s_Veillonella_parvula	1.25E-03	4.49E-03	4.92E-04	0.00E+00	4.92E-04
GLCMANNANAUT-PWY: superpathway of N-acetylglucosamine, N-acetylmannosamine and N-acetylneuraminate degrada-tion|g_Escherichia.s_Escherichia_coli	1.25E-03	4.49E-03	4.66E-04	0.00E+00	4.66E-04

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
HISTSYN-PWY: L-histidine biosynthesis|g_Bifidobacterium.s_Bifidobacterium_bifidum	1.25E-03	4.49E-03	3.92E-04	0.00E+00	3.92E-04
NONMEVIPP-PWY: methylerythritol phosphate pathway I|g_Paraprevotella.s_Paraprevotella_clara	1.25E-03	4.49E-03	0.00E+00	1.51E-03	1.51E-03
NONOXIPENT-PWY: pentose phosphate pathway (non-oxidative branch)|g_Blautia.s_Ruminococcus_gnavus	1.25E-03	4.49E-03	6.92E-04	0.00E+00	6.92E-04
PANTO-PWY: phosphopantothenate biosynthesis I|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	1.25E-03	4.49E-03	0.00E+00	8.16E-04	8.16E-04
PANTO-PWY: phosphopantothenate biosynthesis I|g_Blautia.s_Ruminococcus_gnavus	1.25E-03	4.49E-03	6.83E-04	0.00E+00	6.83E-04
PANTO-PWY: phosphopantothenate biosynthesis I|g_Ruminococcus.s_Ruminococcus_callidus	1.25E-03	4.49E-03	0.00E+00	1.26E-03	1.26E-03
PEPTIDOGLYCANSYN-PWY: peptidoglycan biosynthesis I (meso-diaminopimelate containing)|g_Coprococcus.s_Coprococcus_comes	1.25E-03	4.49E-03	0.00E+00	8.07E-04	8.07E-04
PWY-1042: glycolysis IV (plant cytosol)|g_Escherichia.s_Escherichia_coli	1.25E-03	4.49E-03	4.33E-04	0.00E+00	4.33E-04
PWY-2942: L-lysine biosynthesis III|g_Catenibacterium.s_Catenibacterium_mitsuokai	1.25E-03	4.49E-03	0.00E+00	4.71E-04	4.71E-04
PWY-4702: phytate degradation I|g_Escherichia.s_Escherichia_coli	1.25E-03	4.49E-03	5.17E-04	0.00E+00	5.17E-04
PWY-5188: tetrapyrrole biosynthesis I (from glutamate)|g_Fusobacterium.s_Fusobacterium_varium	1.25E-03	4.49E-03	4.20E-04	0.00E+00	4.20E-04
PWY-5384: sucrose degradation IV (sucrose phosphorylase)|g_Escherichia.s_Escherichia_coli	1.25E-03	4.49E-03	5.81E-04	0.00E+00	5.81E-04
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Barnesiella.s_Barnesiella_intestinihominis	1.25E-03	4.49E-03	0.00E+00	9.02E-04	9.02E-04
PWY-5686: UMP biosynthesis|g_Eubacterium.s_Eubacterium_ventriosum	1.25E-03	4.49E-03	7.34E-04	0.00E+00	7.34E-04
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I|g_Erysipelotrichaceae_noname.s_Eubacterium_biforme	1.25E-03	4.49E-03	0.00E+00	6.96E-04	6.96E-04
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I|g_Ruminococcus.s_Ruminococcus_callidus	1.25E-03	4.49E-03	0.00E+00	8.30E-04	8.30E-04
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Flavonifractor.s_Flavonifractor_plautii	1.25E-03	4.49E-03	4.89E-04	0.00E+00	4.89E-04
PWY-6125: superpathway of guanosine nucleotides de novo biosynthesis II|g_Bacteroides.s_Bacteroides_dorei	1.25E-03	4.49E-03	0.00E+00	1.10E-03	1.10E-03
PWY-6126: superpathway of adenosine nucleotides de novo biosynthesis II|g_Bacteroides.s_Bacteroides_dorei	1.25E-03	4.49E-03	0.00E+00	7.75E-04	7.75E-04
PWY-6151: S-adenosyl-L-methionine cycle I|g_Roseburia.s_Roseburia_inulinivorans	1.25E-03	4.49E-03	4.66E-04	0.00E+00	4.66E-04
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Flavonifractor.s_Flavonifractor_plautii	1.25E-03	4.49E-03	4.89E-04	0.00E+00	4.89E-04
PWY-6305: putrescine biosynthesis IV|g_Escherichia.s_Escherichia_coli	1.25E-03	4.49E-03	4.05E-04	0.00E+00	4.05E-04
PWY-6385: peptidoglycan biosynthesis III (mycobacteria)|g_Bacteroides.s_Bacteroides_fragilis	1.25E-03	4.49E-03	6.27E-04	0.00E+00	6.27E-04
PWY-6386: UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)|g_Bacteroides.s_Bacteroides_fragilis	1.25E-03	4.49E-03	7.84E-04	0.00E+00	7.84E-04
PWY-6386: UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)|g_Blautia.s_Ruminococcus_gnavus	1.25E-03	4.49E-03	8.20E-04	0.00E+00	8.20E-04

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-6386: UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)|g_Coprococcus.s_Coprococcus_comes	1.25E-03	4.49E-03	0.00E+00	8.30E-04	8.30E-04
PWY-6386: UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)|g_Eubacterium.s_Eubacterium_ventriosum	1.25E-03	4.49E-03	9.12E-04	0.00E+00	9.12E-04
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate contain-ing)|g_Blautia.s_Ruminococcus_gnavus	1.25E-03	4.49E-03	8.51E-04	0.00E+00	8.51E-04
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate contain-ing)|g_Coprococcus.s_Coprococcus_comes	1.25E-03	4.49E-03	0.00E+00	8.11E-04	8.11E-04
PWY-6609: adenine and adenosine salvage III|g_Alistipes.s_Alistipes_shahii	1.25E-03	4.49E-03	0.00E+00	8.51E-04	8.51E-04
PWY-6609: adenine and adenosine salvage III|g_Erysipelotrichaceae_noname.s_Eubacterium_biforme	1.25E-03	4.49E-03	0.00E+00	5.10E-04	5.10E-04
PWY-6609: adenine and adenosine salvage III|g_Eubacterium.s_Eubacterium_ventriosum	1.25E-03	4.49E-03	8.69E-04	0.00E+00	8.69E-04
PWY-6737: starch degradation V|g_Erysipelotrichaceae_noname.s_Eubacterium_biforme	1.25E-03	4.49E-03	0.00E+00	5.65E-04	5.65E-04
PWY-6737: starch degradation V|g_Fusobacterium.s_Fusobacterium_varium	1.25E-03	4.49E-03	4.05E-04	0.00E+00	4.05E-04
PWY-6936: seleno-amino acid biosynthe-sis|g_Escherichia.s_Escherichia_coli	1.25E-03	4.49E-03	5.43E-04	0.00E+00	5.43E-04
PWY-7219: adenosine ribonucleotides de novo biosynthe-sis|g_Alistipes.s_Alistipes_shahii	1.25E-03	4.49E-03	0.00E+00	1.07E-03	1.07E-03
PWY-7219: adenosine ribonucleotides de novo biosynthe-sis|g_Flavonifractor.s_Flavonifractor_plautii	1.25E-03	4.49E-03	4.85E-04	0.00E+00	4.85E-04
PWY-7221: guanosine ribonucleotides de novo biosynthe-sis|g_Blautia.s_Ruminococcus_gnavus	1.25E-03	4.49E-03	6.31E-04	0.00E+00	6.31E-04
PWY-7221: guanosine ribonucleotides de novo biosynthe-sis|g_Coprococcus.s_Coprococcus_comes	1.25E-03	4.49E-03	0.00E+00	6.83E-04	6.83E-04
PWY-7221: guanosine ribonucleotides de novo biosynthe-sis|g_Fusobacterium.s_Fusobacterium_varium	1.25E-03	4.49E-03	3.81E-04	0.00E+00	3.81E-04
PWY-7228: superpathway of guanosine nucleotides de novo biosynthesis I|g_Parabacteroides.s_Parabacteroides_distasonis	1.25E-03	4.49E-03	0.00E+00	1.81E-03	1.81E-03
PWY-7229: superpathway of adenosine nucleotides de novo biosynthesis I|g_Bacteroides.s_Bacteroides_dorei	1.25E-03	4.49E-03	0.00E+00	7.47E-04	7.47E-04
PWY-7282: 4-amino-2-methyl-5-phosphomethylpyrimidine biosynthesis (yeast)|g_Bacteroides.s_Bacteroides_fragilis	1.25E-03	4.49E-03	6.21E-04	0.00E+00	6.21E-04
PWY-7663: gondoate biosynthesis (anaero-bic)|g_Escherichia.s_Escherichia_coli	1.25E-03	4.49E-03	4.92E-04	0.00E+00	4.92E-04
PWY0-1241: ADP-L-glycero-β-D-manno-heptose biosynthesis|g_Escherichia.s_Escherichia_coli	1.25E-03	4.49E-03	4.09E-04	0.00E+00	4.09E-04
PWY0-1261: anhydromuropeptides recy-cling|g_Escherichia.s_Escherichia_coli	1.25E-03	4.49E-03	5.28E-04	0.00E+00	5.28E-04
PWY0-1296: purine ribonucleosides degrada-tion|g_Erysipelotrichaceae_noname.s_Eubacterium_biforme	1.25E-03	4.49E-03	0.00E+00	6.08E-04	6.08E-04
PWY0-1296: purine ribonucleosides degrada-tion|g_Fusobacterium.s_Fusobacterium_varium	1.25E-03	4.49E-03	3.16E-04	0.00E+00	3.16E-04
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Barnesiella.s_Barnesiella_intestinihominis	1.25E-03	4.49E-03	0.00E+00	9.02E-04	9.02E-04

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY0-845: superpathway of pyridoxal 5'-phosphate biosynthesis and salvage|g_Bacteroides.s_Bacteroides_fragilis	1.25E-03	4.49E-03	6.38E-04	0.00E+00	6.38E-04
PYRIDOXSYN-PWY: pyridoxal 5'-phosphate biosynthesis I|g_Bacteroides.s_Bacteroides_fragilis	1.25E-03	4.49E-03	8.00E-04	0.00E+00	8.00E-04
RHAMCAT-PWY: L-rhamnose degradation I|g_Escherichia.s_Escherichia_coli	1.25E-03	4.49E-03	5.40E-04	0.00E+00	5.40E-04
SER-GLYSYN-PWY: superpathway of L-serine and glycine biosynthesis I|g_Escherichia.s_Escherichia_coli	1.25E-03	4.49E-03	6.36E-04	0.00E+00	6.36E-04
SO4ASSIM-PWY: sulfate reduction I (assimilato-ry)|g_Escherichia.s_Escherichia_coli	1.25E-03	4.49E-03	4.64E-04	0.00E+00	4.64E-04
UNINTEGRATED|g_Alistipes.s_Alistipes_onderdonkii	1.25E-03	4.49E-03	0.00E+00	1.20E-03	1.20E-03
UNINTEGRATED|g_Flavonifractor.s_Flavonifractor_plautii	1.25E-03	4.49E-03	8.37E-04	0.00E+00	8.37E-04
VALSYN-PWY: L-valine biosynthe-sis|g_Ruminococcus.s_Ruminococcus_lactaris	1.25E-03	4.49E-03	0.00E+00	5.95E-04	5.95E-04
ENTBACSYN-PWY: enterobactin biosynthesis	1.25E-03	4.49E-03	1.75E-03	0.00E+00	1.75E-03
GLYCOL-GLYOXDEG-PWY: superpathway of glycol metabolism and degradation	1.25E-03	4.49E-03	1.66E-03	0.00E+00	1.66E-03
POLYISOPRENSYN-PWY: polyisoprenoid biosynthesis (E. coli)	1.25E-03	4.49E-03	1.85E-03	0.00E+00	1.85E-03
PWY-6891: thiazole biosynthesis II (Bacillus)	1.25E-03	4.49E-03	1.35E-03	0.00E+00	1.35E-03
PWY-6895: superpathway of thiamin diphosphate biosynthesis II	1.25E-03	4.49E-03	1.45E-03	0.00E+00	1.45E-03
PWY-7254: TCA cycle VII (acetate-producers)	1.25E-03	4.49E-03	1.48E-03	0.00E+00	1.48E-03
PWY-7198: pyrimidine deoxyribonucleotides de novo biosynthesis IV	1.28E-03	4.58E-03	3.36E-03	1.13E-03	2.22E-03
P461-PWY: hexitol fermentation to lactate, formate, ethanol and acetate|g_Anaerostipes.s_Anaerostipes_hadrus	1.33E-03	4.77E-03	6.54E-04	5.83E-05	5.95E-04
ARO-PWY: chorismate biosynthesis I|g_Roseburia.s_Roseburia_intestinalis	1.34E-03	4.79E-03	9.98E-04	1.31E-04	8.67E-04
COMPLETE-ARO-PWY: superpathway of aromatic amino acid biosynthesis|g_Roseburia.s_Roseburia_intestinalis	1.34E-03	4.79E-03	9.95E-04	1.34E-04	8.61E-04
PPGPPMET-PWY: ppGpp biosynthesis	1.39E-03	4.96E-03	3.00E-03	8.38E-04	2.17E-03
PWY-6703: preQ0 biosynthe-sis|g_Anaerostipes.s_Anaerostipes_hadrus	1.43E-03	5.07E-03	7.81E-04	3.22E-05	7.49E-04
PWY-7204: pyridoxal 5'-phosphate salvage II (plants)	1.43E-03	5.07E-03	1.90E-03	4.51E-05	1.86E-03
BRANCHED-CHAIN-AA-SYN-PWY: superpathway of branched amino acid biosynthesis	1.43E-03	5.08E-03	3.85E-03	6.35E-03	2.50E-03
PWY-7219: adenosine ribonucleotides de novo biosynthe-sis|g_Anaerostipes.s_Anaerostipes_hadrus	1.47E-03	5.21E-03	9.28E-04	1.02E-04	8.27E-04
PWY-6163: chorismate biosynthesis from 3-dehydroquinate|g_Roseburia.s_Roseburia_intestinalis	1.58E-03	5.61E-03	1.00E-03	1.41E-04	8.62E-04
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Anaerostipes.s_Anaerostipes_hadrus	1.59E-03	5.61E-03	7.21E-04	6.50E-05	6.56E-04
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Anaerostipes.s_Anaerostipes_hadrus	1.59E-03	5.61E-03	7.21E-04	6.50E-05	6.56E-04
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Dorea.s_Dorea_longicatena	1.64E-03	5.79E-03	4.68E-04	1.30E-03	8.29E-04
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Dorea.s_Dorea_longicatena	1.64E-03	5.79E-03	4.68E-04	1.30E-03	8.29E-04

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-621: sucrose degradation III (sucrose invertase)|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	1.66E-03	5.83E-03	1.15E-03	1.95E-04	9.57E-04
PWY-7383: anaerobic energy metabolism (invertebrates, cytosol)	1.69E-03	5.95E-03	1.13E-03	5.39E-03	4.26E-03
MET-SAM-PWY: superpathway of S-adenosyl-L-methionine biosynthesis|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	1.71E-03	6.00E-03	7.37E-04	5.51E-05	6.82E-04
METSYN-PWY: L-homoserine and L-methionine biosynthesis|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	1.71E-03	6.00E-03	7.29E-04	4.81E-05	6.81E-04
PWY-5347: superpathway of L-methionine biosynthesis (transsulfuration)|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	1.71E-03	6.00E-03	6.83E-04	4.73E-05	6.35E-04
PWY-6588: pyruvate fermentation to acetone|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	1.76E-03	6.14E-03	7.47E-04	1.10E-04	6.37E-04
PWY-5384: sucrose degradation IV (sucrose phosphorylase)	1.76E-03	6.14E-03	1.47E-03	2.72E-04	1.20E-03
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I|g_Anaerostipes.s_Anaerostipes_hadrus	1.89E-03	6.60E-03	7.14E-04	7.64E-05	6.37E-04
PWY-5686: UMP biosynthesis|g_Anaerostipes.s_Anaerostipes_hadrus	2.06E-03	7.15E-03	6.87E-04	5.01E-05	6.37E-04
PWY-7328: superpathway of UDP-glucose-derived O-antigen building blocks biosynthesis	2.06E-03	7.15E-03	2.69E-03	2.65E-04	2.43E-03
ASPASN-PWY: superpathway of L-aspartate and L-asparagine biosynthesis|g_Roseburia.s_Roseburia_intestinalis	2.06E-03	7.17E-03	9.45E-04	1.29E-04	8.16E-04
POLYAMSYN-PWY: superpathway of polyamine biosynthesis I	2.08E-03	7.21E-03	3.59E-03	7.56E-03	3.97E-03
CALVIN-PWY: Calvin-Benson-Bassham cycle|g_Anaerostipes.s_Anaerostipes_hadrus	2.10E-03	7.25E-03	9.47E-04	1.31E-04	8.16E-04
THRESYN-PWY: superpathway of L-threonine biosynthesis|g_Anaerostipes.s_Anaerostipes_hadrus	2.10E-03	7.25E-03	7.45E-04	1.33E-04	6.12E-04
PWY-5863: superpathway of phylloquinol biosynthesis	2.10E-03	7.25E-03	2.75E-03	4.98E-04	2.25E-03
COBALSYN-PWY: adenosylcobalamin salvage from cobinamide I|g_Roseburia.s_Roseburia_intestinalis	2.15E-03	7.41E-03	1.08E-03	1.62E-04	9.22E-04
RHAMCAT-PWY: L-rhamnose degradation I|g_Bacteroides.s_Bacteroides_xylanisolvens	2.22E-03	7.64E-03	1.32E-03	4.33E-04	8.87E-04
GLCMANNANAUT-PWY: superpathway of N-acetylglucosamine, N-acetylmannosamine and N-acetylneuraminate degradation|g_Dorea.s_Dorea_longicatena	2.24E-03	7.71E-03	5.13E-04	1.73E-03	1.22E-03
VALSYN-PWY: L-valine biosynthesis|g_Anaerostipes.s_Anaerostipes_hadrus	2.24E-03	7.71E-03	8.21E-04	1.00E-04	7.20E-04
P164-PWY: purine nucleobases degradation I (anaerobic)	2.24E-03	7.71E-03	3.52E-04	3.46E-03	3.11E-03
PWY-6703: preQ0 biosynthesis|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	2.33E-03	8.00E-03	9.86E-04	2.54E-04	7.32E-04
PWY-6163: chorismate biosynthesis from 3-dehydroquinate|g_Roseburia.s_Roseburia_inulinivorans	2.46E-03	8.41E-03	5.17E-04	2.79E-05	4.89E-04
PWY-7357: thiamin formation from pyrithiamine and oxythiamine (yeast)|g_Roseburia.s_Roseburia_inulinivorans	2.46E-03	8.41E-03	4.13E-04	2.23E-05	3.91E-04

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-7357: thiamin formation from pyrithiamine and oxythiamine (yeast)	2.49E-03	8.50E-03	6.12E-03	8.51E-03	2.38E-03
DTDPRHAMSYN-PWY: dTDP-L-rhamnose biosynthesis I|g_Anaerostipes.s_Anaerostipes_hadrus	2.66E-03	9.06E-03	7.68E-04	1.17E-04	6.51E-04
PWY66-422: D-galactose degradation V (Leloir pathway)|g_Anaerostipes.s_Anaerostipes_hadrus	2.66E-03	9.06E-03	7.83E-04	1.12E-04	6.71E-04
COMPLETE-ARO-PWY: superpathway of aromatic amino acid biosynthesis|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	2.74E-03	9.28E-03	7.86E-04	1.79E-04	6.07E-04
PWY-5676: acetyl-CoA fermentation to butanoate II	2.82E-03	9.28E-03	2.69E-03	5.45E-03	2.76E-03
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I|g_Roseburia.s_Roseburia_intestinalis	2.83E-03	9.28E-03	9.20E-04	1.62E-04	7.58E-04
MET-SAM-PWY: superpathway of S-adenosyl-L-methionine biosynthesis	2.96E-03	9.28E-03	2.22E-03	4.25E-03	2.02E-03
PWY-6897: thiamin salvage II	2.96E-03	9.28E-03	6.23E-03	8.85E-03	2.62E-03
PWY-7199: pyrimidine deoxyribonucleosides salvage	2.96E-03	9.28E-03	4.03E-03	7.03E-03	3.00E-03
PWY-6737: starch degradation V|g_Roseburia.s_Roseburia_inulinivorans	3.00E-03	9.28E-03	5.91E-04	1.13E-04	4.78E-04
1CMET2-PWY: N10-formyl-tetrahydrofolate biosynthesis|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	6.12E-04	0.00E+00	6.12E-04
ARGININE-SYN4-PWY: L-ornithine de novo biosynthesis|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	5.55E-04	0.00E+00	5.55E-04
ARGININE-SYN4-PWY: L-ornithine de novo biosynthesis|g_Bacteroides.s_Bacteroides_xylanisolvens	3.08E-03	9.28E-03	0.00E+00	9.31E-04	9.31E-04
AST-PWY: L-arginine degradation II (AST pathway)|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	4.73E-04	0.00E+00	4.73E-04
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	3.08E-03	9.28E-03	0.00E+00	1.25E-03	1.25E-03
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	4.55E-04	0.00E+00	4.55E-04
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Odoribacter.s_Odoribacter_splanchnicus	3.08E-03	9.28E-03	8.84E-04	0.00E+00	8.84E-04
COA-PWY-1: coenzyme A biosynthesis II (mammalian)|g_Ruminococcus.s_Ruminococcus_callidus	3.08E-03	9.28E-03	0.00E+00	7.23E-04	7.23E-04
COA-PWY: coenzyme A biosynthesis I|g_Odoribacter.s_Odoribacter_splanchnicus	3.08E-03	9.28E-03	8.09E-04	0.00E+00	8.09E-04
DTDPRHAMSYN-PWY: dTDP-L-rhamnose biosynthesis I|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	5.81E-04	0.00E+00	5.81E-04
DTDPRHAMSYN-PWY: dTDP-L-rhamnose biosynthesis I|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	4.77E-04	0.00E+00	4.77E-04
FUCCAT-PWY: fucose degradation|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	5.79E-04	0.00E+00	5.79E-04
GLUCOSE1PMETAB-PWY: glucose and glucose-1-phosphate degradation|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	3.72E-04	0.00E+00	3.72E-04
GLYCOL-GLYOXDEG-PWY: superpathway of glycol metabolism and degradation|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	5.17E-04	0.00E+00	5.17E-04
HEMESYN2-PWY: heme biosynthesis II (anaerobic)|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	5.25E-04	0.00E+00	5.25E-04
P461-PWY: hexitol fermentation to lactate, formate, ethanol and acetate|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	4.94E-04	0.00E+00	4.94E-04

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PANTO-PWY: phosphopantothenate biosynthesis I|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	5.18E-04	0.00E+00	5.18E-04
PANTOSYN-PWY: pantothenate and coenzyme A biosynthesis I|g_Eubacterium.s_Eubacterium_ventriosum	3.08E-03	9.28E-03	8.37E-04	0.00E+00	8.37E-04
PEPTIDOGLYCANSYN-PWY: peptidoglycan biosynthesis I (meso-diaminopimelate containing)|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	3.08E-03	9.28E-03	0.00E+00	1.02E-03	1.02E-03
PEPTIDOGLYCANSYN-PWY: peptidoglycan biosynthesis I (meso-diaminopimelate containing)|g_Bacteroides.s_Bacteroides_dorei	3.08E-03	9.28E-03	0.00E+00	5.91E-04	5.91E-04
PEPTIDOGLYCANSYN-PWY: peptidoglycan biosynthesis I (meso-diaminopimelate containing)|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	5.35E-04	0.00E+00	5.35E-04
POLYISOPRENSYN-PWY: polyisoprenoid biosynthesis (E. coli)|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	5.90E-04	0.00E+00	5.90E-04
PWY-1042: glycolysis IV (plant cytosol)|g_Bacteroides.s_Bacteroides_xylanisolvens	3.08E-03	9.28E-03	0.00E+00	1.56E-03	1.56E-03
PWY-2941: L-lysine biosynthesis II|g_Catenibacterium.s_Catenibacterium_mitsuokai	3.08E-03	9.28E-03	0.00E+00	4.69E-04	4.69E-04
PWY-2942: L-lysine biosynthesis III|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	6.52E-04	0.00E+00	6.52E-04
PWY-2942: L-lysine biosynthesis III|g_Veillonella.s_Veillonella_parvula	3.08E-03	9.28E-03	5.75E-04	0.00E+00	5.75E-04
PWY-3841: folate transformations II|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	6.16E-04	0.00E+00	6.16E-04
PWY-4242: pantothenate and coenzyme A biosynthesis III|g_Eubacterium.s_Eubacterium_ventriosum	3.08E-03	9.28E-03	8.73E-04	0.00E+00	8.73E-04
PWY-4984: urea cycle|g_Flavonifractor.s_Flavonifractor_plautii	3.08E-03	9.28E-03	6.71E-04	0.00E+00	6.71E-04
PWY-5083: NAD/NADH phosphorylation and dephosphorylation|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	5.97E-04	0.00E+00	5.97E-04
PWY-5097: L-lysine biosynthesis VI|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	6.07E-04	0.00E+00	6.07E-04
PWY-5097: L-lysine biosynthesis VI|g_Barnesiella.s_Barnesiella_intestinihominis	3.08E-03	9.28E-03	0.00E+00	1.17E-03	1.17E-03
PWY-5097: L-lysine biosynthesis VI|g_Catenibacterium.s_Catenibacterium_mitsuokai	3.08E-03	9.28E-03	0.00E+00	4.61E-04	4.61E-04
PWY-5189: tetrapyrrole biosynthesis II (from glycine)|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	3.82E-04	0.00E+00	3.82E-04
PWY-5659: GDP-mannose biosynthesis|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	4.36E-04	0.00E+00	4.36E-04
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	5.91E-04	0.00E+00	5.91E-04
PWY-5667: CDP-diacylglycerol biosynthesis I|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	3.70E-04	0.00E+00	3.70E-04
PWY-5686: UMP biosynthesis|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	4.93E-04	0.00E+00	4.93E-04
PWY-5686: UMP biosynthesis|g_Catenibacterium.s_Catenibacterium_mitsuokai	3.08E-03	9.28E-03	0.00E+00	4.03E-04	4.03E-04
PWY-5686: UMP biosynthesis|g_Erysipelotrichaceae_noname.s_Eubacterium_biforme	3.08E-03	9.28E-03	0.00E+00	5.06E-04	5.06E-04

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-5695: urate biosynthesis/inosine 5'-phosphate degradation|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	4.57E-04	0.00E+00	4.57E-04
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I|g_Coprococcus.s_Coprococcus_comes	3.08E-03	9.28E-03	0.00E+00	6.36E-04	6.36E-04
PWY-6121: 5-aminoimidazole ribonucleotide biosynthesis I|g_Fusobacterium.s_Fusobacterium_varium	3.08E-03	9.28E-03	3.74E-04	0.00E+00	3.74E-04
PWY-6122: 5-aminoimidazole ribonucleotide biosynthesis II|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	3.68E-04	0.00E+00	3.68E-04
PWY-6123: inosine-5'-phosphate biosynthesis I|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	5.27E-04	0.00E+00	5.27E-04
PWY-6124: inosine-5'-phosphate biosynthesis II|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	5.25E-04	0.00E+00	5.25E-04
PWY-6124: inosine-5'-phosphate biosynthesis II|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	5.23E-04	0.00E+00	5.23E-04
PWY-6151: S-adenosyl-L-methionine cycle I|g_Catenibacterium.s_Catenibacterium_mitsuokai	3.08E-03	9.28E-03	0.00E+00	4.52E-04	4.52E-04
PWY-6151: S-adenosyl-L-methionine cycle I|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	4.80E-04	0.00E+00	4.80E-04
PWY-6277: superpathway of 5-aminoimidazole ribonucleotide biosynthesis|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	3.68E-04	0.00E+00	3.68E-04
PWY-6385: peptidoglycan biosynthesis III (mycobacteria)|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	4.80E-04	0.00E+00	4.80E-04
PWY-6385: peptidoglycan biosynthesis III (mycobacteria)|g_Eubacterium.s_Eubacterium_ventriosum	3.08E-03	9.28E-03	7.97E-04	0.00E+00	7.97E-04
PWY-6386: UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	3.08E-03	9.28E-03	0.00E+00	1.05E-03	1.05E-03
PWY-6386: UDP-N-acetylmuramoyl-pentapeptide biosynthesis II (lysine-containing)|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	4.86E-04	0.00E+00	4.86E-04
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)|g_Adlercreutzia.s_Adlercreutzia_equolifaciens	3.08E-03	9.28E-03	0.00E+00	1.09E-03	1.09E-03
PWY-6387: UDP-N-acetylmuramoyl-pentapeptide biosynthesis I (meso-diaminopimelate containing)|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	5.23E-04	0.00E+00	5.23E-04
PWY-6608: guanosine nucleotides degradation III|g_Coprococcus.s_Coprococcus_catus	3.08E-03	9.28E-03	0.00E+00	7.98E-04	7.98E-04
PWY-6609: adenine and adenosine salvage III|g_Blautia.s_Ruminococcus_gnavus	3.08E-03	9.28E-03	8.06E-04	0.00E+00	8.06E-04
PWY-6703: preQ0 biosynthesis|g_Ruminococcus.s_Ruminococcus_lactaris	3.08E-03	9.28E-03	0.00E+00	9.42E-04	9.42E-04
PWY-6859: all-trans-farnesol biosynthesis|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	5.95E-04	0.00E+00	5.95E-04
PWY-6936: seleno-amino acid biosynthesis|g_Erysipelotrichaceae_noname.s_Eubacterium_biforme	3.08E-03	9.28E-03	0.00E+00	5.76E-04	5.76E-04
PWY-6936: seleno-amino acid biosynthesis|g_Eubacterium.s_Eubacterium_ventriosum	3.08E-03	9.28E-03	7.85E-04	0.00E+00	7.85E-04
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Barnesiella.s_Barnesiella_intestinihominis	3.08E-03	9.28E-03	0.00E+00	7.71E-04	7.71E-04
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Coprococcus.s_Coprococcus_catus	3.08E-03	9.28E-03	0.00E+00	7.74E-04	7.74E-04

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_7_1_58FAA	3.08E-03	9.28E-03	6.60E-04	0.00E+00	6.60E-04
PWY-7111: pyruvate fermentation to isobutanol (engineered)|g_Veillonella.s_Veillonella_parvula	3.08E-03	9.28E-03	4.09E-04	0.00E+00	4.09E-04
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Alistipes.s_Alistipes_indistinctus	3.08E-03	9.28E-03	0.00E+00	6.40E-04	6.40E-04
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	5.72E-04	0.00E+00	5.72E-04
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Clostridium.s_Clostridium_bolteae	3.08E-03	9.28E-03	5.77E-04	0.00E+00	5.77E-04
PWY-7219: adenosine ribonucleotides de novo biosynthesis|g_Ruminococcus.s_Ruminococcus_lactaris	3.08E-03	9.28E-03	0.00E+00	6.60E-04	6.60E-04
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	5.24E-04	0.00E+00	5.24E-04
PWY-7221: guanosine ribonucleotides de novo biosynthesis|g_Flavonifractor.s_Flavonifractor_plautii	3.08E-03	9.28E-03	5.62E-04	0.00E+00	5.62E-04
PWY-7242: D-fructuronate degradation|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	4.64E-04	0.00E+00	4.64E-04
PWY-7282: 4-amino-2-methyl-5-phosphomethylpyrimidine biosynthesis (yeast)|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	6.27E-04	0.00E+00	6.27E-04
PWY-7663: gondoate biosynthesis (anaerobic)|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	6.40E-04	0.00E+00	6.40E-04
PWY0-1296: purine ribonucleosides degradation|g_Catenibacterium.s_Catenibacterium_mitsuokai	3.08E-03	9.28E-03	0.00E+00	3.61E-04	3.61E-04
PWY0-1296: purine ribonucleosides degradation|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	4.20E-04	0.00E+00	4.20E-04
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	5.91E-04	0.00E+00	5.91E-04
PWY0-1319: CDP-diacylglycerol biosynthesis II|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	3.70E-04	0.00E+00	3.70E-04
PWY0-1586: peptidoglycan maturation (meso-diaminopimelate containing)|g_Veillonella.s_Veillonella_parvula	3.08E-03	9.28E-03	6.19E-04	0.00E+00	6.19E-04
PWY0-845: superpathway of pyridoxal 5'-phosphate biosynthesis and salvage|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	6.14E-04	0.00E+00	6.14E-04
PWY66-400: glycolysis VI (metazoan)|g_Bacteroides.s_Bacteroides_xylanisolvens	3.08E-03	9.28E-03	0.00E+00	1.30E-03	1.30E-03
PYRIDOXSYN-PWY: pyridoxal 5'-phosphate biosynthesis I|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	6.07E-04	0.00E+00	6.07E-04
THRESYN-PWY: superpathway of L-threonine biosynthesis|g_Escherichia.s_Escherichia_coli	3.08E-03	9.28E-03	4.77E-04	0.00E+00	4.77E-04
THRESYN-PWY: superpathway of L-threonine biosynthesis|g_Veillonella.s_Veillonella_parvula	3.08E-03	9.28E-03	5.37E-04	0.00E+00	5.37E-04
UNINTEGRATED|g_Bacteroides.s_Bacteroides_finegoldii	3.08E-03	9.28E-03	0.00E+00	1.79E-03	1.79E-03
UNINTEGRATED|g_Bacteroides.s_Bacteroides_sp_3_2_5	3.08E-03	9.28E-03	5.50E-04	0.00E+00	5.50E-04
UNINTEGRATED|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_7_1_58FAA	3.08E-03	9.28E-03	7.74E-04	0.00E+00	7.74E-04
UNINTEGRATED|g_Ruminococcus.s_Ruminococcus_bromii	3.08E-03	9.28E-03	0.00E+00	5.86E-04	5.86E-04
VALSYN-PWY: L-valine biosynthesis|g_Barnesiella.s_Barnesiella_intestinihominis	3.08E-03	9.28E-03	0.00E+00	7.71E-04	7.71E-04

Test.Variable	P.Value	pFDR (BH)	Sub1 var mean	Sub2 var mean	Absolute Difference
VALSYN-PWY: L-valine biosynthesis|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_7_1_58FAA	3.08E-03	9.28E-03	6.60E-04	0.00E+00	6.60E-04
VALSYN-PWY: L-valine biosynthesis|g_Veillonella.s_Veillonella_parvula	3.08E-03	9.28E-03	4.09E-04	0.00E+00	4.09E-04
AST-PWY: L-arginine degradation II (AST pathway)	3.08E-03	9.28E-03	1.20E-03	0.00E+00	1.20E-03
P105-PWY: TCA cycle IV (2-oxoglutarate decarboxylase)	3.08E-03	9.28E-03	1.42E-03	0.00E+00	1.42E-03
PWY-2723: trehalose degradation V	3.08E-03	9.28E-03	1.57E-03	0.00E+00	1.57E-03
PWY-561: superpathway of glyoxylate cycle and fatty acid degradation	3.08E-03	9.28E-03	1.27E-03	0.00E+00	1.27E-03
PWY-5723: Rubisco shunt	3.08E-03	9.28E-03	1.42E-03	0.00E+00	1.42E-03
PWY-6731: starch degradation III	3.08E-03	9.28E-03	1.95E-03	0.00E+00	1.95E-03
PWY-6823: molybdenum cofactor biosynthesis	3.08E-03	9.28E-03	1.34E-03	0.00E+00	1.34E-03
PWY-7269: NAD/NADP-NADH/NADPH mitochondrial interconversion (yeast)	3.08E-03	9.28E-03	1.20E-03	0.00E+00	1.20E-03
PWY-7315: dTDP-N-acetylthomosamine biosynthesis	3.08E-03	9.28E-03	1.51E-03	0.00E+00	1.51E-03
PWY-6317: galactose degradation I (Leloir pathway)|g_Anaerostipes.s_Anaerostipes_hadrus	3.14E-03	9.46E-03	8.15E-04	1.20E-04	6.95E-04
PWY-6969: TCA cycle V (2-oxoglutarate:ferredoxin oxidoreductase)	3.19E-03	9.61E-03	1.29E-03	5.63E-03	4.33E-03
ARO-PWY: chorismate biosynthesis I|g_Lachnospiraceae_noname.s_Lachnospiraceae_bacterium_5_1_63FAA	3.22E-03	9.67E-03	7.73E-04	1.81E-04	5.92E-04
PWY-5686: UMP biosynthesis|g_Bifidobacterium.s_Bifidobacterium_longum	3.22E-03	9.67E-03	6.26E-04	7.82E-05	5.48E-04
NAD-BIOSYNTHESIS-II: NAD salvage pathway II|g_Escherichia.s_Escherichia_coli	3.33E-03	9.96E-03	5.96E-04	9.58E-06	5.86E-04
PWY0-1533: methylphosphonate degradation I|g_Escherichia.s_Escherichia_coli	3.33E-03	9.96E-03	4.30E-04	5.90E-06	4.24E-04
TRPSYN-PWY: L-tryptophan biosynthesis|g_Escherichia.s_Escherichia_coli	3.33E-03	9.96E-03	5.20E-04	9.23E-06	5.11E-04

Supplementary Table 8.2 | Metabolic pathways with significantly varied relative abundance between participants (pFDR = 0.01).

Appendix D

PUBLICATIONS

A Prospective Metagenomic and Metabolomic Analysis of the Impact of Exercise and/or Whey Protein Supplementation on the Gut Microbiome of Sedentary Adults

AMERICAN SOCIETY FOR MICROBIOLOGY | mSystems®

RESEARCH ARTICLE
Host-Microbe Biology

A Prospective Metagenomic and Metabolomic Analysis of the Impact of Exercise and/or Whey Protein Supplementation on the Gut Microbiome of Sedentary Adults

Owen Cronin,[a,b] Wiley Barton,[a,b,c] Peter Skuse,[c] Nicholas C. Penney,[d,e] Isabel Garcia-Perez,[d] Eileen F. Murphy,[f] Trevor Woods,[g] Helena Nugent,[b] Aine Fanning,[a] Silvia Melgar,[a] Eanna C. Falvey,[b,h] Elaine Holmes,[d] Paul D. Cotter,[a,c] Orla O'Sullivan,[a,c] Michael G. Molloy,[a,b] Fergus Shanahan[a,b]

[a]APC Microbiome Ireland, National University of Ireland, Cork, Ireland
[b]Department of Medicine, National University of Ireland, Cork, Ireland
[c]Teagasc Food Research Centre, Moorepark, Fermoy, Cork, Ireland
[d]Section of Biomolecular Medicine, Division of Computational Systems Medicine, Department of Surgery and Cancer, Imperial College London, London, United Kingdom
[e]Division of Surgery, Department of Surgery and Cancer, Imperial College London, London, United Kingdom
[f]Alimentary Health Ltd., Cork, Ireland
[g]Department of Sport and Physical Activity, Human Performance Laboratory, National University of Ireland, Cork, Ireland
[h]Department of Sports Medicine, Sports Surgery Clinic, Santry, Dublin, Ireland

ABSTRACT Many components of modern living exert influence on the resident intestinal microbiota of humans with resultant impact on host health. For example, exercise-associated changes in the diversity, composition, and functional profiles of microbial populations in the gut have been described in cross-sectional studies of habitual athletes. However, this relationship is also affected by changes in diet, such as changes in dietary and supplementary protein consumption, that coincide with exercise. To determine whether increasing physical activity and/or increased protein intake modulates gut microbial composition and function, we prospectively challenged healthy but sedentary adults with a short-term exercise regime, with and without concurrent daily whey protein consumption. Metagenomics- and metabolomics-based assessments demonstrated modest changes in gut microbial composition and function following increases in physical activity. Significant changes in the diversity of the gut virome were evident in participants receiving daily whey protein supplementation. Results indicate that improved body composition with exercise is not dependent on major changes in the diversity of microbial populations in the gut. The diverse microbial characteristics previously observed in long-term habitual athletes may be a later response to exercise and fitness improvement.

IMPORTANCE The gut microbiota of humans is a critical component of functional development and subsequent health. It is important to understand the lifestyle and dietary factors that affect the gut microbiome and what impact these factors may have. Animal studies suggest that exercise can directly affect the gut microbiota, and elite athletes demonstrate unique beneficial and diverse gut microbiome characteristics. These characteristics are associated with levels of protein consumption and levels of physical activity. The results of this study show that increasing the fitness levels of physically inactive humans leads to modest but detectable changes in gut microbiota characteristics. For the first time, we show that regular whey protein intake leads to significant alterations to the composition of the gut virome.

KEYWORDS bacteriophages, exercise, metabolism, microbial communities, next-generation sequencing, whey protein

Received 30 March 2018 **Accepted** 2 April 2018 **Published** 24 April 2018

Citation Cronin O, Barton W, Skuse P, Penney NC, Garcia-Perez I, Murphy EF, Woods T, Nugent H, Fanning A, Melgar S, Falvey EC, Holmes E, Cotter PD, O'Sullivan O, Molloy MG, Shanahan F. 2018. A prospective metagenomic and metabolomic analysis of the impact of exercise and/or whey protein supplementation on the gut microbiome of sedentary adults. mSystems 3:e00044-18. https://doi.org/10.1128/mSystems.00044-18.

Editor Jack A. Gilbert, University of Chicago

Copyright © 2018 Cronin et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International license.

Address correspondence to Owen Cronin, owen.cronin@hotmail.com.

O.C. and W.B. contributed equally to this article.

Whey protein-derived bacteriophages alter the gut virome of physically inactive volunteers.

Downloaded from http://msystems.asm.org/ on April 27, 2018 by guest

Most of the elements of human lifestyle and environment influence the composition or function of the gut microbiome (1, 2). Indeed, the microbiome has been viewed as a transducer of nutrient and other environmental signals for the host (3). Therefore, several investigators have begun to explore whether a sedentary lifestyle or, more specifically, a lifestyle that includes exercise and fitness is associated with changes in the gut microbiota (4–7). This has been assessed in cross-sectional studies of habitual exercisers (8–10) and professional athletes (10–12) in addition to experimental models.

In elite athletes, distinct compositional and functional microbial characteristics, including increased α-diversity, enhanced microbial production of short-chain fatty acids, and greater metabolic capacity, are evident in the gut (11, 12). These microbial features positively correlate with the athletes' levels of physical activity, in addition to the quantity of dietary protein consumed. In many professional sporting disciplines, as well as amateur sport, intentional protein supplementation (e.g., whey protein) provides a sizeable proportion of athletes' daily protein intake (12).

Evidence from animal studies highlights the potential for taxonomic manipulation of colonic microbiota following exercise interventions, both with and without concurrent dietary alterations (6, 13, 14). Previously, we have proposed several mechanisms by which exercise and resultant fitness may directly influence the gut microbiota, including effects on gastrointestinal transit time (15), a known driver of the diversity of microbial populations in the gut (16, 17). It appears that physical activity initiated in the juvenile period of development demonstrates a greater potential for fostering a preferential microbiota than exercise commenced in adulthood (6, 18).

However, the relationship between exercise and alterations in the microbiome in humans is compounded by changes in dietary consumption that often accompany physical activity, e.g., increased protein supplement intake.

Building on previous work (11, 12), the present study sought to interrogate correlations between the gut microbiome and levels of physical activity and protein consumption. To do so, using a combination of next-generation shotgun sequencing and metabolomic analysis, we prospectively examined the impact of exercise, with and without whey protein supplementation, on the adult human gut microbiome. We report that 8 weeks of combined aerobic and resistance training led to modest alterations in the composition and activity of the gut microbiome of sedentary individuals. Participants consuming whey protein daily did experience a marked alteration in the diversity of their gut virome following 8 weeks of oral supplementation.

RESULTS

Study overview. Following local advertisement, healthy Irish male and female Caucasian volunteers ($n = 90$) aged 18 to 40 years and with a body mass index (BMI) of between 22 and 35 kg/m² (predominantly overweight or obese) were recruited between January and August 2014 (Fig. 1). The study was conducted in accordance with the Declaration of Helsinki, and, prior to commencement, ethical approval was granted by the Cork Clinical Research Ethics Committee (CREC). All volunteers provided written informed consent. To prospectively measure the effect of *de novo* exercise training on gut microbiota, subjects were required to be physically inactive for at least 3 months prior to study entry (i.e., not engaged in regular structured or unstructured exercise beyond the light physical activities of daily life). All participants were screened for specific exclusion criteria, including regular medication use and history of cardiovascular disease (CVD), diabetes mellitus, or autoimmune disorders (see Table S1 in the supplemental material). Volunteers who had received oral antibiotics or bowel preparations or had suffered gastroenteritis 1 month prior to study enrollment were excluded.

Eligible volunteers were randomized into 2 intervention groups: an exercise-only group (E group) and an exercise plus daily whey protein supplementation group (EP group) (Fig. 1). A separate parallel group consuming whey protein supplementation but not participating in exercise programs (P group) was included in the study as a control. To encourage recruitment to this control group, volunteers were offered an exercise

Downloaded from http://msystems.asm.org/ on April 27, 2018 by guest

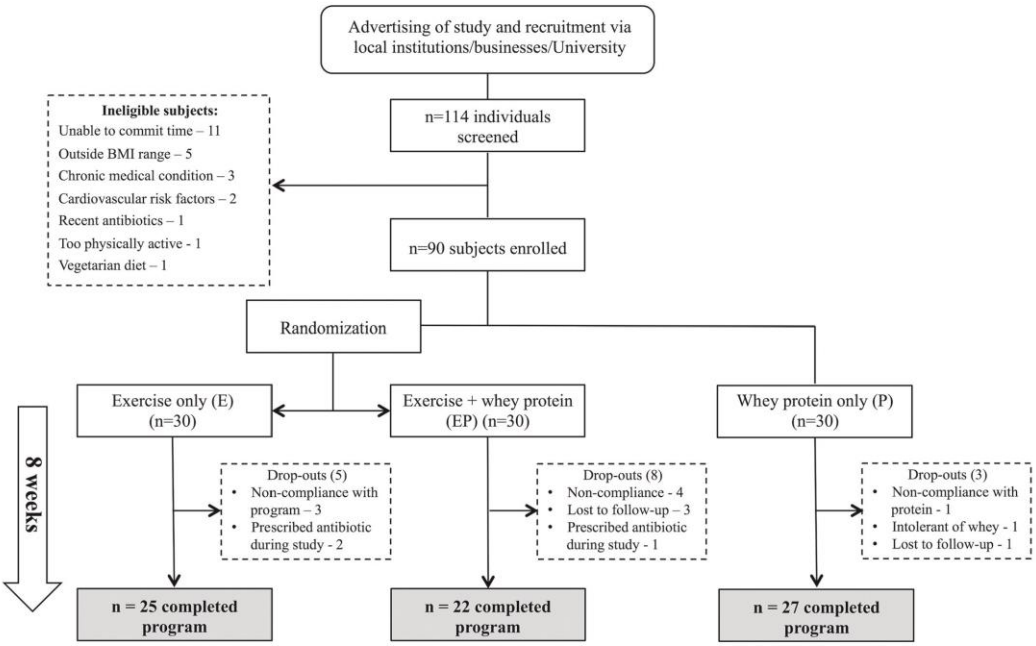


FIG 1 Study design. The figure presents details of study recruitment and allocation of participants to intervention groups as follows: exercise-only group (E), exercise and protein supplementation group (EP), and whey protein supplementation-only group (P). Reasons for volunteer dropout and completion numbers are also outlined. See also Table S1.

program at a later date, but their participation in the program after the conclusion of the study was not followed extensively. All participants were observed and measured for 8 weeks (n = 30 for each group). The exercise-only group (E) participated in an 8-week mixed aerobic and resistance exercise training program. The exercise plus whey protein supplementation group (EP) followed the same exercise program, in addition to consuming the once-daily whey protein supplement. All volunteers were asked to maintain their usual *ad libitum* dietary intake during the intervention period and to refrain from taking additional vitamin, dietary, or herbal supplements.

Participants in the E and EP groups were required to train 3 times per week for 8 weeks. The exercise program consisted of combined aerobic and resistance training. Aerobic exercise was standardized, progressive, and similar in energy expenditure to a "couch-to-5-km-running" program. The intensity of aerobic exercise was moderate, being graded at between 5 and 7 of 10 on the modified Borg rating of perceived exertion (RPE) scale (19). Resistance training consisted of 7 machine-based resistance exercises. Starting weights were calculated at induction at 70% of the individual's one-repetition maximum (1RM) value. Subjects were required to perform a minimum 3 sets of 8 repetitions. Resistance training was progressive, with aims of increasing resistance weight by 15% to 20% over the 8-week period.

To ensure a uniform and consistent increase in daily protein consumption, subjects in the P and EP arms of the study were required to take a daily 30-g protein supplement containing 24 g of whey protein (donated as an unrestricted grant by Carbery Group, Ballineen, Co. Cork, Ireland). The supplement comprised a blend of whey protein concentrate, isolate, and hydrolyzed whey protein concentrate (see Table S2 for full nutritional details). Subjects' compliance to daily whey protein supplementation was encouraged using daily text message reminders. Volunteers were required to return

Downloaded from http://msystems.asm.org/ on April 27, 2018 by guest

TABLE 1 Baseline demographic and anthropometric characteristics of the study participants with comparisons between the 3 intervention groups[a]

Patient characteristic	Values			*P* value
	Exercise (E) only (n = 25)	Exercise + protein (EP) (n = 22)	Protein only (P) (n = 27)	
Age (yrs)	35 (28, 38)	32 (28, 35)	34 (28, 36)	0.528
No. (%) of females[b]	n = 14 (56)	n = 12 (55)	n = 11 (41)	0.48
Height (cm)	172 (165, 181)	169 (166, 183)	172 (163, 178)	0.67
Weight (kg)	78.8 (70.1, 94.5)	82.3 (69, 98.9)	76.4 (69.8, 87)	0.67
BMI (kg/m^2)	27.9 (25.1, 29.2)	27.5 (25.7, 30)	27 (24.9, 28.7)	0.761
Resting heart rate (bpm)	72 (65, 81)	68 (61, 79)	74 (66, 78)	0.36
Systolic BP (mm Hg)	128 (117, 134)	125 (121, 136)	125 (118, 130)	0.706
Diastolic BP (mm Hg)	78 (74, 89)	76 (72, 84)	79 (75, 84)	0.543
Waist/hip ratio	0.85 (0.83, 0.89)	0.84 (0.8, 0.93)	0.83 (0.78, 0.88)	0.365
Body fat (%)	32.8 (29, 38.7)	34.7 (29, 37.2)	34.5 (29.3, 39.4)	0.659
Fat mass (kg)	26.3 (22.6, 30.6)	26 (23, 33.1)	26.8 (20.7, 32.9)	0.96
Fat mass (trunk) (kg)	14.1 (10.8, 16.8)	14.1 (11.2, 17.6)	13.7 (9.4, 17.1)	0.878
Lean tissue mass (kg)	52.4 (40.7, 61.4)	51.3 (41.5, 61.5)	47.2 (42.9, 53.3)	0.44
Weekly PA (METS)	462 (298, 1,139)	564 (413, 844)	657 (424, 1,145)	0.599
Weekly PA (kCal)	761 (381, 1,618)	748 (525, 1,127)	762 (512, 1,773)	0.767
Sitting time (h per wk)	56 (40, 61)	62 (47, 76)	51 (33, 62)	0.114
Motorized transport (h per wk)	5 (3.25, 8.3)	3.5 (2, 6)	4.1 (0.8, 7)	0.27

[a]Values represent medians (interquartile ranges) except where otherwise indicated. *P* values represent results of Kruskal-Wallis tests or chi-square tests. BMI, body mass index; IPAQ, International Physical Activity Questionnaire; METS, metabolic equivalents.
[b]Data indicate chi-square test results.

empty whey protein sachets to the study site fortnightly before the issuing of further supplement. Subjects with a compliance rate of less than 90% were excluded from the study.

Baseline measurements were not significantly different among the three study groups (Table 1).

Participants were predominantly overweight, with body fat percentages above 30%. There were no significant differences in the participants' baseline levels of physical activity as assessed using the International Physical Activity Questionnaire (20). All baseline values are expressed as medians and interquartile ranges (IQR).

Eight weeks of aerobic and resistance training improves body composition and cardiorespiratory fitness profiles in sedentary subjects. A total of 74 of the 90 participants enrolled in the study completed the 8-week study period (reasons for dropping out are detailed in Fig. 1). At entry, the intervention groups shared similar clinical and anthropometric characteristics. Following the intervention period, both E and EP group participants demonstrated significant and similar improvements in predicted maximal aerobic capacity (VO_{2max}) (Fig. 2A). Furthermore, resting heart rate was significantly reduced following the intervention period in both of the exercising groups (E and EP) compared with the protein-only group (P = 0.005) (Table S3). Compliance with the prescribed exercise program was high, with a median of 21 sessions (87.5%) performed in both the E and EP groups. The types and levels of exercise training undertaken in both groups were similar, with no statistically significant differences in the aerobic- and resistance-training workloads recorded (Table S4).

In contrast to the protein-only group, the exercise-only group and the exercise plus protein supplementation group experienced significant decreases in percentage body fat, total fat mass, and trunk fat mass during the intervention period (Fig. 2B and C), in addition to an increase in total lean tissue mass (Fig. 2D) (all $P < 0.001$; see also Table S3). Compliance with daily whey protein supplementation in the EP and P groups was high with only one participant excluded due to poor adherence to whey protein supplementation. Whey protein supplementation aside, dietary frequency patterns did not deviate from the volunteers' usual intake at study entry (Fig. S1). Addition of the 30 g daily whey protein supplement did not favor the EP over E group with respect to body composition improvement; however, the study was not designed to test this

Downloaded from http://msystems.asm.org/ on April 27, 2018 by guest

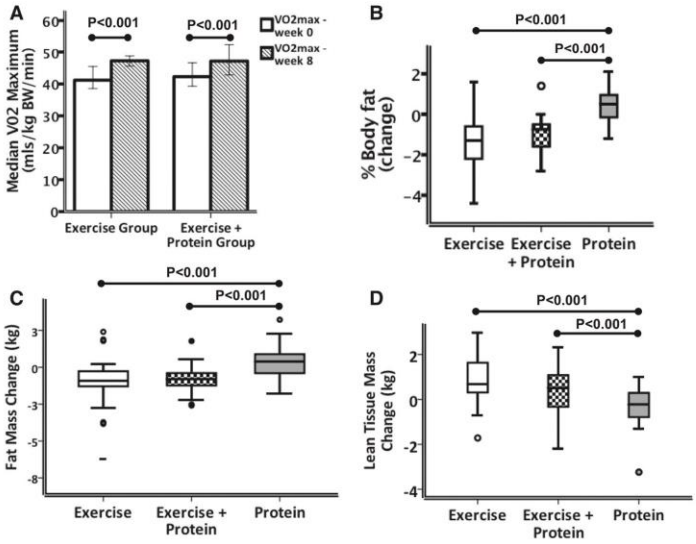


FIG 2 Alterations in cardiorespiratory fitness and body composition following exercise interventions, protein interventions, and combined interventions. (A) Peak aerobic capacity (VO_{2max}) per kilogram of body weight as predicted using the Rockport 1-mi walk test was higher in both the E and EP groups following the intervention period, indicating improved levels of cardiorespiratory fitness. Within-group comparisons were tested using the Wilcoxon signed-rank test ($P < 0.001$). (B) Changes in percentages of body fat following the intervention period as measured using DEXA. Percent body fat reduction was significantly greater in the exercise-only group and in the exercise plus protein supplementation group compared to the protein-only group. (C) Absolute changes in body fat mass (in kilograms) following the intervention period demonstrated a significantly greater reduction in both the exercise and exercise plus protein supplementation groups. (D) Absolute change in lean tissue mass (kg), measured using a three-compartment model, indicating significantly greater lean mass accretion in the E and EP groups than in the P group. Error bars represent 95% confidence intervals. See also Tables S3 to S5.

hypothesis. No clinically relevant differences in resting-state serum proinflammatory markers were evident following any of the interventions (Fig. S2; see also Table S5).

Metagenomic assessment of microbiota after exercise and/or dietary adjustment. Postintervention alterations (percent Δ) in gut microbial α-diversity did not identify significant modulation in taxonomic composition or metabolic pathways for any of the intervention groups compared to baseline (Fig. 3A to D). A trend of median increase in bacterial diversity was observed for the E and EP groups (Fig. 3B). These findings of moderate alterations of α-diversity were consistent across pairwise comparisons of the groups, with a few notable exceptions. Increased α-diversity of *Archaea* species in the P group following intervention was observed, as was a moderate enhancement of archaeal diversity in the P group compared to the EP group ($P < 0.05$ and $P < 0.01$, respectively; Fig. 3E). After the intervention period, bacterial diversity was greater in the EP group than in the P group ($P < 0.05$; Fig. 3F), while the diversity of virus species was lower in EP group than in the E group ($P < 0.05$; Fig. 3G).

Principal-coordinate analysis (PCoA) was used to present separation of measures from the taxonomic composition and metabolic pathway models (Fig. 3H to O). Prior to intervention, all 3 groups demonstrated similarity in measures of microbial metabolic pathways and taxonomic β-diversity (Fig. 3H to K). A significant separation between the intervention groups was detected in the Bray-Curtis-derived dissimilarity matrices generated from participants postintervention for metabolic pathways ($P = 0.054$; Fig. 3L), the entirety of detected species ($P < 0.001$; Fig. 3M), and species of bacteria ($P < 0.05$) and viruses ($P < 0.001$) (Fig. 3N and O, respectively). *Archaea* species did not differentiate with intervention (data not presented).

Downloaded from http://msystems.asm.org/ on April 27, 2018 by guest

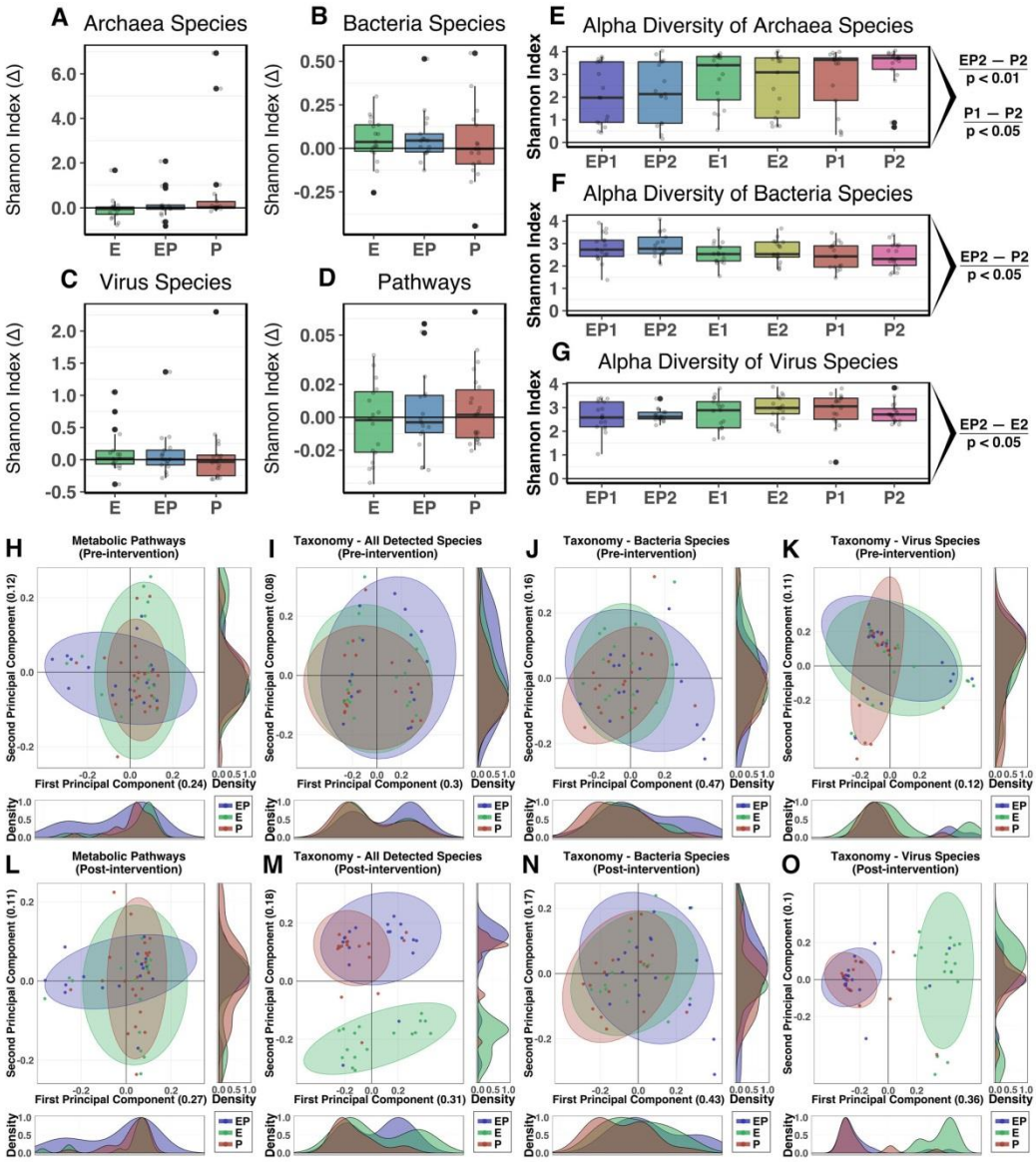


FIG 3 Intervention effects on taxonomic and functional pathway diversity of the intestinal microbiome. (A to D) Percent change (Δ) of Shannon α-diversity H-index values following intervention. No significant variations were presented for taxonomic measurements (A to C) or metabolic pathways (D). (E to G) Pairwise statistical assessment of taxonomy α-diversity demonstrates equal data with respect to the presence of taxonomy between groups at baseline. EP1, combined exercise and protein supplementation group, week 0; EP2, combined exercise and protein supplementation group, week 8; E1, exercise-only group, week 0; E2, exercise-only group, week 8; P1, protein-only group, week 0; P2, protein-only group, week 8. (E) The diversity of *Archaea* was significantly altered after intervention within the P group ($P < 0.05$) and, similarly, was greater in the P group (P2) than in the EP group (EP2) ($P < 0.01$). (F) Postintervention bacterial diversity was greater in the EP group (EP2) in testing against the P group (P2) ($P < 0.05$). (G) Similar levels of virus diversity were presented in the protein supplementation groups (EP and P) following the intervention, with significantly lower diversity in the EP group than in the E group ($P < 0.05$). (H to O) Principal-coordinate analysis (PCoA) of relative abundance profiles for taxonomic and metabolic pathway constructions of the three groups demonstrates the influence of the interventions on the diversity of microbial populations. (H to K) Prior to intervention, group profiles of taxonomic and metabolic pathway

(Continued on next page)

Downloaded from http://msystems.asm.org/ on April 27, 2018 by guest

Pairwise analysis of taxonomy compared according to high-level phylogeny (*Archaea*, *Bacteria*, and viruses) demonstrated significant alterations of detected virus species in both the EP and P groups following the intervention period that were absent from the exercise-only group ($P < 0.001$; Fig. 4). There were no further significant separations for *Archaea* or *Bacteria* species or for metabolic pathways (Fig. S3A to I). An unsupervised partial-least-squares–discriminant analysis (PLS-DA) approach was used to identify underlying features of the metabolic pathways before and after the intervention period (Fig. S3J and K). Pathways associated with *Prevotella copri* were shown to cluster with the E group prior to intervention. Following intervention, this cluster was still present but, in addition, separate clusters of *P. copri*- and *Bacteroides vulgatus*-associated pathways were apparent within the EP and P groups, respectively. Forty-eight species were detected as being differentially abundant within the three groups (false-discovery rate [FDR] = 0.05). The majority of identified taxa were virus species, predominantly *Lactococcus* phage, within the P and EP groups. No *Archaea* were found to have significantly varied in abundance with treatment in any of the groups (Table S6).

Of the total 23,019 unique (e.g., coenzyme A biosynthesis) and taxonomically specific (e.g., coenzyme A biosynthesis in *Akkermansia muciniphila*) metabolic pathways included in the metagenomic construction of the models, 619 were identified as having significantly differed among the three intervention groups at either the pretreatment or posttreatment time point ($P < 0.05$). Significantly altered pathways were organized according to a metabolic pathway hierarchy defined by the MetaCyc database and were structured as a heat map of low-level categories of classification (e.g., nucleotide biosynthesis) (Fig. S4). Scaled group means of pathway relative abundances demonstrated modest alterations of microbial metabolic potential. A complete list of the categorized pathways can be found in the supplemental material (Table S6). Further assessment of pathways differing among all groups was performed both within each group (before and after treatment) and between the separate groups. No significant variation within groups was evident following *P* value correction for multiple testing.

Untargeted metabolomic analysis of participant fecal-water and urinary samples revealed no significant separations either within each group pre- and postintervention or between groups at each time point with analysis of the full spectrum of metabolites. Subsequent targeted metabolomic quantification, guided by previous findings (11, 70, 71), revealed significant changes following intervention in the amount of glutamate (fecal water) and *trans*-aconitate (urine) in the protein-only group ($P < 0.01$ and $P < 0.05$, respectively; Table S6). Comparisons of differences (percent Δ) in metabolite quantifications between all groups demonstrated significant variation in the levels of phenylacetylglycine (PAG) and trimethylamine N-oxide (TMAO) ($P < 0.01$ and $P < 0.05$, respectively) in urine, as well as of glutamate ($P < 0.05$) in fecal water, within all groups (Table S6). Such differences were also present in the paired comparisons. Levels of both PAG and TMAO were significantly reduced ($\Delta = -0.196$ and -0.518, respectively) in the E group following the intervention period in comparison to the P group ($P < 0.05$, Table S6).

Characterization of whey protein supplement microbial content. Metagenomic sequencing of the whey protein supplement and of a non-dairy-based dietary supplement control revealed a taxonomic profile in the former that was characterized by high proportions of bacteriophage associated with lactic acid bacteria. Notably, these phage were also enriched in participants in receipt of the whey supplement (Table S6). The

FIG 3 Legend (Continued)
diversity were not significantly differentiated. (L to O) Following intervention, a significant separation was identified between the groups for measures of (L) metabolic pathways ($P = 0.054$), (M) all detected species unsegregated by phylogeny ($P < 0.001$), (N) bacteria ($P < 0.05$), and (O) virus species ($P < 0.001$). Specific separations in diversity per intervention group are outlined further in Fig. 4 for virus species and Fig. S3 for all other comparisons. Statistical assessment of PCoA dissimilarity matrices was performed with the Adonis2 permutational multivariate analysis of variance (PERMANOVA) test. (H to O) Density plots were derived from kernel density estimates and scaled to a maximum estimated value of 1 and display concentrations of plotted data along the corresponding plot axis. *P* values were calculated for α-diversity comparisons using the Wilcoxon signed-rank test.

Downloaded from http://msystems.asm.org/ on April 27, 2018 by guest

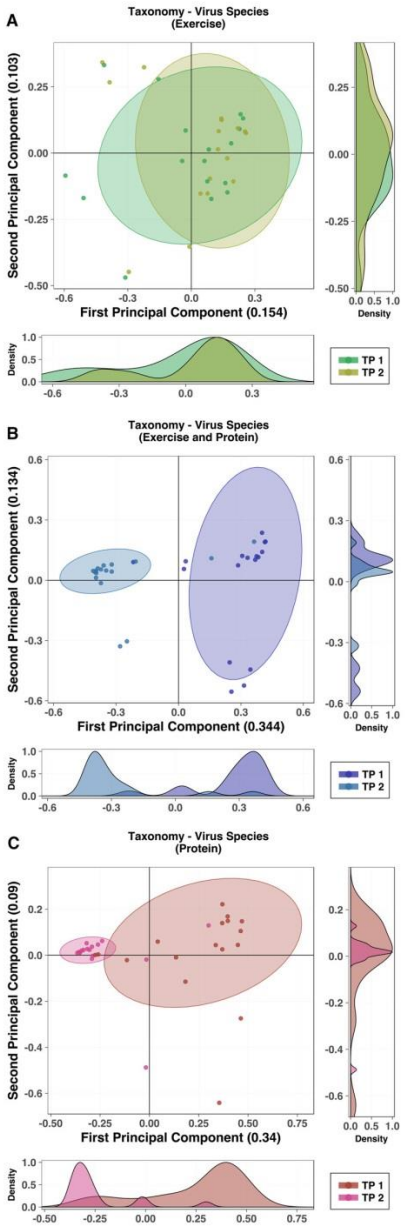


FIG 4 Pairwise analysis of detected virus taxonomy prior to and following intervention. (A to C) PCoA of virus species for each group, comparing virus profiles before and after the intervention period (time point 1 [TP1] [week 0] and time point 2 [TP2] [week 8], respectively). (A) The exercise-only group had virus diversity that was not significantly altered by intervention. (B and C) Diversity of viruses was significantly affected during the intervention period for both groups receiving protein supplementation ($P < 0.001$). The exercise plus protein supplementation group (B) and the protein-only group (C) demonstrated

(Continued on next page)

Downloaded from http://msystems.asm.org/ on April 27, 2018 by guest

taxonomic composition of the whey protein and control demonstrated highly divergent microbial contents of the supplements, including taxa detected in participants.

DISCUSSION

To accurately and consistently increase daily protein intake, we selected a whey protein supplement. Whey protein, a widely used commercial supplement in elite sport and amateur fitness milieu, is known for its muscle accretion effects (21), in addition to its positive influence on energy metabolism (22–24) and, more recently, on appetite control (25). In addition, its use facilitated analysis of the effect of a widely available exercise adjunct on the diversity, composition, and activity of microbial populations in the gut. Somewhat unexpectedly, individuals in the whey protein supplementation-only group (P) experienced a significant alteration in the β-diversity of the gut virome (Fig. 4C). Furthermore, this change was mirrored in the combined exercise and protein supplementation (EP) group (Fig. 4B), suggesting a robust effect of whey protein on the taxonomic richness of the gut virome. To explore this dynamic, a sample of the whey supplement and a sample of a non-dairy-based dietary supplement were sequenced for microbial content. Intriguingly, all bacteriophage and two of the four bacterial species that were significantly altered in the groups receiving whey protein were present in high relative abundance within the whey protein supplement but not the control supplement. Further in-depth experimentation is required to determine whether virus particles from whey protein conclusively transmit to the human gut from consumption and, if so, whether they remain biologically active. However, the overlap in the taxonomic compositions of the whey supplement and participants' gut microbiome provides a convincing explanation for the source of virome changes observed.

While this examination did not identify a significant impact of short-term combined aerobic and resistance exercise on the diversity of bacterial or archaeal constituents of the gut microbiome, subtle compositional and functional changes were detected in this analysis (Fig. 3 and 4; see also Fig. S3 and S4). Although the results were not statistically significant, the groups engaged in exercise demonstrated less change in archaeal diversity than the protein-only group after the intervention period. In the case of the exercise-only group, a reduction in *Archaea* diversity was observed, suggesting that exercise acts against intestinal *Archaea*. More-extensive investigation is necessary to resolve this issue, but in view of a putative role for *Archaea* in intestinal disorders and as modulators of TMAO concentrations, such an inquiry is justified (26). Changes in *Bacteria* diversity were similarly below the threshold of statistical significance; however, the median differences between groups indicated that those undertaking exercise had increases in bacterial diversity that were absent from the intervention group excluded from exercise. Curiously, the diversity of bacterial species was elevated in the EP group after intervention but the diversity of virus species was uniformly lower. The inverse relationship of these measures is counterintuitive given the predominance of bacteriophage in the detected viruses. However, the influx of such bacteriophage may explain the overall reduction of virus diversity within the group. Furthermore, this increase in the levels of bacteriophage may have been insufficient to profoundly influence the overall diversity of the *Bacteria* due to their selective targeting of only a few bacterial species.

The absence of substantial modulation of the diversity of microbial populations in the gut following the 8-week exercise intervention mirrors recent findings in mice (27). To date, most of the work in humans has focused on elite or professional athletes (10–12) and as a result has explored the relationship between established physical

FIG 4 Legend (Continued)
reduced variability of diversity following intervention. Results of pairwise analysis of additional taxonomic and metabolic pathway profiles are presented in Fig. S3. Statistical assessment of PCoA dissimilarity matrices was performed with the Adonis2 PERMANOVA test. (A to C) Density plots were derived from kernel density estimates and scaled to a maximum estimated value of 1 and display concentrations of plotted data along the corresponding plot axis.

Downloaded from http://msystems.asm.org/ on April 27, 2018 by guest

"fitness" and the gut microbiota. Few prospective studies have examined the effect of exercise on the gut microbiota of physically inactive human volunteers (28). The current study is the largest to have done so. It should be acknowledged that the unperturbed adult intestinal microbiome is resilient (29) and may not be subject to significant alteration following an 8-week intervention period. It is likely that the diverse, metabolically favorable intestinal microbiome evident in the elite athlete is the cumulative manifestation of many years of optimized nutrition and of high degrees of physical condition throughout youth and adolescence and during adult participation in professional sports (30). Initial examination of the acute effects of extreme and prolonged endurance exercise, such as in trained military regiments, suggests that prolonged physical stress negatively impacts intestinal permeability and gut microbiota composition (31). However, the results of the present study indicate that exercise at moderate intensity does not exert a deleterious effect on gut microbial composition or function in the untrained subject. Furthermore, the results of this study signify that exercise-induced improvements in cardiorespiratory fitness and body composition are not dependent on substantial alteration of the diversity of microbial populations in the gut. Whether the limited changes in microbiome composition and function detected in this study contributed to the witnessed improvements in body composition and fitness profiles remains unknown.

An intriguing exception to the otherwise minimal differences in metabolomic modification is represented by the controversial metabolite TMAO. Associations between TMAO and cardiovascular disease (CVD) have framed the metabolite as a disease factor; however, high levels have also been observed in populations with low CVD risk (32–35). Elevated levels of TMAO have previously been found in elite athletes (11), and while the presence of TMAO may or may not have deleterious health implications, we demonstrate a potential modulatory effect of exercise on urinary TMAO levels. Participants in the exercise-only group showed levels of urinary TMAO that were reduced below baseline with intervention, while the groups receiving whey protein had increased levels of the metabolite, with the combined-treatment group demonstrating lower levels than the protein-only group. The TMAO precursor phosphatidylcholine comprised less than 0.1% of the constituents of the whey protein supplement used in this study, suggesting a possible direct effect of whey protein and/or exercise on TMAO production. While the data represent a promising paradigm, further work will be necessary to determine the specific mechanisms involved and to rule out unintended dietary influence or influences of host biology (e.g., altered absorption of TMAO with exercise). Additionally, known microbial producers of TMAO (36) were absent from the taxonomic profiling. PAG concentrations were similarly reduced in the exercise-only group, although the metabolite has previously been associated with lean body composition and has been found to be present in increased concentrations in athletes. It has also recently been shown to decrease in urine in thoroughbred racehorses following exercise (37).

Likewise, the data reflecting an increase in the abundance of the *P. copri*-associated pathways detected in the EP group postintervention supports the work of others which suggested an active role for *Prevotella* species in host metabolic (38, 39) and immune health (40). Studies have linked *Prevotella* with inflammatory and metabolic disorders, including rheumatoid arthritis (41, 42), ankylosing spondylitis (43), and type 2 diabetes mellitus (44). Conversely, and consistent with our findings, increased physical activity has been associated with increases in *Prevotella*-related metabolic pathways in the gut microbiome (10).

It is pertinent to acknowledge the difficulty in controlling all potential confounders of gut microbial composition and activity in this investigation (e.g., diet, wide BMI range). This study attempted to control for the potential impact of dietary variation by instruction of volunteers to maintain their usual *ad libitum* dietary intake. Food frequency questionnaire (FFQ) dietary analysis indicated stability in the volunteers' dietary patterns; however, like all self-reported methods of dietary intake assessment, FFQ assessment is subject to its limitations, including recall bias (45).

Downloaded from http://msystems.asm.org/ on April 27, 2018 by guest

In conclusion, this prospective examination demonstrated that short-to-medium-term combined exercise in healthy, physically inactive adults does not induce drastic alterations in the diversity of bacterial, viral, and archaeal populations in the gut. We highlight an interaction between whey protein intake and the β-diversity of the adult gut virome which requires further exploration. Furthermore, the functional activity of the gut microbiota does not appear to be extensively manipulated by short-term, moderate-intensity exercise and/or whey protein supplementation, although some changes, including alteration of levels of urinary TMAO and PAG excretion, were evident. The alterations in the diversity, composition, and metabolomic profiles of microbial gut populations that we and others have observed in habitual exercisers and professional athletes may represent late responses to exercise or fitness.

MATERIALS AND METHODS

Experimental models and subject details. A description of the human study model used here is outlined in Results under "Study overview." Male and female volunteers were enrolled. The Cork Clinical Research Ethics Committee (CREC) approved the study before it commenced. Recruitment and assignment of interventions are outlined in detail below.

Study recruitment and safe participation. Male and female participants, aged between 18 and 40 years (inclusive), were recruited via online, e-mail, and poster advertisement of the study details. This information was circulated to the study institutions (University College Cork and Cork University Hospital) and local businesses in Cork City, Ireland. Participants were informed that free gymnasium membership would be supplied for the study period. Interested individuals contacted researchers via the study telephone line and were screened initially for inclusion criteria (see Table S1 in the supplemental material). Baseline levels of physical activity were assessed using the International Physical Activity Questionnaire short form (46). If appropriate, a subsequent screening visit at the study site was arranged for further assessment of the exclusion criteria. Safe participation in the exercise program was ensured by medical screening of all participants using an adapted version of the safe participation questionnaire of the American College of Sports Medicine (47).

Intervention group allocation. Eligible volunteers were randomized into 2 intervention groups, namely, an exercise-only group (E group) and an exercise plus daily whey protein supplementation group (EP group) (Fig. 1). A separate parallel group consuming whey protein supplementation alone (P group) was included in the study as a control. Participants in the P group were instructed to maintain their usual levels of light physical activity. To encourage recruitment to the control group, volunteers were offered an exercise program at a later date but were not followed extensively during that period. All participants were observed and their responses measured for 8 weeks ($n = 30$ for each group).

Combined exercise intervention. Combined aerobic and resistance training was performed at the Mardyke Arena gymnasium at University College Cork, Ireland. All exercise sessions took place at this venue. Volunteers in the P group were asked to maintain their usual levels of physical inactivity for the 8-week period. Participants in the E and EP groups were instructed to adhere to the assigned exercise program and to avoid additional (moderate to vigorous) physical activity outside that prescribed. Participants were instructed to train 3 times per week for the 8-week intervention period. Participants received instructions with respect to the format of the required training program during a 90-min induction session with designated gym instructors. This induction included demonstration of all aerobic and resistance training equipment and the opportunity to ask questions when required. Resistance training machines were customized for individual differences in ranges of motion, and the participants were observed during use of all of the machines, with instructor feedback and correction. For resistance machines, 1-repetition maximum (1RM) values were calculated using standardized methods (48).

The outline of the exercise sessions was as follows. After a 5-min warmup on the treadmill (brisk walking at approximately 4 km/h; modified Borg RPE, 3 to 5/10), participants underwent aerobic training of moderate intensity (modified Borg RPE, 5 to 7/10). To encourage compliance with the prescribed RPE scales, volunteers were reminded of the desired intensities on each of their weekly exercise training programs. The aerobic exercise progressed in duration on a weekly basis but remained of moderate intensity. Initially, aerobic exercise lasted approximately 18 min; by week 8 of the intervention period, the duration of aerobic exercise increased to approximately 32 min depending on the type of aerobic activity chosen by the volunteer. Participants were provided with a choice of aerobic activities, including treadmill jogging/running, use of a cross-trainer device (with no added resistance), use of a stepper machine, and stationary cycling (with mild resistance). The duration for each of these activities was calculated based on the 2011 Compendium of Physical Activities (49) to ensure similar levels of energy expenditure across all activities. To allow variety and to maintain interest, participants were provided with an option of aerobic activities, provided that they did not change exercises within a given training session.

Upon completion of the aerobic exercise activity, participants undertook machine-based resistance training. In summary, participants were required to perform a minimum of 3 sets of 8 repetitions on 7 different resistance machines (3 exercising the upper body, 3 exercising the lower body, and 1 exercising the core muscles). The allowed resistance machine options were as follows: for the upper body, shoulder press, chest press, lateral pulldowns, and seated rowing; for the lower body, leg extension, leg curl, gluteal kick-back, and leg press; for the core muscles, abdominal curls and torso

Downloaded from http://msystems.asm.org/ on April 27, 2018 by guest

rotation. A minimum limit of 3 sets of 8 repetitions was instituted, with a maximum limit of 3 sets of 12 repetitions. Starting weights were calculated at induction to correspond to 70% of the individual's one-repetition maximum (1RM) value. Resistance training was progressive, with the aim of increasing the resistance weight by 15% to 20% over an 8-week period. Free-weight use was not permitted.

Compliance and withdrawal from the study. Compliance with the prescribed exercise program was monitored remotely by the investigators using a FitLinxx activity monitoring system (Activelinxx, Shelton, CT). All volunteers were provided with a unique identification number for the FitLinxx physical activity recording system and were required to log in and record all activities undertaken at the Mardyke Arena gym during training. Using this tracking system, compliance with the prescribed exercise program was monitored by the investigators. Similarly, the quantity of aerobic and resistance training performed by participants was recorded (Table S4). The FitLinxx software program enables accurate recording of the duration and frequency of training and provides an estimate of energy expenditure during aerobic training. The facility's FitLinxx software and hardware were subjected to maintenance and recalibration prior to commencement of the study.

Participants noted to have not complied with the exercise regime for more than 7 consecutive days were withdrawn from the study. Individuals requiring antibiotics during the intervention period were also withdrawn from the study, as were participants not complying with whey protein intake requirement in the EP and P groups.

Measurement visits. Measurement visits took place at 2 sites: Cork University Hospital and the Mardyke Arena, University College Cork. Baseline measurement was conducted within the 4 days prior to the commencement of the intervention period and once more after the 8-week intervention. Participants were asked to refrain from the use of alcohol and medication and moderate to vigorous physical activity for at least 24 h prior to measurement. To minimize potential effects of diurnal variation, measurement visits took place between 7:00 a.m. and 10:30 a.m. Initially, participants attended the Department of Medicine research facility at Cork University Hospital and sat restfully in a quiet environment. Participants proceeded to participate in measurement of clinical variables, e.g., recording of weight, blood pressure, and heart rate, before undergoing phlebotomy by a trained nurse using universal precautions. Approximately 16 ml of venous blood was withdrawn. Plasma and serum samples were transported immediately to the clinic laboratories at the Mercy University Hospital, Cork. Standardized laboratory techniques were employed for the measurement of hematology and biochemistry indices. Following phlebotomy, individuals underwent a total body dual-energy X-ray absorptiometry (DEXA) scan to assess body composition. When possible, volunteers were asked to provide fresh urine and fecal samples, which were transported at room temperature to Teagasc Moorepark, Fermoy, Co. Cork, where DNA extraction took place. Following completion of the body composition assessment, participants proceeded to the indoor track at the Mardyke Arena gymnasium to undergo a submaximal cardiorespiratory fitness assessment as described below.

VO_{2max} and body composition measurement. To prevent injury from unaccustomed vigorous exercise, we chose a submaximal assessment of peak aerobic capacity. Baseline and postintervention levels of cardiorespiratory fitness were measured using a validated submaximal fitness test (50). The Rockport 1-mi walk test was performed in a standardized temperature environment at the indoor running track of the Mardyke Arena. This test was used to estimate maximal oxygen uptake (VO_{2max}).

A Lunar iDXA machine (GE Healthcare, Madison, WI) at the Bone Densitometry Unit, Cork University Hospital, was used. Body composition was analyzed using enCORE software (V.13.4, 2010) and a three-compartment (fat mass, bone mass, lean tissue) body composition model. Volunteers were scanned postvoiding and dressed in light clothing, with metal-wear removed where present. Quality control (QC) analysis was performed on the iDXA machine before use of the machine on each measurement day.

Inflammatory cytokine measurement. Blood samples (4 ml) from participants were collected in serum separator clot activator blood collection tubes (Greiner Bio-One, Stonehouse, United Kingdom; reference no. 454071). The blood samples were allowed to rest upright on the laboratory bench for 30 min before centrifugation was performed at 1,000 × *g* for 20 min at room temperature. Approximately 2 ml of supernatant sera was then harvested by pipette, frozen, and stored at −80°C in polypropylene cryogenic vials. At a later date, following a complete thaw, resting levels of proinflammatory cytokines were measured using a mesoscale discovery (MSD) platform (Meso Scale Discovery, Rockville, MD). The MSD system is an electrochemiluminescence-based solid-phase multiplex assay. An ultrasensitive human proinflammatory I, V-Plex immunoassay panel containing interleukin-6 (IL-6), IL-8, tumor necrosis factor alpha (TNF-α), IL-10, and gamma interferon was used to measure serum levels. Samples were diluted 1:2 according to the manufacturer's protocol, and samples from all 3 intervention groups were dispersed across each MSD plate. The lower limit of detection was <1 pg/ml for all assays. All plasma samples were measured in duplicate, and the mean cytokine concentration of the duplicates (in picograms per milliliter) was used for analysis.

Dietary data collection. Dietary data were collected by means of a 146-item food frequency questionnaire (FFQ) as outlined previously (12). Participants were asked to record their usual pattern of dietary intake over the previous 8 weeks. The FFQ used was an adapted version of the questionnaire used in the United Kingdom arm of the European Prospective Investigation into Cancer (EPIC) study (51), which was based on the original Willet FFQ (52). Completed FFQs were coded and dietary data were visualized with correspondence analysis using the ade4 package (53) in the R programing environment (V.3.3.2).

DNA extraction and metagenomic sequencing of fecal microbiome and whey protein supplement. DNA was extracted from the donated fresh fecal samples received at the Teagasc Moorepark

Downloaded from http://msystems.asm.org/ on April 27, 2018 by guest

research facility using a QIAmp DNA stool minikit (Qiagen, Crawley, West Sussex, United Kingdom) (54). Samples were provided by participants as partial evacuations into sterile containers and, when not immediately transported for DNA extraction, were held at 4°C for no more than 12 h. Samples were prepared for DNA extraction by manual homogenization of a portion of the sample representing all microenvironments (i.e., core and external surface) of the feces. The provided manufacturer's protocol was used with modification whereby a zirconia bead (Stratech Scientific) cell disruption bead-beating step (performed three times for 30 s each time) was introduced in order to enhance homogenization of the extraction material. DNA extracts and the remaining fecal samples were subsequently stored at −80°C until they were prepared for sequencing.

Metagenomic libraries were prepared using an Illumina Nextera XT DNA library preparation kit (Illumina Inc., USA) in complete accordance with the manufacturer's protocol (15031942; Illumina). Library input DNA was normalized to 0.2 ng/μl using fluorometric quantification and Qubit system 2.0 (Thermo, Fisher). Library fragmentation and amplification were performed using a G-Storm GS1 thermal cycler, with subsequent library purification achieved with AMPure magnetic beads (Beckman Coulter, Inc.) at a DNA/AMPure ratio of 1:1.8. Confirmation of the sizes of the library fragments was carried out on an Agilent 2100 Bioanalyzer system, with final assessment of library molar concentration (2 nM) performed using a Roche LightCycler 480 instrument (Roche Applied Science) and a Kapa library quantification kit (Kapa Biosystems). A total of 8 equimolar library pools of samples were made prior to shipping of the pools on dry ice for sequencing on an Illumina Hiseq 2500 (chemistry V.4.0) sequencing platform (Beckman Coulter, Inc.; Genomics Inc., Danvers, MA). High-throughput sequencing was performed using the high-output run mode for 2 × 125-bp paired-end reads with the addition of a PhiX library (1%) to estimate sequence quality. A sample of the whey protein used in the study and a sample of an oat-based nutritional supplement used as a control were both processed in a manner identical to that used with the fecal samples for the extraction of microbial DNA and preparation of metagenomic libraries. Sequencing of the supplement libraries was performed using an Illumina MiSeq (chemistry V.3.0) platform in high-output run mode for 2 × 300-bp paired-end reads (Teagasc sequencing facility).

Bioinformatic processing of microbial metagenomic sequencing. Processing of metagenomic FASTQ sequence files proceeded with the removal of human-derived contaminant sequences with NCBI Best Match Tagger (BMTagger) software, while trimming and removal of duplicate reads or of reads of substandard quality were performed with Picard and SAM tools. Functional profiling of high-quality processed reads was facilitated by use of the Human Microbiome Project (HMP) Unified Metabolic Analysis Network (HUMAnN2 V.0.99) pipeline (55). Models of microbial metabolic pathways produced by HUMAnN2 were derived from the MetaCyc database (56) and were the basis for analyses performed on microbial metabolic profiling. Taxonomic profiling was facilitated by use of the Kraken taxonomy assignment software tool (V.0.10.6) (57).

Metabolomic sample preparation. Samples were stored at −80°C prior to analysis. Urine samples were subjected to vortex mixing and then centrifuged at 1,600 × *g* for 10 min to remove precipitated proteins and particulates. For metabolic profiling analysis by reversed-phase (RP) and hydrophilic interaction chromatography (HILIC) ultraperformance liquid chromatography-mass spectrometry (UPLC-MS), samples were prepared as follows: 200 μl of supernatant was diluted (1:1) with high-purity (ultraperformance liquid chromatography [HPLC]-grade) water, subjected to vortex mixing, centrifuged at 2,700 × *g* for 20 min, and divided into aliquots for analysis. Quality control (QC) samples were prepared by pooling 50-μl volumes of each sample. For ^{1}H nuclear magnetic resonance (^{1}H-NMR) spectroscopy, each sample contained 540 μl of urine mixed with 60 μl of phosphate buffer (pH 7.4; 80% D_2O) containing a 1 mM concentration of the internal standard, 3-(trimethylsilyl)-[2,2,3,3,-2H4]-propionic acid (TSP)–2 mM sodium azide (Na^3N), as described previously (58). During the analyses, samples were maintained at 4°C in the autosampler.

Fecal samples underwent two freeze-thaw cycles. Following the freeze-thaw cycles, 100 mg of homogenized sample was placed in a microtube containing 250 μl of 25% acetonitrile (ACN) (1:2 ACN/H_2O), 2 mM sodium azide, and ~0.05 g 1-mm-diameter zirconia beads. Each microtube was processed for 10 s in a Biospec bead beater. Samples were then centrifuged at 16,000 × *g* for 20 min. The fecal-water supernatant was subsequently centrifuged through centrifuge tube filters (cellulose acetate membrane; pore size, 0.22 μm) to remove any remaining particulate matter. The centrifuge tube filters were washed three times with 25% acetonitrile prior to use. The resulting fecal water was prepared for UPLC-MS profiling using HILIC by diluting 3:1 with acetonitrile and for bile acid profiling by diluting 1:1 with isopropanol. Samples were subjected to vortex mixing and incubated at −20°C for 1 h. Following the incubation step, samples were centrifuged at 4°C at 16,000 × *g* for 1 h and divided into aliquots for analysis. QC samples were prepared by pooling 20-μl volumes of each fecal-water sample followed by preparation as described above. For ^{1}H-NMR spectroscopy, 50 μl of the filtered fecal water was added to a glass tube (Pyrex), which was placed under a nitrogen gas flow for 30 min or until all the liquid had evaporated. The dried sample was reconstituted with 540 μl of D_2O and 60 μl of phosphate buffer solution as described above. The solution was mixed and sonicated for 5 min before undergoing further centrifugation at 14,000 rpm for 10 min, and then 600 μl of the supernatant was transferred to an NMR tube for ^{1}H-NMR spectral acquisition.

Metabolomic analysis. RP, HILIC, and bile acid UPLC-MS metabolic profiling experiments were performed using a Waters Acquity Ultra Performance LC system (Waters, Milford, MA) coupled to a Xevo G2 quadrupole-time of flight (Q-TOF) mass spectrometer (Waters, Milford, MA) with an electrospray source. Samples were analyzed in randomized order, with QC analyzes performed every 10 samples. First, urine samples were analyzed using UPLC-MS and an RP chromatographic method with both positive and negative MS ionization modes. Second, to separate and detect the more polar molecules, a HILIC

Downloaded from http://msystems.asm.org/ on April 27, 2018 by guest

chromatographic stage was used with the positive MS ionization mode. Fecal-water samples underwent analysis using HILIC and bile acid profiling chromatographic methods in positive and negative ionization modes, respectively. HILIC, RP, and bile acid profiling liquid chromatographic separation procedures were performed as previously described (59, 60). Mass spectrometry was performed with the following settings. Capillary and cone voltages were set at 1.5 kV and 30 V, respectively. The desolvation gas level was set at 1,000 liters/h at a temperature of 600°C. The cone gas level was set to 50 liters/h. The source temperature was set to 120°C. To ensure the accuracy of the mass data, a lock-spray interface was used, with leucine enkephalin {556.27741 Da ([M+H]+), 554.2615 Da ([M−H]−)} solution used as the lock mass at a concentration of 2,000 ng/ml and a flow rate of 15 μl/min.

^{1}H-NMR spectroscopy was performed on the aqueous-phase extracts at 300 K on a Bruker 600-MHz spectrometer (Bruker Biospin, Germany) using a standard one-dimensional (1D) pulse sequence corresponding to RD − g_{z1} − 90° − t_1 −90° − t_m − g_{z2} − 90°− ACQ (58), where the value of 90° represents the applied 90° radio frequency pulse; the relaxation delay (RD) was set at 4 s, the interpulse delay (t_1) was set at 4 μs, the mixing time (t_m) was set at 10 ms, the magnetic field gradients (g_{z1} and g_{z2}) were applied for 1 ms, and the acquisition period (AQA) was 2.7 s. Water suppression was achieved through irradiation of the water signal during RD and t_m. Urine sample spectra were acquired using 4 dummy scans followed by 32 scans whereas fecal spectra were acquired using 256 scans and 4 dummy scans and collected into 64 K data points. A spectral width of 12,000 Hz was used for all the samples. Prior to Fourier transformation, the free induction decay (FID) values were multiplied by an exponential function corresponding to a line broadening of 0.3 Hz.

Metabolomic data treatment. The raw mass spectrometric data acquired were preprocessed using xcms in R. centWave peak picking methods were used to detect chromatographic peaks (61). The xcms-centWave parameters were data set specific. Feature grouping across samples was performed using the "nearest" method within xcms. Peak filling and MinFrac (0.5), and coefficient of variation (CV) (0.3) filters were applied to the features. Data were normalized using median fold change normalization to the median data set (62).

^{1}H-NMR spectra were automatically corrected for phase and baseline distortions and referenced to the TSP singlet at δ 0.0 using TopSpin 3.1 software. Spectra were then digitized into 20 K data points at a resolution of 0.0005 ppm using an in-house MatLab R2014a (MathWorks, Inc.) script. Subsequently, spectral regions corresponding to the internal standard (δ −0.5 to 0.5) and water (δ 4.6 to 5) peaks were removed. In addition, urea spectra (δ 5.4 to 6.3) were removed from the urinary spectra. Spectra were normalized using median fold change normalization to the median spectrum (62). Combinations of data-driven strategies, such as SubseT optimization by reference matching (STORM) (63) and Statistical TOtal Correlation SpectroscopY (STOCSY) (64), and analytical identification strategies were used to identify metabolites of interest from ^{1}H-NMR data sets. Specifically, a catalogue of 1D ^{1}H-NMR and 2D NMR experiments was performed using techniques such as J-RESolved spectroscopy, ^{1}H-^{1}H TOtal Correlation SpectroscopY (TOCSY), ^{1}H-^{1}H COrrelation SpectroscopY (COSY), ^{1}H-^{13}C Hetero-nuclear Single Quantum Coherence (HSQC), and ^{1}H-^{13}C Heteronuclear Multiple-Bond Correlation (HMBC) spectroscopy. Finally, for those metabolites giving ambiguous data, e.g., TMAO, the metabolites were confirmed using *in situ* spiking experiments and authentic chemical standards. Semiquantification data corresponding to the identified metabolites were calculated through peak intensity measurements of the normalized ^{1}H-NMR spectra using an in-house script. GC-MS data were processed using MassHunter Quantitative Analysis (Agilent Technologies) software.

Quantification and statistical analysis. Statistical analysis was carried out using the Statistical Package for the Social Sciences V.23 (SPSS, Inc., Chicago, Illinois) and the R statistical programing environment (V.3.3.2). Due to the predominance of non-normally distributed data, nonparametric analyses were performed to compare baseline clinical and demographic variables between groups. Similarly, nonparametric statistical tests were employed in the analysis of microbiome and metabolomics data. Clinical data are presented as medians and interquartile ranges (IQR), unless stated otherwise. Between-group differences in baseline, follow-up, and postintervention changes (Δ) in clinical and demographic data were compared using the Kruskal-Wallis test. For significantly different results, a Mann-Whitney *U* test was performed to determine the groups between which this difference applied. Where stated, the Wilcoxon signed-rank test was used to compare baseline and postintervention values within intervention groups. A type I error rate of $\leq$0.05 was considered significant in all cases. Correction of *P* values relating to microbiome and metabolomic analysis was performed using the Benjamini-Hochberg false-discovery rate (FDR) (65) in the base *stats* package in R. Statistical assessment of dissimilarity matrices (Bray-Curtis) derived from microbial data was facilitated with the adonis2 function in the vegan R package (V.2.4-3) (66). Identification of statistically relevant taxonomic features was performed with the analysis of composition of microbiomes (ANCOM) test as implemented in the R package of the same name (V.1.1-3) (67). Detection of underlying features of metabolic pathways was performed with unsupervised cross-validated partial-least-squares–discriminant analysis (PLS-DA) and the KODAMA algorithm from the R package of the same name (V.1.4) (68). Measurements of alpha diversity and calculations of relative abundances were also performed with the vegan package. Relative-abundance data were generated separately for identified species within each phylogenetic domain (e.g., *Bacteria*).

For metabolomic analysis, the resulting ^{1}H-NMR and LC-MS data sets were imported into MatLab to conduct multivariate statistical analysis. Data were centered and scaled to account for the repeated-measures design and then modeled using partial-least-squares–discriminant analysis (PLS-DA) with Monte Carlo cross-validation (MCCV) (69). The fit and predictability of the models obtained were determined and expressed as R2 and Q2 values, respectively.

Downloaded from http://msystems.asm.org/ on April 27, 2018 by guest

Data availability. The microbial DNA sequences have been deposited in the European Nucleotide Database (ENA) database under ID code PRJEB20054.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at https://doi.org/10.1128/mSystems.00044-18.

FIG S1, PDF file, 1 MB.
FIG S2, GIF file, 0.01 MB.
FIG S3, PDF file, 2.6 MB.
FIG S4, PDF file, 0.1 MB.
TABLE S1, DOCX file, 0.1 MB.
TABLE S2, DOCX file, 0.04 MB.
TABLE S3, DOCX file, 0.1 MB.
TABLE S4, DOCX file, 0.1 MB.
TABLE S5, DOCX file, 0.1 MB.
TABLE S6, XLSX file, 2 MB.

ACKNOWLEDGMENTS

We thank the staff of the Mardyke Arena and Mercy University Hospital, Cork, Ireland, for their assistance in conducting this study, in addition to the study volunteers for their time and efforts.

This research was funded by Science Foundation Ireland (SFI) in the form of a center grant (APC microbiome Ireland grant SFI/12/RC/2273) and by the Irish Centre for Arthritis Research and Education (I CARE). Research in the Cotter laboratory is funded by SFI through the PI award "Obesibiotics" (11/PI/1137). O.O. and W.B. are funded by SFI through a Starting Investigator research grant award (13/SIRG/2160). N.C.P. is funded by the Diabetes Research & Wellness Foundation through the Sutherland-Earl Clinical Research Fellowship 2015. We thank the Imperial-National Institute for Health Research (NIHR) Clinical Phenome Centre for support. The center is supported by the NIHR Imperial Biomedical Research Centre based at Imperial College London Healthcare National Health Service (NHS) Trust and Imperial College London. The views expressed are ours and not necessarily those of the NHS, the NIHR, or the Department of Health. Carbery Foods Ltd., Ballineen, Co. Cork, donated whey protein supplements for use in this study.

F.S. is a founder shareholder in Atlantia Food Clinical Trials, Tucana Health Ltd., and Alimentary Health Ltd. He is director of the APC Microbiome Ireland, a research center that is funded in part by Science Foundation Ireland (APC/SFI/12/RC/2273) and which is/has recently been in receipt of research grants from Abbvie, Alimentary Health, Cremo, Danone, Janssen, Friesland Campina, General Mills, Kerry, MeadJohnson, Nutricia, 4D Pharma, Second Genome, and Sigmoid Pharma.

F.S., M.G.M., P.D.C., O.O., E.C.F., E.F.M., and O.C. conceived the study design. O.C. and T.W. formulated the exercise intervention with input from E.C.F. O.C. and H.N. conducted study recruitment, enrollment, interventions, and clinical measurement. P.S., W.B., and O.O. conducted DNA extraction from feces, subsequent sequencing, and metagenomic analysis. N.C.P., I.G.-P., and E.H. were responsible for metabolomic analysis. A.F. and S.M. performed and contributed to inflammatory cytokine measurement and analysis. O.C., W.B., O.O., N.C.P., P.D.C., and F.S. wrote the manuscript with review and editing from all of us.

REFERENCES

1. Debelius J, Song SJ, Vazquez-Baeza Y, Xu ZZ, Gonzalez A, Knight R. 2016. Tiny microbes, enormous impacts: what matters in gut microbiome studies? Genome Biol 17:217. https://doi.org/10.1186/s13059-016-1086-x.
2. Sonnenburg JL, Bäckhed F. 2016. Diet-microbiota interactions as moderators of human metabolism. Nature 535:56–64. https://doi.org/10.1038/nature18846.
3. Shanahan F, van Sinderen D, O'Toole PW, Stanton C. 2017. Feeding the microbiota: transducer of nutrient signals for the host. Gut 66:1709–1717. https://doi.org/10.1136/gutjnl-2017-313872.
4. Bermon S, Petriz B, Kajėnienė A, Prestes J, Castell L, Franco OL. 2015. The microbiota: an exercise immunology perspective. Exerc Immunol Rev 21:70–79.
5. Cerdá B, Pérez M, Pérez-Santiago JD, Tornero-Aguilera JF, González-

Downloaded from http://msystems.asm.org/ on April 27, 2018 by guest

Soltero R, Larrosa M. 2016. Gut microbiota modification: another piece in the puzzle of the benefits of physical exercise in health? Front Physiol 7:51. https://doi.org/10.3389/fphys.2016.00051.
6. Mika A, Van Treuren W, González A, Herrera JJ, Knight R, Fleshner M. 2015. Exercise is more effective at altering gut microbial composition and producing stable changes in lean mass in juvenile versus adult male F344 rats. PLoS One 10:e0125889. https://doi.org/10.1371/journal.pone.0125889.
7. Cronin O, Molloy MG, Shanahan F. 2016. Exercise, fitness, and the gut. Curr Opin Gastroenterol 32:67–73. https://doi.org/10.1097/MOG.0000000000000240.
8. Estaki M, Pither J, Baumeister P, Little JP, Gill SK, Ghosh S, Ahmadi-Vand Z, Marsden KR, Gibson DL. 2016. Cardiorespiratory fitness as a predictor of intestinal microbial diversity and distinct metagenomic functions. Microbiome 4:42. https://doi.org/10.1186/s40168-016-0189-7.
9. Bressa C, Bailén-Andrino M, Pérez-Santiago J, González-Soltero R, Pérez M, Montalvo-Lominchar MG, Maté-Muñoz JL, Domínguez R, Moreno D, Larrosa M. 2017. Differences in gut microbiota profile between women with active lifestyle and sedentary women. PLoS One 12:e0171352. https://doi.org/10.1371/journal.pone.0171352.
10. Petersen LM, Bautista EJ, Nguyen H, Hanson BM, Chen L, Lek SH, Sodergren E, Weinstock GM. 2017. Community characteristics of the gut microbiomes of competitive cyclists. Microbiome 5:98. https://doi.org/10.1186/s40168-017-0320-4.
11. Barton W, Penney NC, Cronin O, Garcia-Perez I, Molloy MG, Holmes E, Shanahan F, Cotter PD, O'Sullivan O. 2018. The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level. Gut 67:625–633. https://doi.org/10.1136/gutjnl-2016-313627.
12. Clarke SF, Murphy EF, O'Sullivan O, Lucey AJ, Humphreys M, Hogan A, Hayes P, O'Reilly M, Jeffery IB, Wood-Martin R, Kerins DM, Quigley E, Ross RP, O'Toole PW, Molloy MG, Falvey E, Shanahan F, Cotter PD. 2014. Exercise and associated dietary extremes impact on gut microbial diversity. Gut 63:1913–1920. https://doi.org/10.1136/gutjnl-2013-306541.
13. Evans CC, LePard KJ, Kwak JW, Stancukas MC, Laskowski S, Dougherty J, Moulton L, Glawe A, Wang Y, Leone V, Antonopoulos DA, Smith D, Chang EB, Ciancio MJ. 2014. Exercise prevents weight gain and alters the gut microbiota in a mouse model of high fat diet-induced obesity. PLoS One 9:e92193. https://doi.org/10.1371/journal.pone.0092193.
14. Queipo-Ortuño MI, Seoane LM, Murri M, Pardo M, Gomez-Zumaquero JM, Cardona F, Casanueva F, Tinahones FJ. 2013. Gut microbiota composition in male rat models under different nutritional status and physical activity and its association with serum leptin and ghrelin levels. PLoS One 8:e65465. https://doi.org/10.1371/journal.pone.0065465.
15. O'Sullivan O, Cronin O, Clarke SF, Murphy EF, Molloy MG, Shanahan F, Cotter PD. 2015. Exercise and the microbiota. Gut Microbes 6:131–136. https://doi.org/10.1080/19490976.2015.1011875.
16. Vandeputte D, Falony G, Vieira-Silva S, Tito RY, Joossens M, Raes J. 2016. Stool consistency is strongly associated with gut microbiota richness and composition, enterotypes and bacterial growth rates. Gut 65:57–62. https://doi.org/10.1136/gutjnl-2015-309618.
17. Song BK, Cho KO, Jo Y, Oh JW, Kim YS. 2012. Colon transit time according to physical activity level in adults. J Neurogastroenterol Motil 18:64–69. https://doi.org/10.5056/jnm.2012.18.1.64.
18. Mika A, Fleshner M. 2016. Early-life exercise may promote lasting brain and metabolic health through gut bacterial metabolites. Immunol Cell Biol 94:151–157. https://doi.org/10.1038/icb.2015.113.
19. Borg G. 1998. Borg's perceived exertion and pain scales. Human Kinetics, Champaign, IL.
20. Kim Y, Park I, Kang M. 2013. Convergent validity of the international physical activity questionnaire (IPAQ): meta-analysis. Public Health Nutr 16:440–452. https://doi.org/10.1017/S1368980012002996.
21. Morton RW, McGlory C, Phillips SM. 2015. Nutritional interventions to augment resistance training-induced skeletal muscle hypertrophy. Front Physiol 6:245. https://doi.org/10.3389/fphys.2015.00245.
22. McAllan L, Keane D, Schellekens H, Roche HM, Korpela R, Cryan JF, Nilaweera KN. 2013. Whey protein isolate counteracts the effects of a high-fat diet on energy intake and hypothalamic and adipose tissue expression of energy balance-related genes. Br J Nutr 110:2114–2126. https://doi.org/10.1017/S0007114513001396.
23. Sousa GT, Lira FS, Rosa JC, de Oliveira EP, Oyama LM, Santos RV, Pimentel GD. 2012. Dietary whey protein lessens several risk factors for metabolic diseases: a review. Lipids Health Dis 11:67. https://doi.org/10.1186/1476-511X-11-67.
24. Brimelow RE, West NP, Williams LT, Cripps AW, Cox AJ. 2017. A role for whey-derived lactoferrin and immunoglobulins in the attenuation of obesity-related inflammation and disease. Crit Rev Food Sci Nutr 57:1593–1602. https://doi.org/10.1080/10408398.2014.995264.
25. Reimer RA, Willis HJ, Tunnicliffe JM, Park H, Madsen KL, Soto-Vaca A. 2017. Inulin-type fructans and whey protein both modulate appetite but only fructans alter gut microbiota in adults with overweight/obesity: a randomized controlled trial. Mol Nutr Food Res 61. https://doi.org/10.1002/mnfr.201700484.
26. Nkamga VD, Henrissat B, Drancourt M. 2017. Archaea: essential inhabitants of the human digestive microbiota. Hum Microbiome J 3:1–8. https://doi.org/10.1016/j.humic.2016.11.005.
27. Lamoureux EV, Grandy SA, Langille MGI. 2017. Moderate exercise has limited but distinguishable effects on the mouse microbiome. MSystems 2:e00006-17. https://doi.org/10.1128/mSystems.00006-17.
28. Allen JM, Mailing LJ, Niemiro GM, Moore R, Cook MD, White BA, Holscher HD, Woods JA. 2018. Exercise alters gut microbiota composition and function in lean and obese humans. Med Sci Sports Exerc 50:747–757. https://doi.org/10.1249/MSS.0000000000001495.
29. Faith JJ, Guruge JL, Charbonneau M, Subramanian S, Seedorf H, Goodman AL, Clemente JC, Knight R, Heath AC, Leibel RL, Rosenbaum M, Gordon JI. 2013. The long-term stability of the human gut microbiota. Science 341:1237439. https://doi.org/10.1126/science.1237439.
30. Cronin O, O'Sullivan O, Barton W, Cotter PD, Molloy MG, Shanahan F. 2017. Gut microbiota: implications for sports and exercise medicine. Br J Sports Med 51:700–701. https://doi.org/10.1136/bjsports-2016-097225.
31. Karl JP, Margolis LM, Madslien EH, Murphy NE, Castellani JW, Gundersen Y, Hoke AV, Levangie MW, Kumar R, Chakraborty N, Gautam A, Hammamieh R, Martini S, Montain SJ, Pasiakos SM. 2017. Changes in intestinal microbiota composition and metabolism coincide with increased intestinal permeability in young adults under prolonged physiological stress. Am J Physiol Gastrointest Liver Physiol 312:G559–G571. https://doi.org/10.1152/ajpgi.00066.2017.
32. Koeth RA, Wang Z, Levison BS, Buffa JA, Org E, Sheehy BT, Britt EB, Fu X, Wu Y, Li L, Smith JD, DiDonato JA, Chen J, Li H, Wu GD, Lewis JD, Warrier M, Brown JM, Krauss RM, Tang WH, Bushman FD, Lusis AJ, Hazen SL. 2013. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. Nat Med 19:576–585. https://doi.org/10.1038/nm.3145.
33. Tang WH, Wang Z, Levison BS, Koeth RA, Britt EB, Fu X, Wu Y, Hazen SL. 2013. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. N Engl J Med 368:1575–1584. https://doi.org/10.1056/NEJMoa1109400.
34. Wang Z, Klipfell E, Bennett BJ, Koeth R, Levison BS, Dugar B, Feldstein AE, Britt EB, Fu X, Chung YM, Wu Y, Schauer P, Smith JD, Allayee H, Tang WH, DiDonato JA, Lusis AJ, Hazen SL. 2011. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. Nature 472:57–63. https://doi.org/10.1038/nature09922.
35. Bennett BJ, de Aguiar Vallim TQ, Wang Z, Shih DM, Meng Y, Gregory J, Allayee H, Lee R, Graham M, Crooke R, Edwards PA, Hazen SL, Lusis AJ. 2013. Trimethylamine-N-oxide, a metabolite associated with atherosclerosis, exhibits complex genetic and dietary regulation. Cell Metab 17:49–60. https://doi.org/10.1016/j.cmet.2012.12.011.
36. Romano KA, Vivas EI, Amador-Noguez D, Rey FE. 2015. Intestinal microbiota composition modulates choline bioavailability from diet and accumulation of the proatherogenic metabolite trimethylamine-N-oxide. MBio 6:e02481. https://doi.org/10.1128/mBio.02481-14.
37. Jang HJ, Kim DM, Kim KB, Park JW, Choi JY, Oh JH, Song KD, Kim S, Cho BW. 2017. Analysis of metabolomic patterns in thoroughbreds before and after exercise. Asian-Australas J Anim Sci 30:1633–1642. https://doi.org/10.5713/ajas.17.0167.
38. Mithieux G. 24 October 2017. Gut microbiota and host metabolism: what relationship? Neuroendocrinology https://doi.org/10.1159/000484526.
39. De Vadder F, Kovatcheva-Datchary P, Zitoun C, Duchampt A, Bäckhed F, Mithieux G. 2016. Microbiota-produced succinate improves glucose homeostasis via intestinal gluconeogenesis. Cell Metab 24:151–157. https://doi.org/10.1016/j.cmet.2016.06.013.
40. Marungruang N, Tovar J, Björck I, Hållenius FF. 2017. Improvement in cardiometabolic risk markers following a multifunctional diet is associated with gut microbial taxa in healthy overweight and obese subjects. Eur J Nutr. https://doi.org/10.1007/s00394-017-1563-3.
41. Scher JU, Sczesnak A, Longman RS, Segata N, Ubeda C, Bielski C, Rostron T, Cerundolo V, Pamer EG, Abramson SB, Huttenhower C, Littman DR. 2013. Expansion of intestinal Prevotella copri correlates with enhanced

Downloaded from http://msystems.asm.org/ on April 27, 2018 by guest

susceptibility to arthritis. Elife 2:e01202. https://doi.org/10.7554/eLife.01202.
42. Pianta A, Arvikar S, Strle K, Drouin EE, Wang Q, Costello CE, Steere AC. 2017. Evidence of the immune relevance of Prevotella copri, a gut microbe, in patients with rheumatoid arthritis. Arthritis Rheumatol 69:964–975. https://doi.org/10.1002/art.40003.
43. Wen C, Zheng Z, Shao T, Liu L, Xie Z, Le Chatelier E, He Z, Zhong W, Fan Y, Zhang L, Li H, Wu C, Hu C, Xu Q, Zhou J, Cai S, Wang D, Huang Y, Breban M, Qin N, Ehrlich SD. 2017. Quantitative metagenomics reveals unique gut microbiome biomarkers in ankylosing spondylitis. Genome Biol 18:142. https://doi.org/10.1186/s13059-017-1271-6.
44. Leite AZ, Rodrigues NC, Gonzaga MI, Paiolo JCC, de Souza CA, Stefanutto NAV, Omori WP, Pinheiro DG, Brisotti JL, Matheucci Junior E, Mariano VS, de Oliveira GLV. 2017. Detection of increased plasma interleukin-6 levels and prevalence of Prevotella copri and Bacteroides vulgatus in the feces of type 2 diabetes patients. Front Immunol 8:1107. https://doi.org/10.3389/fimmu.2017.01107.
45. Sempos CT. 1992. Invited commentary: some limitations of semiquantitative food frequency questionnaires. Am J Epidemiol 135:1127–1132. https://doi.org/10.1093/oxfordjournals.aje.a116212.
46. Craig CL, Marshall AL, Sjöström M, Bauman AE, Booth ML, Ainsworth BE, Pratt M, Ekelund U, Yngve A, Sallis JF, Oja P. 2003. International physical activity questionnaire: 12-country reliability and validity. Med Sci Sports Exerc 35:1381–1395. https://doi.org/10.1249/01.MSS.0000078924.61453.FB.
47. Pescatello LS. 2014. ACSM's guidelines for exercise testing and prescription, 9th ed. Wolters Kluwer/Lippincott Williams & Wilkins (Health), Philadelphia, PA.
48. Brzycki M. 1998, A practical approach to strength training, 3rd ed. McGraw-Hill, New York, NY.
49. Ainsworth BE, Haskell WL, Herrmann SD, Meckes N, Bassett DR, Jr, Tudor-Locke C, Greer JL, Vezina J, Whitt-Glover MC, Leon AS. 2011. 2011 Compendium of Physical Activities: a second update of codes and MET values. Med Sci Sports Exerc 43:1575–1581. https://doi.org/10.1249/MSS.0b013e31821ece12.
50. Kline GM, Porcari JP, Hintermeister R, Freedson PS, Ward A, McCarron RF, Ross J, Rippe JM. 1987. Estimation of VO2max from a one-mile track walk, gender, age, and body weight. Med Sci Sports Exerc 19:253–259.
51. Ekelund U, Ward HA, Norat T, Luan J, May AM, Weiderpass E, Sharp SJ, Overvad K, Østergaard JN, Tjønneland A, Johnsen NF, Mesrine S, Fournier A, Fagherazzi G, Trichopoulou A, Lagiou P, Trichopoulos D, Li K, Kaaks R, Ferrari P, Licaj I, Jenab M, Bergmann M, Boeing H, Palli D, Sieri S, Panico S, Tumino R, Vineis P, Peeters PH, Monnikhof E, Bueno-de-Mesquita HB, Quirós JR, Agudo A, Sánchez MJ, Huerta JM, Ardanaz E, Arriola L, Hedblad B, Wirfält E, Sund M, Johansson M, Key TJ, Travis RC, Khaw KT, Brage S, Wareham NJ, Riboli E. 2015. Physical activity and all-cause mortality across levels of overall and abdominal adiposity in European men and women: the European Prospective Investigation into Cancer and Nutrition Study (EPIC). Am J Clin Nutr 101:613–621. https://doi.org/10.3945/ajcn.114.100065.
52. Willett WC, Sampson L, Stampfer MJ, Rosner B, Bain C, Witschi J, Hennekens CH, Speizer FE. 1985. Reproducibility and validity of a semiquantitative food frequency questionnaire. Am J Epidemiol 122:51–65. https://doi.org/10.1093/oxfordjournals.aje.a114086.
53. Dray S, Dufour A-B. 2007. The ade4 package: implementing the duality diagram for ecologists. J Stat Softw 22:1–20.
54. Mirsepasi H, Persson S, Struve C, Andersen LO, Petersen AM, Krogfelt KA. 2014. Microbial diversity in fecal samples depends on DNA extraction method: easyMAG DNA extraction compared to QIAamp DNA stool mini kit extraction. BMC Res Notes 7:50. https://doi.org/10.1186/1756-0500-7-50.
55. Abubucker S, Segata N, Goll J, Schubert AM, Izard J, Cantarel BL, Rodriguez-Mueller B, Zucker J, Thiagarajan M, Henrissat B, White O, Kelley ST, Methé B, Schloss PD, Gevers D, Mitreva M, Huttenhower C. 2012. Metabolic reconstruction for metagenomic data and its application to the human microbiome. PLoS Comput Biol 8:e1002358. https://doi.org/10.1371/journal.pcbi.1002358.
56. Caspi R, Foerster H, Fulcher CA, Kaipa P, Krummenacker M, Latendresse M, Paley S, Rhee SY, Shearer AG, Tissier C, Walk TC, Zhang P, Karp PD. 2008. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res 36:D623–D631. https://doi.org/10.1093/nar/gkm900.
57. Wood DE, Salzberg SL. 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol 15:R46. https://doi.org/10.1186/gb-2014-15-3-r46.
58. Dona AC, Jiménez B, Schäfer H, Humpfer E, Spraul M, Lewis MR, Pearce JT, Holmes E, Lindon JC, Nicholson JK. 2014. Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping. Anal Chem 86:9887–9894. https://doi.org/10.1021/ac5025039.
59. Want EJ, Wilson ID, Gika H, Theodoridis G, Plumb RS, Shockcor J, Holmes E, Nicholson JK. 2010. Global metabolic profiling procedures for urine using UPLC-MS. Nat Protoc 5:1005–1018. https://doi.org/10.1038/nprot.2010.50.
60. Sarafian MH, Lewis MR, Pechlivanis A, Ralphs S, McPhail MJ, Patel VC, Dumas ME, Holmes E, Nicholson JK. 2015. Bile acid profiling and quantification in biofluids using ultra-performance liquid chromatography tandem mass spectrometry. Anal Chem 87:9662–9670. https://doi.org/10.1021/acs.analchem.5b01556.
61. Tautenhahn R, Böttcher C, Neumann S. 2008. Highly sensitive feature detection for high resolution LC/MS. BMC Bioinformatics 9:504. https://doi.org/10.1186/1471-2105-9-504.
62. Veselkov KA, Vingara LK, Masson P, Robinette SL, Want E, Li JV, Barton RH, Boursier-Neyret C, Walther B, Ebbels TM, Pelczer I, Holmes E, Lindon JC, Nicholson JK. 2011. Optimized preprocessing of ultra-performance liquid chromatography/mass spectrometry urinary metabolic profiles for improved information recovery. Anal Chem 83:5864–5872. https://doi.org/10.1021/ac201065j.
63. Posma JM, Garcia-Perez I, De Iorio M, Lindon JC, Elliott P, Holmes E, Ebbels TM, Nicholson JK. 2012. Subset optimization by reference matching (STORM): an optimized statistical approach for recovery of metabolic biomarker structural information from 1H NMR spectra of biofluids. Anal Chem 84:10694–10701. https://doi.org/10.1021/ac302360v.
64. Cloarec O, Dumas ME, Craig A, Barton RH, Trygg J, Hudson J, Blancher C, Gauguier D, Lindon JC, Holmes E, Nicholson J. 2005. Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H NMR data sets. Anal Chem 77:1282–1289. https://doi.org/10.1021/ac048630x.
65. Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol 57:289–300. http://www.jstor.org/stable/2346101.
66. Oksanen J, Blanchet FG, Friendly M, Kindt R, Legendre P, McGlinn D, Minchin PR, O'Hara R, Simpson GL, Solymos P, Henry M, Stevens H, Szoecs E, Wagner H. 2017. Vegan: community ecology package. R package version 2.4-3. https://cran.r-project.org/web/packages/vegan/index.html. Accessed 9 May 2017.
67. Mandal S, Van Treuren W, White RA, Eggesbø M, Knight R, Peddada SD. 2015. Analysis of composition of microbiomes: a novel method for studying microbial composition. Microb Ecol Health Dis 26:27663. https://doi.org/10.3402/mehd.v26.27663.
68. Cacciatore S, Tenori L, Luchinat C, Bennett PR, MacIntyre DA. 2017. KODAMA: an R package for knowledge discovery and data mining. Bioinformatics 33:621–623. https://doi.org/10.1093/bioinformatics/btw705.
69. Garcia-Perez I, Posma JM, Gibson R, Chambers ES, Hansen TH, Vestergaard H, Hansen T, Beckmann M, Pedersen O, Elliott P, Stamler J, Nicholson JK, Draper J, Mathers JC, Holmes E, Frost G. 2017. Objective assessment of dietary patterns by use of metabolic phenotyping: a randomised, controlled, crossover trial. Lancet Diabetes Endocrinol 5:184–195. https://doi.org/10.1016/S2213-8587(16)30419-3.
70. Lillefosse HH, Clausen MR, Yde CC, Ditlev DB, Zhang X, Du ZY, Bertram HC, Madsen L, Kristiansen K, Liaset B. 2014. Urinary loss of tricarboxylic acid cycle intermediates as revealed by metabolomics studies: an underlying mechanism to reduce lipid accretion by whey protein ingestion? J Proteome Res 13(5):2560–2570. https://doi.org/10.1021/pr500039t.
71. Piccolo BD, Comerford KB, Karakas SE, Knotts TA, Fiehn O, Adams SH. 2015. Whey protein supplementation does not alter plasma branched-chained amino acid profiles but results in unique metabolomics patterns in obese women enrolled in an 8-week weight loss trial. J Nutr 145(4):691–700. https://doi.org/10.3945/jn.114.203943.

Downloaded from http://msystems.asm.org/ on April 27, 2018 by guest

The Microbiome of Professional Athletes Differs from that of More Sedentary Subjects in Composition and Particularly at the Functional Metabolic Level

Downloaded from http://gut.bmj.com/ on March 30, 2017 - Published by group.bmj.com
Gut Online First, published on March 30, 2017 as 10.1136/gutjnl-2016-313627

ORIGINAL ARTICLE

The microbiome of professional athletes differs from that of more sedentary subjects in composition and particularly at the functional metabolic level

Wiley Barton,[1,2,3] Nicholas C Penney,[4,5] Owen Cronin,[1,3] Isabel Garcia-Perez,[4] Michael G Molloy,[1,3] Elaine Holmes,[4] Fergus Shanahan,[1,3] Paul D Cotter,[1,2] Orla O'Sullivan[1,2]

▶ Additional material is published online only. To view please visit the journal online (http://dx.doi.org/10.1136/gutjnl-2016-313627).

[1]Alimentary Pharmabiotic Centre Microbiome Institute, University College Cork, National University of Ireland, Cork, Ireland
[2]Teagasc Food Research Centre, Cork, Ireland
[3]Department of Medicine, University College Cork, National University of Ireland, Cork, Ireland
[4]Section of Biomolecular Medicine, Division of Computational Systems Medicine, Department of Surgery and Cancer, Imperial College London, London, UK
[5]Division of Surgery, Department of Surgery and Cancer, Imperial College London, London, UK

Correspondence to
Professor Fergus Shanahan, APC Microbiome Institute, University College Cork, National University of Ireland, Cork, T12 DC4A Ireland; f.shanahan@ucc.ie

Received 20 December 2016
Revised 31 January 2017
Accepted 6 March 2017

ABSTRACT

Objective It is evident that the gut microbiota and factors that influence its composition and activity effect human metabolic, immunological and developmental processes. We previously reported that extreme physical activity with associated dietary adaptations, such as that pursued by professional athletes, is associated with changes in faecal microbial diversity and composition relative to that of individuals with a more sedentary lifestyle. Here we address the impact of these factors on the functionality/metabolic activity of the microbiota which reveals even greater separation between exercise and a more sedentary state.

Design Metabolic phenotyping and functional metagenomic analysis of the gut microbiome of professional international rugby union players (n=40) and controls (n=46) was carried out and results were correlated with lifestyle parameters and clinical measurements (eg, dietary habit and serum creatine kinase, respectively).

Results Athletes had relative increases in pathways (eg, amino acid and antibiotic biosynthesis and carbohydrate metabolism) and faecal metabolites (eg, microbial produced short-chain fatty acids (SCFAs) acetate, propionate and butyrate) associated with enhanced muscle turnover (fitness) and overall health when compared with control groups.

Conclusions Differences in faecal microbiota between athletes and sedentary controls show even greater separation at the metagenomic and metabolomic than at compositional levels and provide added insight into the diet–exercise–gut microbiota paradigm.

Significance of this study

What is already known on this subject?
- Taxonomic and functional compositions of the gut microbiome are emerging as biomarkers of human health and disease.
- Physical exercise and associated dietary adaptation are linked with changes in the composition of the gut microbiome.
- Metabolites such as short-chain fatty acids (SCFAs) have an impact on a range of health parameters including immunity, colonic epithelial cell integrity and brain function.

What are the new findings?
- Our original observation of differences in gut microbiota composition in elite athletes is confirmed and the separation between athletes and those with a more sedentary lifestyle is even more evident at the functional or metabolic level. Microbial-derived SCFAs are enhanced within the athletes.

How might it impact on clinical practice in the foreseeable future?
- The findings provide new evidence supporting the link between exercise and metabolic health. The findings provide a platform for the rational design of diets for those engaged in vigorous exercise. The identification of specific alterations in the metabolic profile of subjects engaged in high levels of exercise provides insight necessary for future efforts towards targeted manipulation of the microbiome.

INTRODUCTION

Regular exercise challenges systemic homeostasis resulting in a breadth of multiorgan molecular and physiological responses, including many that centre on immunity, metabolism and the microbiome–gut–brain axis.[1–5] Exercise exhibits systemic and end-organ anti-inflammatory effects as well as contributing to more efficient carbohydrate metabolism, in addition to trophic effects at the level of the central nervous system.[6 7] In fact, increasing physical activity offers an effective treatment and preventative strategy for many chronic conditions in which the gut microbiome has been implicated.[8–10] Conversely, a sedentary lifestyle is a major contributing factor to morbidity in developed Western society and is associated with heightened risk of numerous *diseases of affluence*, such as obesity, diabetes, asthma and cardiovascular disease (CVD).[11–14] Recent evidence supports an influential role for the gut microbiome in these diseases.[15–23]

The concept that regular exercise and sustained levels of increased physical activity foster or assist the maintenance of a preferential intestinal microbiome has recently gained momentum and

To cite: Barton W, Penney NC, Cronin O, *et al*. *Gut* Published Online First: [*please include* Day Month Year] doi:10.1136/gutjnl-2016-313627

Copyright Article author (or their employer) 2017. Produced by BMJ Publishing Group Ltd (& BSG) under licence.

Downloaded from http://gut.bmj.com/ on March 30, 2017 - Published by group.bmj.com

interest.[24–29] Previously, using 16S rRNA amplicon sequencing, we demonstrated taxonomic differences in gut microbiota between an elite athlete cohort of international-level rugby players and a group of age-matched high (>28 kg/m^2) and low (<25 kg/m^2) body mass index (BMI) controls.[26] This analysis illustrated a significantly greater intestinal microbial diversity among the athletes compared with both control groups. This taxonomic diversity significantly correlated with exercise and dietary protein consumption. However, the possibility existed that these differences did not equate to differences at a functional level. Here, we re-examine the microbiome in these participants by whole metagenome shotgun sequencing to provide deeper insight into taxonomic composition and functional potential and by complementary metabolic phenotyping analyses of host-derived and microbial-derived (urine and faecal, respectively) metabolic profiles. This analysis shows that the differences in the gut microbiota between athletes and controls is even more pronounced at the functional metabolic level than at the compositional level as previously reported and provides further rationale for prospective controlled studies to unravel the relationship between diet, exercise and the gut microbiome.

RESULTS

The study groups comprise professional male athletes (n=40) and healthy controls (n=46).[26] To better represent the variability of BMI in the athletes, controls were classified as either low BMI (n=22, BMI≤25.2) or high BMI (n=24, BMI≥26.5). Participants made no report of GI distress or alterations of GI transit time throughout the course of the initial study.

Functional structure of the enteric microbiome correlates with athletic state

Functional metagenomic analysis of faecal samples allowed for the prediction of the operational potential of each individual's microbiota. In total, 19 300 taxonomically linked metabolic pathways were identified in at least one individual. Comparison of phylogenetic constructions derived from the 16S rRNA amplicon data of our previous study and the functional data of this present report reveals a greater level of identification at higher levels of taxonomy (eg, phylum) for 16S sequences,[26] while the metagenomic data had greater fidelity and superior resolution of lower levels of taxonomy (eg, species) (figure 1). Consistent with previous results, the microbiota of the athletes were significantly more diverse than that of both the low-BMI and high-BMI control groups at the functional level (figure 2A). Furthermore, our previous findings of an enrichment of *Akkermansia* in athletes was corroborated by the presence of significantly higher proportions of metabolic pathways associated with this genus in athletes when compared with high-BMI controls ($p<0.001$). Correlation analysis revealed that, of the total 19 300 pathways, 98 were significantly altered between the three cohorts ($p<0.05$) (see online supplementary table S1). Subsequently, large-scale functional dissimilarity between athletes and controls was determined and distinct patterns of pathway composition between groups were revealed (see online supplementary figure S1A). This functional distinction remained true whether applied to total pathway data or to the statistically significant subset of pathways (see online supplementary figure S1B). Correlation of pathways present in at least one member from both cohorts further exemplified the uniformity of the athletes and the division between the athletes and control groups (see online supplementary figure S1C). Separation according to group membership was further illustrated through principal coordinate analysis (PCoA), with statistical support of the significant separation between the athletes and both control groups ($p<0.05$) (figure 2B). This was also the case for the statistically significant subset of pathways (see online supplementary figure S1D). Principal component analysis (PCA) supplemented with a correspondence analysis and k-nearest neighbour semisupervised learning approach cast further light (ie, visualisation of robustly defined class associations of specific individuals within the groups) on the clustering of participants within and between cohorts (see online supplementary figure S1E).

Pathways exhibiting statistically significant variation between the athletes and both control groups were organised according to MetaCyc metabolic pathway hierarchy classification

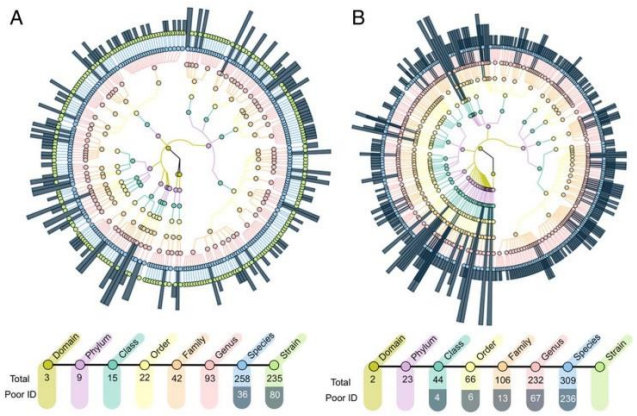


Figure 1 Comparison of phylogenetic constructions from metagenomic and 16S rRNA gene sequencing sourced from all participants. Phylogenetic trees derived from (A) metagenomic sequencing and (B) 16S rRNA amplicon sequencing. Taxonomic levels are assigned from centre out with kingdom-level assignment in centre and strain-level assignment in outer most ring. Dark blue radial highlights correspond to poorly identified taxonomies (ie, 'unknown' and 'unassigned' database entries). Number of assignments at each level of phylogeny is displayed below the respective graph. Taxonomic trees derived from the two sequencing approaches illustrate an advantage of metagenomic sequencing in the number of predictions of lower taxonomic levels and the frequency of full identification of taxa, while 16S rRNA sequencing grants greater insight of high-level phylogenies within the population.

Downloaded from http://gut.bmj.com/ on March 30, 2017 - Published by group.bmj.com

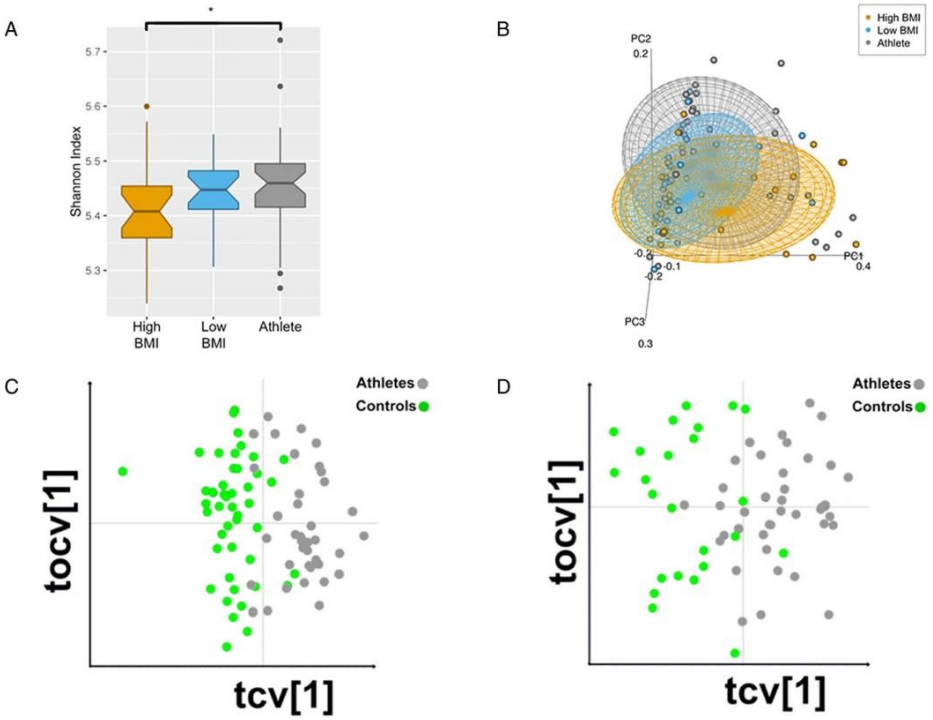


Figure 2 Group-wise comparison of microbial metagenomic and metabolomic profiles. (A) Shannon index of diversity for metabolic pathways from all three groups. Pathway diversity is increased in the athlete group when compared with low-body mass index (BMI) and high-BMI controls. Diversity measures are statistically significant between low-BMI and athlete groups ($p<0.049$), with statistical significance between all groups (Kruskal-Wallis $p<0.05$). (B) Principle coordinate analysis of Bray-Curtis compiled distance matrix of all microbial metabolic pathway relative abundances. Groups show significant variation from one another (Adonis PERMANOVA $p<0.05$). Cross-validated orthogonal partial least squares regression discriminant analysis (OPLS-DA) of full nuclear magnetic resonance (^{1}H-NMR) spectra from urine ($R^2Y=0.86$, $Q^2Y=0.60$) (C) and faecal water ($R^2Y=0.86$, $Q^2Y=0.52$) (D) samples. OPLS-DA displays robust separation between athletes and controls. Models comprise 1 predictive (tcv[1]) and 1 orthogonal (tocv[1]) principal component.

(34 metabolic categories), highlighting a number of differences (figure 3A and online supplementary table S2). Distinct clustering patterns were observed within each cohort, with the high-BMI control group having the lowest average abundance scores across 31 metabolic pathway categories (the exceptions being vitamin biosynthesis (VB), lipid biosynthesis (LB) and amino acid biosynthesis (AAB) categories). The athlete group had the highest mean abundance across 29 of the 34 metabolic categories (eg, carbohydrate biosynthesis, cofactor biosynthesis and energy metabolism) (see online supplementary table S2).

Numerous statistically significant ($p<0.05$) associations were identified between pathway abundances and serum creatine kinase (CK)—an enzymatic marker of muscle activity (IU/L), total bilirubin (IU/L) and dietary macronutrient intake of protein (g/day), fibre (g/day), carbohydrates (g/day), sugars (g/day), starch (g/day), fat (g/day) and total energy (KJ/day) (figure 3B). Each group was represented by distinct association profiles of the correlation between clinical measurements and metagenomic pathways. Dietary factors, sugars and other carbohydrates, as well as energy intake, provide the majority of the correlations for the control groups, whereas the athlete group was predominantly correlated with CK, total bilirubin and total energy intake. Of the total number of metabolic pathways with associations to the clinical data from all three groups (10 760; data not shown), relevant pathways related to the production of secondary metabolites, cofactors and short-chain fatty acids (SCFAs) were identified (eg, biotin biosynthesis and pyruvate fermentation to butanoate).

Distinct differences between host and microbial metabolites in athletes and controls

A combination of multiplatform metabolic phenotyping and multivariate analysis based on orthogonal partial least squares discriminant analysis (OPLS-DA) was used to compare urine and faecal samples from athletes and controls. The cross-validated (CV) OPLS-DA models show strong differences between athletes and controls in urine samples by proton nuclear magnetic resonance (^{1}H-NMR) analysis ($R^2Y=0.86$, $Q^2Y=0.60$, figure 2C), hydrophilic interaction ultra-performance liquid chromatography mass spectroscopy (HILIC UPLC-MS) positive mode analysis ($R^2Y=0.85$, $Q^2Y=0.74$, online supplementary figure S2A) and reversed-phase UPLC-MS (RP UPLC-MS) in both positive

Downloaded from http://gut.bmj.com/ on March 30, 2017 - Published by group.bmj.com

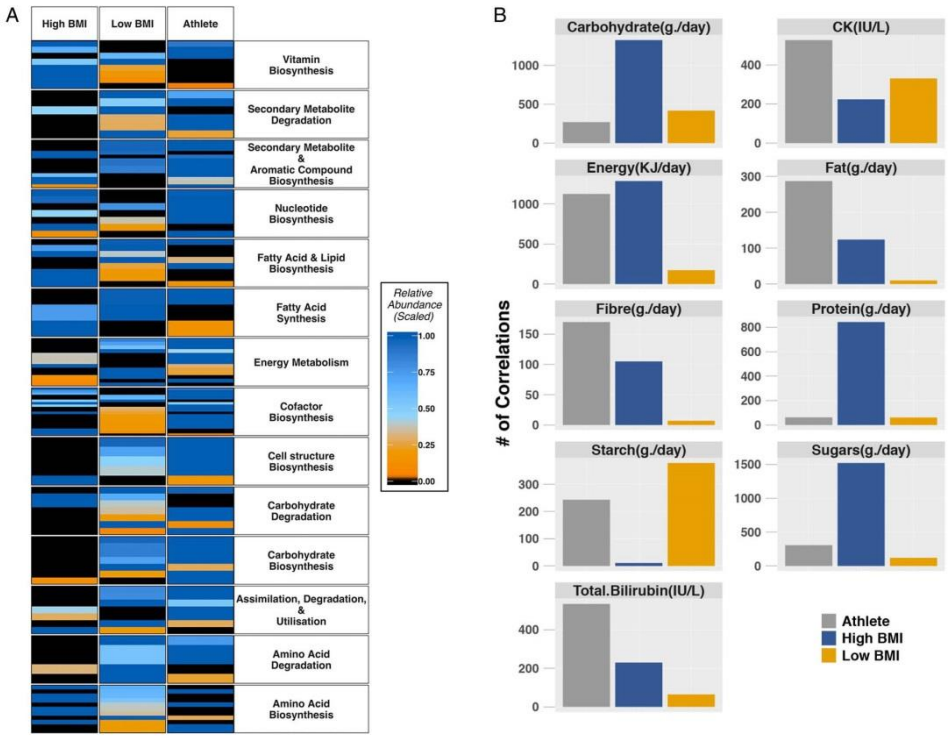


Figure 3 Group variation of microbial metabolic function and associations between pathways and clinical and dietary variables. (A) Mean relative abundance values of statistically significant (Kruskal-Wallis p<0.05) metabolic pathways binned according to categories of metabolic function. (B) Number of metabolic pathways significantly (Benjamini-Hochberg corrected p<0.05) correlated with dietary constituents and blood serum metabolites. BMI, body mass index.

and negative mode analysis ($R^2Y=0.83$, $Q^2Y=0.73$ and $R^2Y=0.83$, $Q^2Y=0.67$, online supplementary figure S2B,C respectively). Likewise, the CV-OPLS-DA models comparing faecal samples, although weaker than the urine models, reveal significant differences between athletes and controls by ^{1}H-NMR analysis ($R^2Y=0.86$, $Q^2Y=0.52$, figure 2D) and HILIC UPLC-MS positive mode analysis ($R^2Y=0.65$, $Q^2Y=0.34$, online supplementary figure S2D).

The loadings of the pairwise OPLS-DA models were used to identify metabolites discriminating between the two classes. Athletes' ^{1}H-NMR metabolic phenotypes were characterised by higher levels of trimethylamine-*N*-oxide (TMAO), L-carnitine, dimethylglycine, O-acetyl carnitine, proline betaine, creatine, acetoacetate, 3-hydroxy-isovaleric acid, acetone, *N*-methylnicotinate, *N*-methylnicotinamide, phenylacetylglutamine (PAG) and 3-methylhistidine in urine samples and higher levels of propionate, acetate, butyrate, trimethylamine (TMA), lysine and methylamine in faecal samples, relative to controls. Athletes were further characterised by lower levels of glycerate, allantoin and succinate and lower levels of glycine and tyrosine relative to controls in urine and faecal samples, respectively (see online supplementary table S3).

While numerous metabolites discriminated significantly between athletes and controls with RP UPLC-MS positive (490) and negative (434) modes for urine, as well as with HILIC UPLC-MS positive mode for urine (196) and faecal water (3), key metabolites were structurally identified using the strategy described below. UPLC-MS analyses revealed higher urinary excretion of *N*-formylanthranilic acid, hydantoin-5-propionic acid, 3-carboxy-4-methyl-5-propyl-2-furanpropionic acid (CMPF), CMPF glucuronide, trimetaphosphoric acid, acetylcarnitine (C2), propionylcarnitine (C3), isobutyrylcarnitine (C4), 2-methylbutyroylcarnitine (C5), hexanoylcarnitine (C6), C9:1-carnitine, L-valine, nicotinuric acid, 4-pyridoxic acid and creatine in athletes relative to controls. Levels of glutamine, 7-methylxanthine, imidazoleacetic acid, isoquinoline/quinolone were lower in athletes' urinary samples relative to controls. In addition, 16 unknown glucuronides were lower in the athlete samples (see online supplementary table S4).

SCFA levels in faeces measured by targeted gas chromatography–mass spectrometry (GC-MS) showed significantly higher levels of acetate (p<0.001), propionate (p<0.001), butyrate (p<0.001) and valerate (p=0.011) in athletes relative to controls. Isobutyrate and isovalerate did not differ significantly between the groups (figure 4B and online supplementary table S5). Furthermore, concentrations of propionate strongly correlated to protein intake, while butyrate was shown to have a strong association with intake of dietary fibre (see online supplementary table S6).

Downloaded from http://gut.bmj.com/ on March 30, 2017 - Published by group.bmj.com

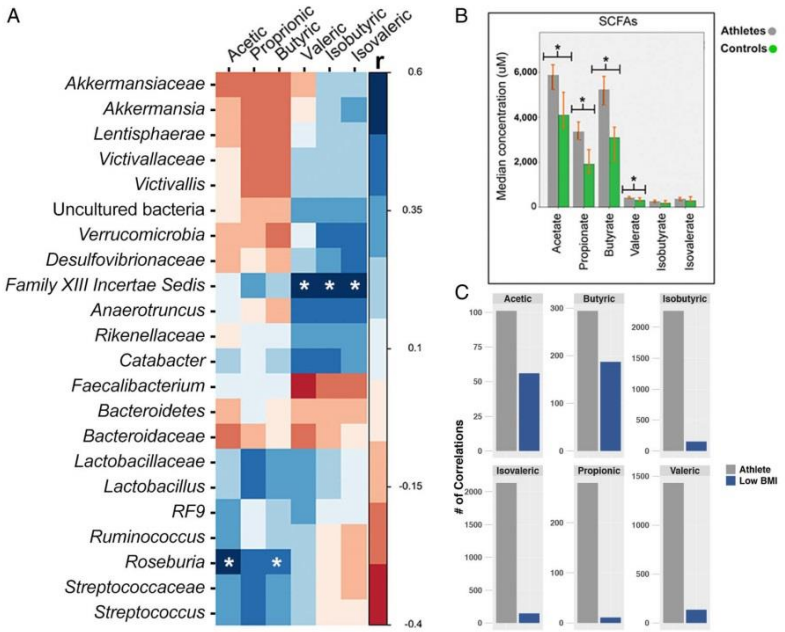


Figure 4 Athletes display a profile of short-chain fatty acids (SCFAs) that alters from that of the controls. (A) Heat map of bacterial taxa (family, genus and species level) that correlate with faecal short-chain fatty acid levels using Spearman's correlation. Cool colours represent positive correlations; hot colours represent negative correlations (r). All taxa shown had a correlation p value<0.01. Those marked with * represent correlations with a false discovery rate <0.01 after Benjamini-Hochberg multiple testing corrections. (B) Median concentrations of GC-MS-derived faecal SCFA. Quantitative analysis of SCFAs in faecal samples shows significant increase in measured concentrations of acetate, propionate, butyrate and valerate in athletes. Error bars represent 95% CIs. (C) Quantification of statistically relevant correlations of metabolic pathways to GC-MS-derived faecal SCFA concentrations (μM). BMI, body mass index; GC-MS, gas chromatography–mass spectrometry.

Correlating metabonomic and metagenomic results

Correlation analysis between targeted measurements of SCFAs and taxonomic data from 16S rRNA sequencing revealed a number of correlations that remained significant following correction; *Roseburia* was positively correlated with acetate ($p=0.004$) and butyrate ($p=0.018$) while *Family XIII Incertae Sedis* was positively correlated with isobutyrate ($p<0.001$), isovaleric acid ($p<0.001$) and valeric acid ($p=0.008$) (figure 4A and online supplementary table S7).

SCFAs were also correlated with pathway relative abundances, with all SCFAs associating with considerably more pathways in the athletes versus the controls (figure 4C). Multiple statistically significant (7948) ($p<0.05$) correlations between the metabolic pathways and SCFAs were identified (see online supplementary table S8). Two distinct blocks of proportionately discriminant correlations were observed with isobutyric and isovaleric acids, which were more abundant in the athletes, while acetic and butyric acids were proportionately more abundant in controls. Correlations of the SCFA concentrations to pathways related to fermentation, biosynthesis or modification of fatty acids were identified among the numerous other associations (see online supplementary table S8 for complete list). Additional correlations of metabolic pathways against well-identified metabolites detected from both faecal water (figure 5A, C) and urine (figure 5B, D) presented numerous significant associations (6186 and 13 412, respectively; data not shown) ($p<0.05$). It was also observed that 16 genera correlated with 12 metabolites (see online supplementary table S9).

DISCUSSION

The results confirm enhancement of microbial diversity in athletes compared with controls. Supporting previous insights into the beneficial influence of physical exercise and associated diet on the compositional structure of the gut microbiota,[25 26 30] this study has extended the paradigm to include links between physical fitness and the functional potential of the gut microbiota and its metabolites. It must be conceded that some athletes, although fit, may not necessarily be more healthy.[31]

Athletes have an increased abundance of pathways that—given an equivalent amount of expression activity—could be exploited by the host for potential health benefit, including biosynthesis of organic cofactors and antibiotics, as well as carbohydrate degradation and secondary metabolite metabolism.[32] Furthermore, athletes have an enriched profile of SCFAs, previously associated with numerous health benefits and a lean phenotype.[33–35] While interpretation of SCFA data can be difficult as levels represent a combination of SCFA production and host-absorption rates, it is notable that, as previously presented, the athletes' diet maintained significantly higher quantities of fibre intake.[26] This along with an increased number of detected SCFA pathways in the athletes would be conducive to an enhanced rate of SCFA production[36]

Downloaded from http://gut.bmj.com/ on March 30, 2017 - Published by group.bmj.com

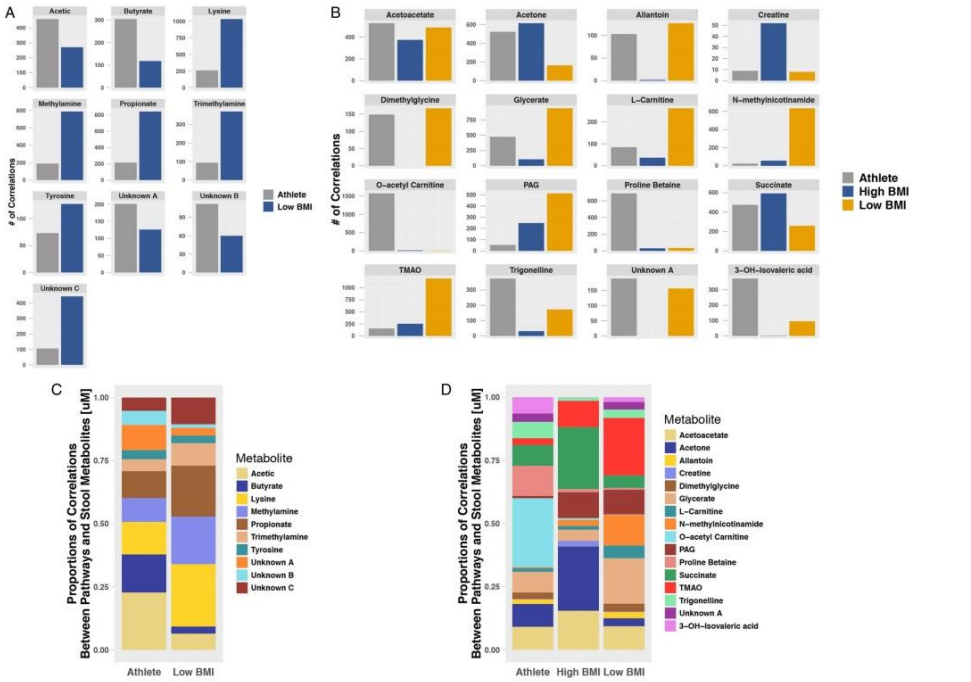


Figure 5 Distinctive association profiles of metabolic pathways to metabolites in athletes and controls. (A) Significant correlations of faecal water-derived metabolites and metabolic pathways, represented by number of correlations for each metabolite. (B) Urine metabolites significantly correlated to pathways and displayed as number of correlations. (C) Significant correlations shown in (A) displayed as proportions of total associations. (D) Correlations presented in (B) given as proportions of total associations. BMI, body mass index; PAG, phenylacetylglutamine; TMAO, trimethylamine-*N*-oxide.

Downloaded from http://gut.bmj.com/ on March 30, 2017 - Published by group.bmj.com

It was noted that athletes excreted proportionately higher levels of the metabolite TMAO, an end product metabolite of dietary protein degradation. Elevated TMAO has been observed in patients with cardiovascular disease and atherosclerosis, highlighting a potential downside to increased protein intake.[15–17 22 37] However, TMAO is also found in high levels in the urine of Japanese populations,[38] who do not have high risk for CVD. Similarly to these populations, the athletes' diet contained a significantly greater proportion of fish. Our current understanding of the implications of this result remains limited and requires elaboration in future studies. Furthermore, pathway abundance in a metagenome merely reflects functional potential and not necessarily increased expression in situ.

Variance of metagenomic composition between athletes and controls was exemplified with unique pathway–pathway correlations between the two groups. Analysis of categorically arranged pathway abundances within the separate cohorts provided additional insight into the previously described dichotomy between the microbiota of athletes and high-BMI controls. The two groups displayed distinct structures of functional capacity, separately oriented to operate under the different physiological milieu of the two groups. Notably, from a functional perspective, the microbiota of the low-BMI group was more similar to the athletes. The low-BMI controls were generally engaged in a modestly active lifestyle, reflected by their leanness and increased levels of CK. It is speculative but not implausible that moderate improvements in physical activity for overweight and obese individuals may confer the beneficial metabolic functions observed within the athlete microbiome.

Dietary contributions to the functional composition of the enteric microbial system are also evident in our study. The relative abundances of pathways related to fundamental metabolic function—AAB, VB and LB—were higher on average within the high-BMI control group when compared with the athlete group. The mechanisms behind these differences are unclear and might reflect chronic adaptation of the athlete gut microbiome; possibly due to a reduced reliance on the corresponding biosynthetic capacities of their gut microbiota. On the contrary, the athlete microbiome presents a functional capacity that is primed for tissue repair and to harness energy from the diet with increased capacity for carbohydrate, cell structure and nucleotide biosynthesis, reflecting the significant energy demands and high cell-turnover evident in elite sport.

Remarkably, our examination of pathway correlation to dietary macronutrients and plasma CK, as a biomarker of exercise,[39] is suggestive of an impact of physical activity on the use of dietary nutrients by the microbiota of the gut. Comparing athletes to both high-BMI and low-BMI controls, a greater number of pathways correlating to specific macronutrients with the controls suggests a shift in the dynamics of these varied metabolic functions. The impact of the athletes' increased protein intake compared with both control groups was evident in the metabolomic phenotyping results. By-products of dietary protein metabolism (mostly by microbes) including TMAO, carnitines, TMA, 3-CMPF and 3-hydroxy-isovaleric acid are all elevated in the athlete cohort. Of particular interest is 3-hydroxy-isovaleric acid (potentially from egg consumption), which has been demonstrated to have efficacy for inhibiting muscle wasting when used in conjunction with physical exercise.[40 41] The compound is also commonly used as a supplement by athletes to increase exercise-induced gains in muscle size, muscle strength and lean body mass, reduce exercise-induced muscle damage and speed recovery from high-intensity exercise.[41] Numerous metabolites associated with muscle turnover, creatine, 3-methylhistidine and L-valine, and host metabolism, carnitine, are elevated in the athlete groups. Metabolites derived from vitamins and recovery supplements common in professional sports, including glutamine, lysine, 4-pyridoxic acid and nicotinamide, are also raised in the athlete group. It is notable that PAG, a microbial conversion product of phenylalanine, has been associated with a lean phenotype and is increased in the athletes.[42] Furthermore, PAG positively correlates with the genus *Erysipelotrichaceae incertae sedis*, which we have previously noted to be present in relatively higher proportions in the athlete group compared with both control groups. PAG is the strongest biomarker postbariatric surgery, where it is associated with an increase in the relative proportions of Proteobacteria as observed here in the athlete group. Within the SCFAs, two distinct clusters were observed; acetic acid, propionic acid and butyric acid correlate with dietary contributors (fibre and protein), while isobutyric acid, isovaleric acid and valeric acid correlate with microbial diversity. The same clusters are observed when correlating with individual taxa, in support of previously observed links between SCFAs and numerous metabolic benefits and a lean phenotype.[33–35]

Our ongoing work in this area with non-athletes engaging in a structured exercise regime looks to further explore components of the exercise and diet–microbiome paradigm, which, along with this study, may inform the design of exercise and fitness programmes, including diet design in the context of optimising microbiota functionality for both athletes and the general population.

MATERIALS AND METHODS

Study population

Elite professional male athletes (n=40) and healthy controls (n=46) matched for age and gender were enrolled in 2011 as previously described in the study.[26] Due to the range of physiques within a rugby team (player position dictates need for a variety of physical constitutions, ie, forward players tend to have larger BMI values than backs, often in the overweight/obese range) the recruited control cohort was subdivided into two groups. To more completely include control participants, the BMI parameter for group inclusion was adjusted to BMI $\leq$25.2 and BMI $\geq$26.5 for the low-BMI and high-BMI groups, respectively. Approval for this study was granted by the Cork Clinical Research Ethics Committee.

Acquisition of clinical, exercise and dietary data

Self-reported dietary intake information was accommodated by a research nutritionist within the parameters of a food frequency questionnaire in conjunction with a photographic food atlas as per the initial investigation.[26] Fasting blood samples were collected and analysed at the Mercy University Hospital clinical laboratories, Cork. As the athletes were involved in a rigorous training camp, we needed to assess the physical activity levels of both control groups. To determine this, we used an adapted version of the EPIC-Norfolk questionnaire.[43] Creatine kinase levels were used as a proxy for level of physical activity across all groups.

Preparation of metagenomic libraries

DNA derived from faecal samples was extracted and purified using the QIAmp DNA Stool Mini Kit (cat. no 51 504) prior to storage at −80°C. DNA libraries were prepared with the Nextera XT DNA Library Kit (cat. no FC-131-1096) prior to processing on the Illumina HiSeq 2500 sequencing platform (see online supplementary methods for further detail).

Downloaded from http://gut.bmj.com/ on March 30, 2017 - Published by group.bmj.com

Metagenomic statistical and bioinformatic analysis

Delivered raw FASTQ sequence files were quality checked as follows: contaminating sequences of human origin were first removed through the NCBI Best Match Tagger (BMTagger). Poor-quality and duplicate read removal, as well as trimming was implemented using a combination of SAM (sequence alignment map) and Picard tools. Processing of raw sequence data produced a total of 2 803 449 392 filtered reads with a mean read count of 32 598 248.74 (±10 639 447 SD) per each of the 86 samples. These refined reads were then subjected to functional profiling by the most recent iteration of the Human Microbiome Project Unified Metabolic Analysis Network (HUMAnN2 V.0.5.0) pipeline.[44] The functional profiling performed by HUMAnN2 composed tabulated files of microbial metabolic pathway abundance and coverage derived from the Metacyc database.[45] Microbial pathway data were statistically analysed in the R software environment (V.3.2.2) (for further details see online supplementary methods) (R Development Core Team. R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Vienna, 2012). 2015. http://www.R-project.org). All presented p values were corrected for multiple comparisons using the Benjamini-Hochberg false discovery rate (pFDR) method.[46]

Metabolic profiling

Urine and faecal samples were prepared for metabonomic analysis as previously described.[47 48] Using established methods, urine samples underwent ^{1}H-NMR, RP and HILIC chromatography profiling experiments. Faecal samples underwent ^{1}H-NMR, HILIC and bile acid UPLC-MS profiling experiments and GC-MS-targeted SCFA analysis.[48–50]

After data preprocessing,[51] the resulting ^{1}H-NMR and LC-MS data sets were imported into SIMCA 14.1 (Umetrics) to conduct multivariate statistical analysis. PCA, followed by OPLS-DA, was performed to examine the data sets and to observe clustering in the results according to the predefined classes. The OPLS-DA models in this study were established based on one PLS component and one orthogonal component. Unit variance scaling was applied to ^{1}H-NMR data, Pareto scaling was applied to MS data. The fit and predictability of the models obtained were determined by the R^2Y and Q^2Y values, respectively. Significant metabolites were obtained from LC-MS OPLS-DA models through division of the regression coefficients by the jack-knife interval SE to give an estimate of the t-statistic. Variables with a t-statistic ≥1.96 (z-score, corresponding to the 97.5 percentile) were considered significant. Significant metabolites were obtained from ^{1}H-NMR OPLS-DA models after investigating correlations with correlation coefficients values higher than 0.4. Univariate statistical analysis (Mann-Whitney U test) was used to examine the SCFA data set. p values were adjusted for multiple testing using the pFDR method.

Confirmation of metabolite identities in the NMR data was obtained using 1D ^{1}H NMR and 2D ^{1}H-^{1}H NMR and ^{1}H-^{13}C NMR experiments. In addition, statistical tools such as SubseT Optimization by Reference Matching (STORM) and Statistical TOtal Correlation SpectroscopY (STOCSY) were also applied.[52 53] Confirmation of metabolites identities in the LC-MS data was obtained using tandem MS (MS/MS) on selected target ions.

Metabolite identification was characterised by a level of assignment (LoA) score that describes how the identification was made.[54] The levels used were as follows: LoA 1: identified compound, confirmed by comparison to an authentic chemical reference. LoA 2: MS/MS precursor and product ions or 1D +2D NMR chemical shifts and multiplicity match to a reference database or literature to putatively annotate compound. LoA 3: chemical shift (δ) and multiplicity matches a reference database to tentatively assign the compound (for further details see online supplementary methods).

Twitter Follow Orla O'Sullivan @OrlaOS

Acknowledgements The authors express gratitude to all participants for the donation of time and samples, in particular staff and players at the Irish Rugby Football Union. Fiona Fouhy for insight provided into the library preparation of metagenomic sequencing. The authors thank the Imperial-National Institute for Health Research (NIHR) Clinical Phenome Centre for support.

Contributors WB prepared DNA samples for metagenomic sequencing. OOS and WB processed and analysed the metagenomic data. EH, IGP and NCP performed metabolomic processing and statistical analysis thereof. FS, PDC, OOS and WB devised experimental design and approach. FS, PDC, OC, OOS, MGM, EH, NCP and WB wrote manuscript. Results discussed by all authors.

Funding This research was funded by Science Foundation Ireland in the form of a centre grant (APC Microbiome Institute Grant Number SFI/12/RC/2273). Research in the Cotter laboratory is funded by SFI through the PI award, 'Obesibiotics' (11/PI/1137). OOS and WB are funded by Science Foundation Ireland through a Starting Investigator Research Grant award (13/SIRG/2160). Nicholas Penney is funded by the Diabetes Research and Wellness Foundation through the Sutherland-Earl Clinical Research Fellowship 2015. The centre is supported by the NIHR Imperial Biomedical Research Centre based at Imperial College Healthcare National Health Service (NHS) Trust and Imperial College London.

Disclaimer The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

Competing interests FS is a founder shareholder in Atlantia Food Clinical Trials, Tucana Health and Alimentary Health. He is director of the *APC Microbiome Institute*, a research centre funded in part by Science Foundation Ireland (APC/SFI/12/RC/2273) and which is/has recently been in receipt of research grants from Abbvie, Alimentary Health, Cremo, Danone, Janssen, Friesland Campina, General Mills, Kerry, MeadJohnson, Nutricia, 4D pharma and Second Genome, Sigmoid pharma.

Ethics approval Cork Clinical Research Ethics Committee.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement In conformation of data accessibility protocol, metagenomic raw sequence data from this study are deposited in EMBL BNucleotide Sequence Database (ENA) (http://www.ebi.ac.uk/ena/data/), accession number PRJEB15388.

REFERENCES

1 Harkin A. Muscling in on depression. *N Engl J Med* 2014;371:2333–4.
2 Benatti FB, Pedersen BK. Exercise as an anti-inflammatory therapy for rheumatic diseases-myokine regulation. *Nat Rev Rheumatol* 2015;11:86–97.
3 Hawley JA, Krook A. Metabolism: one step forward for exercise. *Nat Rev Endocrinol* 2016;12:7–8.
4 Hoffman-Goetz L, Pervaiz N, Packer N, *et al*. Freewheel training decreases pro- and increases anti-inflammatory cytokine expression in mouse intestinal lymphocytes. *Brain Behav Immun* 2010;24:1105–15.
5 Barton W, Shanahan F, Cotter PD, *et al*. The metabolic role of the microbiota. *Clin Liver Dis* 2015;5:91–3.
6 Szuhany KL, Bugatti M, Otto MW. A meta-analytic review of the effects of exercise on brain-derived neurotrophic factor. *J Psychiatr Res* 2015;60:56–64.
7 Ryan SM, Nolan YM. Neuroinflammation negatively affects adult hippocampal neurogenesis and cognition: can exercise compensate? *Neurosci Biobehav Rev* 2016;61:121–31.
8 Johannesson E, Simrén M, Strid H, *et al*. Physical activity improves symptoms in irritable bowel syndrome: a randomized controlled trial. *Am J Gastroenterol* 2011;106:915–22.
9 Robsahm TE, Aagnes B, Hjartaker A, *et al*. Body mass index, physical activity, and colorectal cancer by anatomical subsites: a systematic review and meta-analysis of cohort studies. *Eur J Cancer Prev* 2013;22:492–505.
10 Schwingshackl L, Missbach B, Dias S, *et al*. Impact of different training modalities on glycaemic control and blood lipids in patients with type 2 diabetes: a systematic review and network meta-analysis. *Diabetologia* 2014;57:1789–97.
11 Biswas A, Oh PI, Faulkner GE, *et al*. Sedentary time and its association with risk for disease incidence, mortality, and hospitalization in adults: a systematic review and meta-analysis. *Ann Intern Med* 2015;162:123–32.

Downloaded from http://gut.bmj.com/ on March 30, 2017 - Published by group.bmj.com

12 Same RV, Feldman DI, Shah N, *et al*. Relationship between sedentary behavior and cardiovascular risk. *Curr Cardiol Rep* 2016;18:6.
13 Wilmot EG, Edwardson CL, Achana FA, *et al*. Sedentary time in adults and the association with diabetes, cardiovascular disease and death: systematic review and meta-analysis. *Diabetologia* 2012;55:2895–905.
14 Chen YC, Tu YK, Huang KC, *et al*. Pathway from central obesity to childhood asthma. Physical fitness and sedentary time are leading factors. *Am J Respir Crit Care Med* 2014;189:1194–203.
15 Koeth RA, Wang Z, Levison BS, *et al*. Intestinal microbiota metabolism of L-carnitine, a nutrient in red meat, promotes atherosclerosis. *Nat Med* 2013;19:576–85.
16 Tang WHW, Wang ZE, Levison BS, *et al*. Intestinal microbial metabolism of phosphatidylcholine and cardiovascular risk. *N Engl J Med* 2013;368:1575–84.
17 Tang WHW, Hazen SL. Microbiome, trimethylamine N-oxide, and cardiometabolic disease. *Transl Res* 2017;179:108–15.
18 Woting A, Pfeiffer N, Loh G, *et al*. Clostridium ramosum promotes high-fat diet-induced obesity in gnotobiotic mouse models. *MBio* 2014;5:e01530–14.
19 Utzschneider KM, Kratz M, Damman CJ, *et al*. Mechanisms linking the gut microbiome and glucose metabolism. *J Clin Endocrinol Metab* 2016;101:1445–54.
20 Turnbaugh PJ, Ley RE, Mahowald MA, *et al*. An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* 2006;444:1027–31.
21 Williams NC, Johnson MA, Shaw DE, *et al*. A prebiotic galactooligosaccharide mixture reduces severity of hyperpnoea-induced bronchoconstriction and markers of airway inflammation. *Br J Nutr* 2016;116:798–804.
22 Wang Z, Klipfell E, Bennett BJ, *et al*. Gut flora metabolism of phosphatidylcholine promotes cardiovascular disease. *Nature* 2011;472:57–63.
23 Zhernakova A, Kurilshikov A, Bonder MJ, *et al*. Population-based metagenomics analysis reveals markers for gut microbiome composition and diversity. *Science* 2016;352:565–9.
24 Cronin O, Molloy MG, Shanahan F. Exercise, fitness, and the gut. *Curr Opin Gastroenterol* 2016;32:67–73.
25 Estaki M, Pither J, Baumeister P, *et al*. Cardiorespiratory fitness as a predictor of intestinal microbial diversity and distinct metagenomic functions. *Microbiome* 2016;4:42.
26 Clarke SF, Murphy EF, O'Sullivan O, *et al*. Exercise and associated dietary extremes impact on gut microbial diversity. *Gut* 2014;63:1913–20.
27 O'Sullivan O, Cronin O, Clarke SF, *et al*. Exercise and the microbiota. *Gut Microbes* 2015;6:131–6.
28 Cronin O, O'Sullivan O, Barton W, *et al*. Gut microbiota: implications for sports and exercise medicine. *Br J Sports Med*. Published Online First 11 January 2017.
29 Rankin A, O'Donavon C, Madigan SM, *et al*. 'Microbes in sport'—the potential role of the gut microbiota in athlete health and performance. Br J Sports Med. *Published Online First 25 January* 2017.
30 Petriz BA, Castro AP, Almeida JA, *et al*. Exercise induction of gut microbiota modifications in obese, non-obese and hypertensive rats. *BMC Genomics* 2014;15:511.
31 Maffetone PB, Laursen PB. Athletes: fit but unhealthy? *Sports Med Open* 2016;2:24.
32 Flint HJ, Scott KP, Duncan SH, *et al*. Microbial degradation of complex carbohydrates in the gut. *Gut Microbes* 2012;3:289–306.
33 Koh A, De Vadder F, Kovatcheva-Datchary P, *et al*. From dietary fiber to host physiology: short-chain fatty acids as key bacterial metabolites. *Cell* 2016;165:1332–45.
34 Ridaura VK, Faith JJ, Rey FE, *et al*. Gut microbiota from twins discordant for obesity modulate metabolism in mice. *Science* 2013;341:1241214.
35 Hamer HM, Jonkers DM, Bast A, *et al*. Butyrate modulates oxidative stress in the colonic mucosa of healthy humans. *Clin Nutr* 2009;28:88–93.
36 den Besten G, van Eunen K, Groen AK, *et al*. The role of short-chain fatty acids in the interplay between diet, gut microbiota, and host energy metabolism. *J Lipid Res* 2013;54:2325–40.
37 Bennett BJ, de Aguiar Vallim TQ, Wang Z, *et al*. Trimethylamine-N-oxide, a metabolite associated with atherosclerosis, exhibits complex genetic and dietary regulation. *Cell Metab* 2013;17:49–60.
38 Holmes E, Loo RL, Stamler J, *et al*. Human metabolic phenotype diversity and its association with diet and blood pressure. *Nature* 2008;453:396–400.
39 Brancaccio P, Limongelli FM, Maffulli N. Monitoring of serum enzymes in sport. *Br J Sports Med* 2006;40:96–7.
40 Stratton SL, Bogusiewicz A, Mock MM, *et al*. Lymphocyte propionyl-CoA carboxylase and its activation by biotin are sensitive indicators of marginal biotin deficiency in humans. *Am J Clin Nutr* 2006;84:384–8.
41 Wilson GJ, Wilson JM, Manninen AH. Effects of beta-hydroxy-beta-methylbutyrate (HMB) on exercise performance and body composition across varying levels of age, sex, and training experience: a review. *Nutr Metab (Lond)* 2008;5:1.
42 Holmes E, Li JV, Athanasiou T, *et al*. Understanding the role of gut microbiome-host metabolic signal disruption in health and disease. *Trends Microbiol* 2011;19:349–59.
43 Wareham NJ, Jakes RW, Rennie KL, *et al*. Validity and repeatability of the EPIC-Norfolk Physical Activity Questionnaire. *Int J Epidemiol* 2002;31:168–74.
44 Abubucker S, Segata N, Goll J, *et al*. Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput Biol* 2012;8:e1002358.
45 Caspi R, Altman T, Billington R, *et al*. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of Pathway/Genome Databases. *Nucleic Acids Res* 2014;42:D459–71.
46 Benjamini Y, Hochberg Y. Controlling the false discovery rate—a practical and powerful approach to multiple testing. *J Roy Stat Soc B Met* 1995;57:289–300.
47 Dona AC, Jiménez B, Schäfer H, *et al*. Precision high-throughput proton NMR spectroscopy of human urine, serum, and plasma for large-scale metabolic phenotyping. *Anal Chem* 2014;86:9887–94.
48 García-Villalba R, Giménez-Bastida JA, García-Conesa MT, *et al*. Alternative method for gas chromatography-mass spectrometry analysis of short-chain fatty acids in faecal samples. *J Sep Sci* 2012;35:1906–13.
49 Want EJ, Wilson ID, Gika H, *et al*. Global metabolic profiling procedures for urine using UPLC-MS. *Nat Protoc* 2010;5:1005–18.
50 Sarafian MH, Lewis MR, Pechlivanis A, *et al*. Bile acid profiling and quantification in biofluids using ultra-performance liquid chromatography tandem mass spectrometry. *Anal Chem* 2015;87:9662–70.
51 Veselkov KA, Vingara LK, Masson P, *et al*. Optimized preprocessing of ultra-performance liquid chromatography/mass spectrometry urinary metabolic profiles for improved information recovery. *Anal Chem* 2011;83:5864–72.
52 Posma JM, Garcia-Perez I, De Iorio M, *et al*. Subset optimization by reference matching (STORM): an optimized statistical approach for recovery of metabolic biomarker structural information from 1H NMR spectra of biofluids. *Anal Chem* 2012;84:10694–701.
53 Cloarec O, Dumas ME, Craig A, *et al*. Statistical total correlation spectroscopy: an exploratory approach for latent biomarker identification from metabolic 1H NMR data sets. *Anal Chem* 2005;77:1282–9.
54 Sumner LW, Amberg A, Barrett D, *et al*. Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics* 2007;3:211–21.

The Metabolic Role of the Microbiota

REVIEW

The Metabolic Role of the Microbiota

Wiley Barton B.Sc.,*,† Fergus Shanahan M.D.,†,‡ Paul D. Cotter Ph.D.,*,‡ Orla O'Sullivan Ph.D.*

Expansive cross-disciplinary investigations have highlighted the complex interaction between the human host and its associated species-rich and highly variable population of microorganisms (microbiota; Fig. 1).[1-7] The human gut microbiota has been established as having a particularly important role in health and disease.[1-4,7,8] Significant technological advances have facilitated the characterization of key gut microbial populations, providing information that can be harnessed in a variety of ways from the laboratory bench to the patient's bedside.[1-4,6-9] Indeed, clinicians and investigators can noninvasively acquire clinically relevant samples in the form of patient stool to access insightful biometric data through the extraordinary advances in high-throughput sequencing, nuclear magnetic resonance, and mass spectroscopy platforms.[4,6,7,10] Resulting from the knowledge gained through these and other analyses, novel interventions designed to improve health by changing the gut microbiota (e.g., pre- and probiotics and fecal transplantation) have attracted increasing attention.[1,4,7,8] Furthermore, the acquisition of additional knowledge of the gut microbiota is expected to yield new strategies to address disorders arising from microbiota dysfunction by means of highly specific pharmaceuticals, alterations in diet and perhaps even physical activity.[1-4,8-10]

Diversity of Clinical Relevance

The development of such therapies is highly desirable, given that the gut microbiota has been considered an important factor in a broad variety of disease pathologies.[1,3,4,6-10] Implicated in various neuropsychological conditions, probiotic administration has been shown to alleviate symptoms of anxiety and depression, and normalize levels of proinflammatory cytokines and cortisol.[3] Enrichment of microbiota in colorectal tumor tissue has illustrated a potential link between the microbiome and colorectal cancer.[7] Obesity is a condition with an apparently important microbiota component, with respect to both the onset and maintenance of the disease state.[3,4,7,8,10] One proposed link relates to the presence of microbiota with limited diversity, leading to an overgrowth of key undesirable species that results in metabolic dysfunction.[8] Supporting this, various expressions of inimical dysbiosis have now been associated with obesity.[3,7,10] Furthermore, metagenomic analysis of obese individuals indicates an inverse connection between adiposity content and richness of both enteric bacteria and, as would be expected, bacterial genes.[4,10] Conversely, robust health markers are reported to coincide with greater diversity of the microbiome.[2] In the context of obesity, an assessment of health-associated factors affected by compromised gene density has identified impairment of both oxidative stress remediation and the production of short-chain fatty acids (SCFAs)[4] as well as increased fasting triglyceride serum levels and an altered inflammatory profile.[10]

Of particular interest, the human metabolome, that is, the complete set of small-molecule chemicals generated as a consequence of metabolism, is significantly affected by the gut microbiome.[1,4,6,7] Indeed, our gut microbiota maintains a collective repertoire of genes exceeding the native human genome by 150-fold[3,4] with metabolic contribution of a similar scale.[5] Of the numerous factors affecting this intricate network, diet has consistently been shown to influence the basal metabolomic function of host and microbiota,[1-5,8-10] with subsequent *de novo* metabolic sources produced through elaborate feedback schemes between the host and various microbiota.[1,4,8] Host behavior would also seem to be part of this paradigm, as indicated by a positive correlation between enteric microbiota diversity and physical exercise.[2] Such a relationship may be mediated by inflammatory

Abbreviations: IBD, inflammatory bowel disease; IL6, interferon 6; NAFLD, nonalcoholic fatty liver disease; NASH, nonalcoholic steatohepatitis; SCFA, short-chain fatty acids; TNF-α, tumor necrosis factor α.

*From the *Teagasc Food Research Centre, Moorepark, Fermoy, Cork, Ireland; †Department of Medicine, University College Cork, Ireland; and ‡Alimentary Pharmabiotic Centre, Cork, Ireland.*

Potential conflict of interest: Nothing to report.

View this article online at wileyonlinelibrary.com

© 2015 by the American Association for the Study of Liver Diseases

doi: 10.1002/cld.455

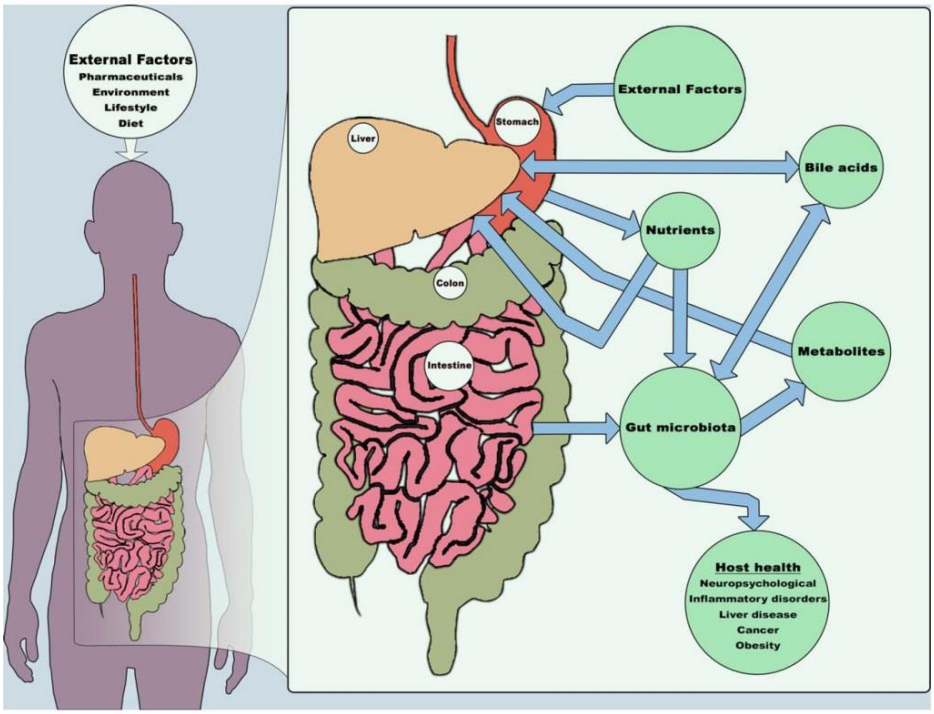


Figure 1 Diagram illustrating a simplification of the gut microbiome-health circuit. External factors such as diet, lifestyle, environment, and pharmaceuticals influence the gut microbiome and subsequently the health of the human host. Nutrients entering the circuit through the stomach stimulate activity in the liver and intestine. The enteric microbiota is affected both by the introduction of dietary nutrients and primary bile acids. Secondary metabolites (e.g., ethanol, short-chain fatty acids, and secondary bile acids) affect host health and contribute to communication between the liver and gut. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

markers such as interleukin 6 (IL6), and tumor necrosis factor α (TNF-α) being lower in active groups.[2] Complementing this, aged persons with compromised health parameters have been observed to have significantly higher levels of TNF-α and IL6 than their healthier counterparts.[5] Various microbial taxonomic groups and a wide number of clinically relevant metabolites may be key in this regard.[4,5,8] For instance, the SCFA butyrate and its associated producers (e.g., *Faecalibacterium praunsnitzii* and *Roseburia hominis*) are considered to mitigate some intestinal inflammatory disorders.[4,6,7] Of these, inflammatory bowel disease (IBD), ulcerative colitis in particular, is accompanied by reduced levels of SCFA and the gut microbes involved in SCFA synthesis from dietary carbohydrate fermentation.[4,6,7] Intriguingly, butyrate is reported to act as an energy source for colonocytes[1,3,6] and a stimulant of regulatory T cell differentiation.[6] This, together with dietary promotion of butyrate-producing species, [9] suggests the potential treatment of IBD with butyrate modulation.

Microbiota Influencing Liver Health

Intriguing insights regarding the impact of microbiota on liver health are being made, including a potential role for the microbiota in the pathogenesis of hepatic disease (Table 1). Microbial-derived ethanol has been shown to play a role in the onset and exacerbation of various stages of non-alcoholic fatty liver disease (NAFLD).[8] In addition, microbial interception of dietary choline can lead to simulation of diet-induced choline deficiency, a condition that contributes to liver disease, including NAFLD and non-alcoholic

TABLE 1 Enteric Microbiota Implicated in Liver Disease

Liver Condition	Implicated Microbiota	Potential Modulating Factors	References
NAFLD	Firmicutes *Lactobacillus* ↑ *Robinsoniella* ↑ *Roseburia* ↑ *Dorea* ↑ *Oscillibacter* ↓	Microbial ethanol productionFiaf inhibition leading to liver accumulation of triglycerides	7
NASH	Bacteroidetes ↑ Proteobacteria ↑	Diet induced dysbiosisHigh fat intakeMicrobial conversion of choline into methylamines	7
Liver cancer	Firmicutes ↓ Proteobacteria *Bilophila wadsworthia* ↑	Diet induced dysbiosisAnimal based dietDCA overproduction	9
Cirrhosis	Firmicutes Lachnospiraceae ↓ Ruminococcaceae ↓ Clostridium Incertae sedis XIV ↓ Lactobacillaceae ↑ Proteobacteria Entereobacteriaceae ↑ Alcaligenaceae ↑ Fusobacteria Fusobacteriaceae ↑	Stunted flow of bileBacterial overgrowthDysbiosis	7

This table illustrates examples of enteric microbiota implicated in various liver conditions. General phyla and included specific taxonomic groups are shown either reduced or increased in respect to the named conditions. Potential modulating factors indicate characteristics and influences of and upon the gut microbiome that contribute to the associated health condition.

steatohepatitis (NASH).[8] With similar clinical relevance, enterohepatic circulation is substantially influenced by the microbiota.[4,8] Bile acid reactions are modulated by enteric microbes[3,8,9] such as Alistipes, Bilophila, and Bacteroides, which in turn are influenced by diet.[9] Microbial gene expression of bile salt hydrolases and sulfite reductases is significantly increased in response to consumption of animal-derived food products.[9] With an abundance of bile salt hydrolases, concentrations of the secondary bile acid deoxycholic acid rapidly increase and with its suggested involvement in hepatic cancer promotion may lead to liver disease.[9] As bile acids exhibit bacteriostatic properties, perturbation of enterohepatic circulation may in turn have a variety of impacts on the host and its microbiota.[8,9] Indeed, that bile acid levels in cirrhosis patients are reduced in fecal samples may promote the selective bacterial overgrowth that is associated with advanced liver disease.[8] With limited availability, the rate of bile acid conjugation is reduced, leading to compromised regulation of epithelial cell integrity factors.[8] Complications of end-stage liver disease are also believed to be influenced by gut microbiota.[9] Hepatic encephalopathy is contributed to by microbial products including endotoxins and ammonia and is greatly remediated by antibiotic treatment.[9]

In summary, with substantial data highlighting its contribution to health and disease, the gut microbiome is poised to continue to be the focus of inquiry-driven research and to be increasingly used for diagnostic and therapeutic applications.■

CORRESPONDENCE
Dr. Paul Cotter, Teagasc Food Research Centre, Moorepark, Fermoy, Co. Cork, Ireland. E-mail: Paul.cotter@teagasc.ie.

References

1. Mani S, Boelsterli UA, Redinbo, MR. Understanding and Modulating Mammalian-Microbial Communication for Improved Human Health. Annu Rev Pharmacol Toxicol 2014;54:559-580.
2. Clarke SF, Murphy EF, O'Sullivan O, Lucey AJ, Humphreys M, Hogan A, et al. Exercise and associated dietary extremes impact on gut microbial diversity. Gut 2014;63:1913-1920.
3. Cryan JF, Dinan TG. Mind-altering microorganisms: the impact of the gut microbiota on brain and behaviour. Nat Rev Neurosci 2012;13:701-712.
4. Ursell LK, Haiser HJ, Van Treuren W, Garg N, Reddivari L, Vanamala J, et al. The intestinal metabolome: an intersection between microbiota and host. Gastroenterology 2014;146:1470-1476.
5. Claesson MJ, Jeffery IB, Conde S, Power SE, O'Connor EM, Cusack S, et al. Gut microbiota composition correlates with diet and health in the elderly. Nature 2012;488:178-184.
6. Dorrestein PC, Mazmanian SK, Knight R. Finding the Missing Links among Metabolites, Microbes, and the Host. Immunity 2014;40: 824-832.
7. Guinane CM, & Cotter PD. Role of the gut microbiota in health and chronic gastrointestinal disease: understanding a hidden metabolic organ. Ther Adv Gastroenterol 2013;6:295-308.
8. Schnabl B, Brenner DA. Interactions between the intestinal microbiome and liver diseases. Gastroenterology 2014;146:1513-1524.
9. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, et al. Diet rapidly and reproducibly alters the human gut microbiome. Nature 2014;505:559-563.
10. Cotillard A, Kennedy SP, Kong LC, Prifti E, Pons N, Le Chatelier E, et al. Dietary intervention impact on gut microbial gene richness. Nature 2013; 500:585-588.

Appendix E

Original Programming

fdatish (format data 'ish)

```
#' Format data in a certain way
#'
#' @param input R object or file - required
#' @param subject ID of individuals - one of: row/col position, NULL, or blank
#' @param group Group variable - one of: row/col position, NULL, or blank
#' @param VarOrient - are variables along rows or cols - one of: "row" or "col"
#' @return list containing numeric data, identification variables as factors, and
 the two     #' combined
#' @seealso N/A
#' @export
#' @examples
#' fthis <- data_frame(
#' ids = letters[1:10],
#' grouping = rep(letters[1:2], 5),
#' var_1 = sample(.1:10),
#' var_2 = sample(.1:10),
#' var_3 = sample(.1:10))
#'
#' fdatish(fthis)
#'
#' wd <- getwd()
#' wd_file <- paste0(wd,"cool_data.txt")
#' fdatish("wd_file")
#'
#ADD OPTION FOR BASE/DATA/NAME ONLY
fdatish <- function(input, subject, group, VarOrient){
  if(missing(input)==TRUE){
    print("Huh, there isn't data supplied, we cannot proceed!

(Hint: input = `your_data.txt`/your_data)")
    stopifnot(missing(input)==FALSE)
  }
  if(is.matrix(input) == FALSE && is.data.frame(input) == FALSE){
```

```
    #DETECT AND DEFINE FILE TYPE WITH INPUT TITLE EG pizza.txt
    print("the data supplied is NOT an object")
    sep.var <- readLines(input, n = 1)
    if (grepl(",", sep.var)) {sep.var = ","}
    if (grepl("\\;", sep.var)) {sep.var = ";"}
    if (grepl("\\t", sep.var)) {sep.var = "\t"}
    if(missing(VarOrient)==TRUE){
      print(sep.var)
      fdat.tmp <- read.table(input, sep = sep.var)
      rowvarP = 0;colvarP = 0
      if(ncol(fdat.tmp) == nrow(fdat.tmp)) {rowvarP = 0;colvarP = 0}
      if(ncol(fdat.tmp) < nrow(fdat.tmp)) {rowvarP = 1} else {colvarP = 1}
      if(length(levels(fdat.tmp[,1])) == nrow(fdat.tmp)) {rowvarP = rowvarP + 1}
      print(rowvarP);print(colvarP)
      if(rowvarP == colvarP) {print("Orientation could not be determined. Specify
 'VarOrient'")}
      if(rowvarP > colvarP) {
        VarOrient = "row"
        print("It was detected that your variable are along the rows. Congrats?")
      } else if(rowvarP < colvarP){
          VarOrient = "col"
        print("It was detected that your variable are along the cols. Congrats?")
}
    }
    if(VarOrient == "col"|VarOrient == "cols"){
      fdat.core <- as.matrix(read.table(input,
                                        row.names = NULL,
                                        quote=NULL,
                                        comment="",
                                        sep=sep.var,
                                        stringsAsFactors=FALSE,
                                        header=TRUE))
    } else if(VarOrient == "row"|VarOrient == "rows"){
      fdat.core <- t(as.matrix(read.table(input,
                                          row.names = 1,
                                          quote=NULL,
                                          comment="",
                                          sep=sep.var,
                                          stringsAsFactors=FALSE,
                                          header=FALSE)))
    }
```

```
} else {
  fdat.core <- input
}
if(missing(subject) == TRUE) {
  print("subject missing")
} else if(is.null(subject)) {
  print("NULL SUBJECT")
} else if(missing(subject) == FALSE){
  pos.ID <- subject
}
if(missing(group) == TRUE){
  print("group missing")
} else if(is.null(group)) {
  print("NULL GROUP")
} else if(missing(group) == FALSE){
  pos.GRP <- group
}
if(exists("pos.GRP")==TRUE|exists("pos.ID")==TRUE){
  print("G|I=T")
  if(exists("pos.GRP")==TRUE&exists("pos.ID")==TRUE){
    print("G&I=T")
    pos.GRP <- as.numeric(pos.GRP)
    pos.ID <- as.numeric(pos.ID)
    if(pos.GRP < pos.ID){
      fdat.names <- fdat.core[,c(pos.GRP,pos.ID)]
    } else if(pos.GRP > pos.ID){
      fdat.names <- fdat.core[,c(pos.ID,pos.GRP)]}
    dimnames(fdat.names)[[2]][pos.GRP] <- "GRP.var"
    dimnames(fdat.names)[[2]][pos.ID] <- "ID.var"
    fdat.names <- as.data.frame(fdat.names)}
  if(exists("pos.GRP")==TRUE&exists("pos.ID")==FALSE){
    fdat.names <- fdat.core[,c(pos.GRP)]
    dimnames(fdat.names)[[2]][pos.GRP] <- "GRP.var"
    fdat.names <- as.data.frame(fdat.names)}
  if(exists("pos.GRP")==FALSE&exists("pos.ID")==TRUE){
    fdat.names <- fdat.core[,c(pos.ID)]
    dimnames(fdat.names)[[2]][pos.ID] <- "ID.var"
    fdat.names <- as.data.frame(fdat.names)}
  #ADD OPT FOR METADATA
}
if((pos.GRP < length(fdat.core)|pos.ID < length(fdat.core))&(pos.GRP == 1|pos.I
```

```
D == 1)){
    fdat.data <- fdat.core[,-(1:length(fdat.names))];print("Base table arranged:
name-data")
  } else {
    fdat.data <- fdat.core[,(1:(length(fdat.core)-length(fdat.names)))];

print("Base table arranged: data-name")
  }
  class(fdat.data) <- "numeric"
  fdat.base <- cbind.data.frame(fdat.names, fdat.data)
  fdat.out <- structure(list(base=fdat.base,data = fdat.data, names = fdat.names)
)
  return(fdat.out)
}
```

fdis (format distance)

```
#' Format data in preparation for ordination
#'
#' @param var.mat matrix of variables, preferably as fdatish object. e.g. taxa -
required
#' @param ID Classifier of subject subset e.g. group - one of: row/col position,
NULL, or blank
#' @param pair.ID Classifier of paired samples - factor list
#' @param nom.hi First list of classifiers defining variable subset e.g. d_bacter
ia - one of: list
#' @param nom.lo Final list of classifiers defining variable subset e.g. _species
 - one of: object
#' @param nom.split Classifier of sample subsetting - factor list
#' @param rm.var Remove empty variables i.e. variables with sum = 0 - one of: TRU
E or FALSE
#' @param rm.unpar Remove sites/subjects unpaired according to 'pair.ID' - one of
: TRUE or FALSE
#' @param opt.split Subset data according to grouping variable - one of: TRUE, FA
LSE or BOTH
#' @param opt.norm Perform normalisation on supplied data - one of: TRUE or FALSE
#' @param opt.comb Combine RA from separate subsets - one of: TRUE, FALSE or BOTH
#' @param opt.Mnorm Perform normalisation on subsets of variables - one of: TRUE
or FALSE
#' @param opt.dist Generate distance matrix for data - one of: TRUE or FALSE
```

```
#' @param opt.comment Provide commentary on function processing - one of: TRUE or
 FALSE
#' @param norm.meth Method of normalisation to be performed by vegan::decostand -
 one of: "total","max","freq","normalize","range","standardize","pa"
#' @param dist.meth Ordination method - one of "PCOA" or "NMDS"
#'
#' @return list containing numeric data, identification variables as factors, and
 the two combined
#' @seealso ampvis2
#' @export
#' @examples
#' fthisDat <- data_frame(
#' ids = letters[1:10],
#' grouping = rep(letters[1:2], 5),
#' TimePoint = rep(1:2,5)
#' var_1 = sample(.1:10),
#' var_2 = sample(.1:10),
#' var_3 = sample(.1:10),
#' var_4 = sample(.1:10),
#' var_5 = sample(.1:10))
#'
#' fdis_out <- fdis(var.mat = fthisDat[4:8], ID = fthisDat$ids, pair.ID = fthisDa
t$TimePoint,
#'                  nom.hi = list("A","B","C"), nom.lo = "D", nom.split,
#'                  rm.var = FALSE, rm.unpar = FALSE,
#'                  opt.split = FALSE, opt.norm = TRUE, opt.Mnorm = FALSE, opt.com
b = FALSE, opt.dist = TRUE, opt.comment = FALSE,
#'                  norm.meth = "total", dist.meth = "NMDS")
#'
#
# NEED TO IMPLEMENT METHOD FOR PATHWAY (| and __S)
fdis <- function(var.mat, ID, pair.ID,
                 nom.hi, nom.lo, nom.split,
                 rm.var = FALSE, rm.unpar = FALSE,
                 opt.split = FALSE, opt.norm, opt.Mnorm, opt.comb, opt.dist, opt.
comment = FALSE, norm.meth, dist.meth){
#VARIABLE ASSESSMENT
  if (opt.comment != TRUE && opt.comment != FALSE){
    opt.comment = FALSE
  }
  if (missing(var.mat)==TRUE){
```

```
    if(opt.comment == TRUE){
    print("Huh, data aren't supplied, we cannot proceed! (Hint: var.mat = `your_d
ata.txt`/your_data)")
    }
    stopifnot(missing(var.mat)==FALSE)
  }
  if (is.matrix(var.mat) == FALSE && is.data.frame(var.mat) == FALSE | is.numeric
(var.mat) == FALSE){
    if (opt.comment == TRUE){
    print("'var.mat' is in unacceptable format, atempting to fix...")
    }
    if (is.numeric(var.mat) == FALSE){
      if (missing(ID) == TRUE | missing(pair.ID) == TRUE | missing(nom.split) ==
TRUE){
        ID.tmp <- var.mat[,sapply(var.mat, is.factor)]
      }
      var.mat <- as.matrix(var.mat[,sapply(var.mat, is.numeric)])
      if (is.numeric(var.mat) == FALSE){
        if (opt.comment == TRUE){
          stop ( "Could not convert 'var.mat' to an acceptable format")
          print("Could not convert 'var.mat' to an acceptable format")
        }
      } else {
        if (opt.comment == TRUE){
          print("'var.mat' successfully converted")
        }
      }
    }
  }
  if (missing(ID) == TRUE){
    if (opt.comment == TRUE){
    print( "Huh, no identification variables are supplied, attempting to locate.w
e cannot proceed!

(Hint: ID = your_data$GROUP)")
    }
    if (exists("ID.tmp") == TRUE){
      if (length(ID.tmp) == 1){
        ID <- ID.tmp
        if (opt.comment == TRUE){
        print(paste("ID variable used:", colnames(ID.tmp)[1]))
```

```
        }
      }
      if (length(ID.tmp) > 1){
        if (length(grep("G",colnames(ID.tmp), value = FALSE, ignore.case = TRUE))
 > 0){
          if (length(grep("G",colnames(ID.tmp), value = FALSE, ignore.case = TRUE
)) > 1){
            if (length(grep("Group",colnames(ID.tmp), value = FALSE, ignore.case
= TRUE)) == 1){
              if (opt.comment == TRUE){
                print(paste("ID variable used:", colnames(ID.tmp)[grep("Group",co
lnames(ID.tmp),
value = FALSE, ignore.case = TRUE)]))
              }
              ID <- ID.tmp[grep("Group", colnames(ID.tmp), value = TRUE, ignore.c
ase = TRUE)]
            }
            if (length(grep("G.p",colnames(ID.tmp), value = FALSE, ignore.case =
TRUE)) == 1){
              if (opt.comment == TRUE){
                print(paste("ID variable used:", colnames(ID.tmp)[grep("G.p",coln
ames(ID.tmp),
value = FALSE, ignore.case = TRUE)]))
              }
              ID <- ID.tmp[grep("G.p", colnames(ID.tmp), value = TRUE, ignore.cas
e = TRUE)]
            }
          } else {
            if (length(grep("G",colnames(ID.tmp), value = FALSE, ignore.case = TR
UE)) == 1){
              ID <- ID.tmp[grep("G", colnames(ID.tmp), value = TRUE, ignore.case
= TRUE)]
              if (opt.comment == TRUE){
                print(paste("ID variable used:", colnames(ID.tmp)[grep("G",colnam
es(ID.tmp),
value = FALSE, ignore.case = TRUE)]))
              }
            }
          }
```

```
        }
        if (length(grep("G",colnames(ID.tmp), value = FALSE, ignore.case = TRUE))
 == 0){
          for (i in 1 : length(ID.tmp)){
            if (exists("ID") == FALSE){
              if (length(levels(ID.tmp[[i]])) > 1){
                ID <- ID.tmp[[i]]
                if (opt.comment == TRUE){
                  print(paste("ID variable used:", colnames(ID.tmp)[[i]]))
                }
              }
            }
          }
        }
      }
    }
  }
  if(is.factor(ID) == FALSE | is.character(ID) == TRUE){
    if(opt.comment == TRUE){
    print("'ID' is in an unacceptable format, atempting to fix...")
    }
    if(is.character(ID) == TRUE){
      ID <- as.factor(ID)
    }
    if(is.factor(ID) == FALSE){
      if(opt.comment == TRUE){
      print("Fix unsuccessful, try suppling as factor!")
      }
    }
  }
  if(missing(nom.hi)== TRUE | is.null(nom.hi) == TRUE && missing(nom.lo) == TRUE
|

is.null(nom.lo) == TRUE){
    if(opt.comment == TRUE){
    print("Variable subsetting not applied")
    }
    nom.hi = NULL
    nom.lo = NULL
  } else {
    if(missing(nom.hi)==TRUE){
```

```
    if(opt.comment == TRUE){
    print("High level variable subsetting not applied")
    }
    nom.hi = NULL
  }
  if(is.list(nom.hi) == FALSE && is.character(nom.hi) == FALSE){
    if(opt.comment == TRUE){
    print("'nom.hi' is in an unacceptable format, atempting to fix...")
    }
    if(is.character(nom.hi) == TRUE){
      nom.hi <- as.list(nom.hi)
    }
    if(is.list(nom.hi) == FALSE){
      if(opt.comment == TRUE){
      print("Fix unsuccessful, try suppling as list!")
      }
    }
  }
  if(missing(nom.lo)==TRUE){
    if(opt.comment == TRUE){
    print("Low level variable subsetting not applied")
    }
    nom.lo = NULL
  }
  if(is.character(nom.lo) == FALSE){
    if(opt.comment == TRUE){
    print("'nom.lo' is in an unacceptable format, atempting to fix...")
    }
    if(is.character(nom.lo) == FALSE){
      nom.lo <- as.character(nom.lo)
    }
    if(is.character(nom.lo) == FALSE){
      if(opt.comment == TRUE){
      print("Fix unsuccessful, try suppling as character!")
      }
    }
  }
}
if(missing(opt.split) == TRUE | is.null(opt.split) | opt.split == FALSE){
  if(opt.comment == TRUE){
  print("Subjects will not be split")
```

```
    }
    nom.split = NULL
    opt.split = FALSE
  }
  if(opt.split == "Both" | opt.split == "both" | opt.split == "B" | opt.split ==
"b"){
    opt.split = "BOTH"
  }
  if(missing(nom.split)==TRUE | is.null(nom.split) == TRUE && opt.split == TRUE |
 opt.split == "BOTH"){
    if(opt.comment == TRUE){
    print("Identification variable for split (nom.split) not found, using 'ID' va
riable")
    }
    nom.split <- ID
  }
  if(missing(rm.var)==TRUE){
    if(opt.comment == TRUE){
    print("Empty variables will not be removed")
    }
    rm.var = FALSE
  }
  if(missing(rm.unpar)==TRUE){
    if(opt.comment == TRUE){
    print("Unpaired subjects will not be removed")
    }
    rm.unpar = FALSE
    stopifnot(missing(rm.unpar)==FALSE)
  }
  if(rm.unpar == TRUE && missing(pair.ID) == TRUE){
    if(opt.comment == TRUE){
    print("ID to match samples is missing, unpaired subjects will not be removed"
)
    }
    rm.unpar = FALSE
  }
  if(rm.unpar == TRUE && missing(pair.ID) == FALSE){
    if(is.factor(pair.ID) == FALSE){
      if(is.character(pair.ID) == FALSE){
      if(opt.comment == TRUE){
      print("'pair.ID' is in an unacceptable format, atempting to fix...")
```

```
    }
    if(is.character(pair.ID) == TRUE){
      pair.ID <- as.factor(pair.ID)
    }
    if(is.factor(pair.ID) == FALSE){
      if(opt.comment == TRUE){
      print("Fix unsuccessful, try suppling as factor!")
      }
      rm.unpar = FALSE
      pair.ID = NULL
    }
  }
  }
}
if(missing(opt.norm)==TRUE){
  if(opt.comment == TRUE){
  print("Data will not be normalised")
  }
  opt.norm = FALSE
}
if(missing(opt.Mnorm) == TRUE){
  if(opt.norm == TRUE){
    if(is.null(nom.hi) == FALSE){
      if(opt.comment == TRUE){
      print("Normalisation performed separately according to 'nom.hi")
      }
      opt.Mnorm = TRUE
    } else {if(is.null(nom.hi) == TRUE){
      if(opt.comment == TRUE){
      print("Normalisation performed on full data set")
      }
      opt.Mnorm = FALSE
    }}
  }
}
if(missing(opt.comb)==TRUE){
  if(opt.comment == TRUE){
  print("Data will not be combined")
  }
  opt.comb = FALSE
}
```

```
  if(opt.comb == "Both" | opt.comb == "both" | opt.comb == "B" | opt.comb == "b")
{
    opt.comb = "BOTH"
  }
  if(missing(opt.dist) == TRUE){
    if(opt.comment == TRUE){
    print("Distance matrix will not be generated")
    }
    opt.dist = FALSE
  }
  if(missing(norm.meth)==TRUE && missing(opt.norm)==TRUE | opt.norm == FALSE){
    if(opt.comment == TRUE){
    print("Data will definitely not be normalised")
    }
  }
  if(missing(norm.meth)==TRUE && opt.norm == TRUE){
    print("Default normalisation - total: divide by margin total")
    norm.meth = "total"
  }
  if(missing(dist.meth)==TRUE && opt.dist == TRUE){
    if(opt.comment == TRUE){
    print("Default ordination - PCoA")
    }
    dist.meth = "PCOA"
  }
  if(missing(dist.meth)==FALSE && opt.dist == TRUE){
    if(dist.meth != "PCOA" && dist.meth != "NMDS"){
    if(opt.comment == TRUE){
      print("Ordination method not supplied as 'PCOA' or 'NMDS', attempting to fi
x")
    }
      if(dist.meth == "pcoa" | dist.meth == "Pcoa" | dist.meth == "PCoA"){
        if(opt.comment == TRUE){
          print("Option determined: 'PCOA'")
        }
        dist.meth = "PCOA"
      }
      if(dist.meth == "nmds" | dist.meth == "Nmds" | dist.meth == "nMDS"){
        if(opt.comment == TRUE){
          print("Option determined: 'NMDS'")
        }
```

```
        dist.meth = "NMDS"
      }
    }
    if(dist.meth != "PCOA" && dist.meth != "NMDS"){
    if(opt.comment == TRUE){
      print("Could NOT determine ordination method. Default used: 'PCOA'")
    }
      dist.meth = "PCOA"}
  }
  if(missing(dist.meth)==TRUE && missing(opt.dist)==TRUE | opt.dist == FALSE){
    if(opt.comment == TRUE){
    print("Ordination data not generated")
    }
  }
  # Remove unpaired samples
  if(rm.unpar == TRUE && is.null(pair.ID) == FALSE){
    var.mat <- var.mat[pair.ID %in% (pair.ID)[duplicated(pair.ID)==TRUE],]
    ID <- ID[pair.ID %in% pair.ID[duplicated(pair.ID)==TRUE]]
    nom.split <- nom.split[pair.ID %in% pair.ID[duplicated(pair.ID)==TRUE]]
  }
  # SUBSET BY VARIABLE NAMES i.e. low lvl extraction -- IMPLEMENT METHOD FOR 'PAT
HWAYS'
  if(is.null(nom.hi) == FALSE | is.null(nom.lo) == FALSE){
    fdis.VAR.sub <- list()
    if(is.null(nom.hi) == FALSE){
      for (j in 1: length(nom.hi)){
        fdis.VAR.sub[[length(fdis.VAR.sub)+1]] <-
          var.mat[,c(grep(nom.hi[[j]], colnames(var.mat)))]
        if(is.null(nom.lo) == FALSE){
          fdis.VAR.sub[[length(fdis.VAR.sub)]] <-
            as.data.frame(fdis.VAR.sub[[length(fdis.VAR.sub)]][,c(grep(nom.lo, co
lnames(fdis.VAR.sub[[length(fdis.VAR.sub)]])))]) #AS>DATA>FRAME ADDED
        }
#        str(fdis.VAR.sub)
        if (exists("rm.var") == FALSE) {rm.var = FALSE}
        if (rm.var == TRUE){
#          if(fdis.VAR.sub[[length(fdis.VAR.sub)]])
          fdis.VAR.sub[[length(fdis.VAR.sub)]] <-
            fdis.VAR.sub[[length(fdis.VAR.sub)]][,c(which(colSums(fdis.VAR.sub[[l
ength(fdis.VAR.sub)]])!=0)), drop = FALSE] #ADDITION OF DROP
        }
```

```
      }
    }
    if(is.null(nom.hi) == TRUE && is.null(nom.lo) == FALSE){
      fdis.VAR.sub[[length(fdis.VAR.sub)+1]] <-
        var.mat[,c(grep(nom.lo, colnames(var.mat)))]
      if (exists("rm.var") == FALSE) {rm.var = FALSE}
      if (rm.var == TRUE){
        fdis.VAR.sub[[length(fdis.VAR.sub)]] <-
          fdis.VAR.sub[[length(fdis.VAR.sub)]][,c(which(colSums(fdis.VAR.sub[[len
gth(fdis.VAR.sub)]])!=0)), drop = FALSE]#ADDITION OF DROP
      }
    }
  }
  # Normalisation of multiple sets # POSSIBLE NEED FOR MARGIN.VAR
  if(opt.norm == TRUE){
    fdis.norm <- list()
    if(exists("fdis.VAR.sub")){
      if(opt.Mnorm == TRUE){
        for (i in 1 : length(fdis.VAR.sub)){
          fdis.norm[[length(fdis.norm) + 1]] <- decostand(fdis.VAR.sub[[i]], meth
od = norm.meth)
        }
      } else {
        if(opt.Mnorm == FALSE){
          if(length(fdis.VAR.sub) > 1){
            for(i in 1 : length(fdis.VAR.sub)){
              if(i == 1){
                fdis.norm.tmp <- fdis.VAR.sub[[i]]
              }
              if(i > 1){
                fdis.norm.tmp <- cbind(fdis.norm.tmp, fdis.VAR.sub[[i]])
              }
            }
            fdis.norm <- decostand(fdis.norm.tmp, method = norm.meth)
          }
          if(length(fdis.VAR.sub) == 1){
            fdis.norm <- decostand(fdis.VAR.sub, method = norm.meth)
          }
        }
      }
    }
```

```
    # Combine sets
    # MAY BE BROKEN - CHECK ID=TRUE colnames!!!!!!!
    # NOT ACCEPTING PATHWAY DATA WITH LOW LENGTH
    if(opt.comb == TRUE | opt.comb == "BOTH"){
      for (i in 1 : length(fdis.norm)){
        if (i == 1) {
          fdis.norm.tmp <- fdis.norm[[i]]
        } else {
          if (i < length(fdis.norm)){
            if (nrow(fdis.norm.tmp)==nrow(fdis.norm[[i]])){
              fdis.norm.tmp <- cbind(fdis.norm.tmp, fdis.norm[[i]])
            }
            if (nrow(fdis.norm.tmp)!=nrow(fdis.norm[[i]]) && ncol(fdis.norm.tmp)=
=ncol(fdis.norm[[i]])){
              fdis.norm.tmp <- rbind(fdis.norm.tmp, fdis.norm[[i]])
            }
          }
        }
        if (i == length(fdis.norm)){
          if (nrow(fdis.norm[[i]])==nrow(fdis.norm[[i-1]])){
            fdis.norm.tmp <- cbind(fdis.norm.tmp, fdis.norm[[i]])
          }
          if (nrow(fdis.norm[[i]])!=nrow(fdis.norm[[i-1]]) && ncol(fdis.norm.tmp)
==ncol(fdis.norm[[i]])){
            fdis.norm.tmp <- cbind(fdis.norm.tmp, fdis.norm[[i]])
          }
        }
      }
      fdis.Mnorm <- list()
      fdis.Mnorm[[1]] <- fdis.norm.tmp
      if(opt.comb == "BOTH"){
        fdis.norm[[length(fdis.norm)+1]] <- fdis.Mnorm
        fdis.Mnorm <- fdis.norm
      }
    }
  }
  # SUBSETTING DATA #!!!!!ADD DETAIL ON SPLIT COMMENT DYSFUNCTIONAL WITH INPUTS O
F 1 COLUMN!!!!
  if(opt.split == TRUE | opt.split == "BOTH"){
    fdis.split <- list()
    for(i in 1 : length(fdis.norm)){
```

```
      fdis.split.tmp <- (split(data.frame(fdis.norm[[i]]), nom.split))
      for(j in 1: length(levels(nom.split))){
        fdis.split[[length(fdis.split) + 1]] <- fdis.split.tmp[[j]]
        if(rm.var == TRUE){
          if(opt.split == TRUE){
            if(opt.comment == TRUE){
              print(paste("Checking for empty variables following data subsetting
 in: ",

levels(nom.split)[j]))
            }
            fdis.split[[length(fdis.split)]] <- fdis.split[[length(fdis.split)]][
,c(which(colSums(fdis.split[[length(fdis.split)]])!=0))]
            #print(ncol(fdis.split[[length(fdis.split)]]))
            #str(fdis.split)
            if(is.null(ncol(fdis.split[[length(fdis.split)]])) == TRUE){
              str(fdis.split[[length(fdis.split)]])
              print(ncol(fdis.split[[length(fdis.split)]]))
              if(opt.comment == TRUE){
                print(paste(nom.hi[[i]],"-",levels(nom.split)[j],"-","All variabl
es removed

after subsetting"))
              }
            }
            if(is.null(ncol(fdis.split[[length(fdis.split)]])) == FALSE){
              if(ncol(fdis.split[[length(fdis.split)]]) != ncol(fdis.split.tmp[[j
]])){
                if(opt.comment == TRUE){
                  print("Empty variables removed after subsetting")
                }
              }
            }
            if(is.null(ncol(fdis.split[[length(fdis.split)]])) == FALSE){
            if(ncol(fdis.split[[length(fdis.split)]]) == ncol(fdis.split.tmp[[j]]
)){
              if(opt.comment == TRUE){
                print("No empty variables detected after subsetting")
              }
            }
            }
```

```
          }
        }
      }
    }
    if(opt.split == "BOTH"){
      for(i in 1 : length(fdis.norm)){
        fdis.split[[length(fdis.split) + 1]] <- fdis.norm[[i]]
        if(rm.var == TRUE){
          if(opt.split == "BOTH"){
          if(opt.comment == TRUE){
            print(paste("Checking for empty variables following data subsetting i
n: ",

levels(nom.split)[i]))
          }
          fdis.split[[length(fdis.split)]] <- fdis.split[[length(fdis.split)]][,c
(which(colSums(fdis.split[[length(fdis.split)]])!=0))]
          if(ncol(fdis.split[[length(fdis.split)]]) != ncol(fdis.norm[[i]])){
            if(opt.comment == TRUE){
              print("Empty variables removed after subsetting")
            }
          }
          if(ncol(fdis.split[[length(fdis.split)]]) == ncol(fdis.norm[[i]])){
            if(opt.comment == TRUE){
              print("No empty variables detected after subsetting")
            }
          }
          }
        }
      }
    }
  }
  # Distance matrix generation OPT for methods
  # ADD OPT - POINT OR FULL WCMDSCALE OUT
  if(opt.dist == TRUE){
    print("467")
    fdis.dist <- list()
    fdis.eig <- list()
    if(exists("fdis.split") == TRUE){
      fdis.norm <- fdis.split
    }
```

```
    if(dist.meth == "PCOA"){
      for(i in 1: length(fdis.norm)){
        fdis.dist[[length(fdis.dist) + 1]] <- vegdist(fdis.norm[[i]], method = "b
ray")
        fdis.dist[[length(fdis.dist)]] <- wcmdscale(fdis.dist[[length(fdis.dist)]
], k=2, w=rep(1, nrow(fdis.norm[[i]])), eig = TRUE)
        fdis.eig[[length(fdis.eig) + 1]] <- eigenvals(fdis.dist[[length(fdis.dist
)]])
        fdis.eig[[length(fdis.eig)]] <- (fdis.eig[[length(fdis.eig)]]/sum(fdis.ei
g[[length(fdis.eig)]]))
        fdis.dist[[length(fdis.dist)]] <- data.frame(fdis.dist[[length(fdis.dist)
]]$points)
      }
    }
    if(dist.meth == "NMDS"){
      for(i in 1: length(fdis.norm)){
        fdis.dist[[length(fdis.dist) + 1]] <- metaMDS(fdis.norm[[i]], k = 2)
      }
    }
  }
  # COMPLETE STRATEGY FOR APPLICATION OF SUBJECT DETAILS TO OUTPUT
#  if(opt.ID.out == TRUE)
  ID.out.tmp <- as.data.frame(cbind(as.character(ID), nom.split))
  colnames(ID.out.tmp) <- c("GRP.var", "nom.split")
  ID.out.tmp <- split(ID.out.tmp, ID.out.tmp$nom.split)
  ID.out <- list()
  for(i in 1 : (length(levels(ID))/length(levels(nom.split)))){
    for(j in 1 : length(levels(nom.split))){
      ID.out[[length(ID.out) + 1]] <- ID.out.tmp[[j]]
      ID.out[[length(ID.out)]]$Variable.ID <- nom.hi[[i]]
    }
  }
#  print(str(fdis.norm))
  if(opt.dist == TRUE){
  for(i in 1 : length(fdis.dist)){
    fdis.dist[[i]] <- cbind(ID.out[[i]], fdis.dist[[i]])
  }
  }
  if(opt.norm == TRUE && opt.dist == TRUE){
    fdis.out <- structure(list(RA = fdis.norm, Distance = fdis.dist, eigan = fdis
.eig))
```

```
  }
  if(opt.norm == TRUE && opt.dist == FALSE){
    fdis.out <- structure(list(RA = fdis.norm))
  }
  if(opt.norm == FALSE && opt.dist == TRUE){
    fdis.out <- structure(list(Data = fdis.norm, Distance = fdis.dist, eigan = fd
is.eig))
  }
  return(fdis.out)
}
```

geom.insec Generate intersection points of a given value (e.g. median) to close geoms in GGPlot2

```
geom.insec <- function(gg.dat, group.var, x.var, y.var, cross.var){
  out.tmp <- NULL
  for(i in 1 : length(unique(group.var))){
    input.tmp <- gg.dat[group.var == levels(group.var)[i],]
    x.tmp <- as.numeric(x.var[group.var == levels(group.var)[i]])
    y.tmp <- y.var[group.var == levels(group.var)[i]]
    cord.tmp <- as.data.frame(cbind(x.tmp, y.tmp))
    grp.tmp <- do.call("rbind",
                       sapply(1:(length(group.var[which(group.var == group.var[i]
)]) - 1), function(j){
                         regres <- lm(x.tmp ~ y.tmp, cord.tmp[j:(j+1),])
                         if (regres$qr$rank < 2) return(NULL)
                         mSec <- predict(regres, newdata = data.frame(y.tmp = uni
que(cross.var[group.var == levels(group.var)[i]])))
                         if (x.tmp[j] < mSec & mSec < x.tmp[j+1])
                           return(data.frame(x = mSec, y = unique(cross.var[group
.var ==

levels(group.var)[i]])))
                         else return(NULL)
                       }
                       )
    )
    merge.tmp <- input.tmp[1:(nrow(grp.tmp)),]
    for (k in 1 : (ncol(input.tmp))){
      if (identical(as.numeric(x.tmp), as.numeric(as.character(input.tmp[1:(nrow(
input.tmp)),k]))) == TRUE |
          identical(as.numeric(y.tmp), as.numeric((input.tmp[,k]))) == TRUE) {
        if (identical(as.numeric(x.tmp), as.numeric(as.character(input.tmp[1:(nro
```

```
w(input.tmp)),k])))  == TRUE){
          merge.tmp[k] <- grp.tmp$x
        }
        if (identical(as.numeric(y.tmp), as.numeric((input.tmp[,k]))) == TRUE){
          merge.tmp[k] <- grp.tmp$y
        }
      }
    }
    out.tmp = rbind(out.tmp, merge.tmp)
  }
  return(out.tmp)
}

GG.data.mSec <- geom.insec(gg.dat = GG.data.in.form,
                           group.var = GG.data.in.form$group.ID,
                           x.var = GG.data.in.form$TP,
                           y.var = GG.data.in.form$multi.sites.raw,
                           cross.var = GG.data.in.form$multi.sites.mean)
```

falph (Normalise and generate alpha diversity on data)

```
falph <- function(site.mat, site.ID, group.ID,
                  div.meth = "shannon", norm.meth = "total", margin.var = 1,
                  opt.comment = FALSE, output.opt) {
  falph.in <- site.mat
  if (is.list(falph.in) == TRUE) {falph.tmp <- list()}
  if (is.list(falph.in) == FALSE) {falph.tmp = NULL}
  if (is.list(falph.in) == TRUE) {
    for (i in 1:length(falph.in)) {
      multi.sites = NULL
      if (is.null(norm.meth) == FALSE) {
        if (norm.meth == "wisconsin") {
          names.tmp <- falph.in[[i]][1:(sum(sapply(falph.in[[i]], is.factor)))]
          data.tmp <- falph.in[[i]][

(sum(sapply(falph.in[[i]], is.factor)) + 1) : ncol(falph.in[[i]])]
          data.tmp <- wisconsin(data.tmp)
          falph.in[[i]] <- cbind(names.tmp, data.tmp)
        }
        if (norm.meth != "wisconsin") {
          names.tmp <- falph.in[[i]][1:(sum(sapply(falph.in[[i]], is.factor)))]
          data.tmp <- falph.in[[i]][
```

```
(sum(sapply(falph.in[[i]], is.factor)) + 1) : ncol(falph.in[[i]])]
          data.tmp <- decostand(data.tmp, method = norm.meth, MARGIN = margin.var
)
          falph.in[[i]] <- cbind(names.tmp, data.tmp)
        }
      }
      for (j in 1:length(site.ID)) {
        site.tmp <- diversity(falph.in[[i]][j,(sum(sapply(falph.in[[i]], is.facto
r))+1) : ncol(falph.in[[i]])], index = div.meth)
        multi.sites <- rbind(multi.sites, site.tmp)
      }
#      ID.tmp <- cbind(as.character(site.ID), as.character(group.ID))
      ID.tmp <- as.data.frame(cbind(as.character(site.ID), as.character(group.ID)
))
      colnames(ID.tmp) <- c("site.ID", "group.ID")
      if (exists("pwy.opt") == FALSE) {pwy.opt = FALSE}
      if (pwy.opt == FALSE){
      ID.tmp$taxa = nom.hi[[i]]
      }
      ID.tmp$Diversity = div.meth
      multi.sites <- cbind(ID.tmp, multi.sites)
      falph.tmp[[length(falph.tmp) + 1]] <- multi.sites
    }
  }
  if (output.opt == "short") {falph.out <- falph.tmp}
  if (output.opt == "long") {
    if (is.list(falph.tmp) == TRUE) {
      melt.tmp = NULL
      for (i in 1:length(falph.tmp)) {
        melt.tmp = rbind(melt.tmp, falph.tmp[[i]])
        melt.tmp$taxa = as.factor(melt.tmp$taxa)
      }
      falph.out = melt.tmp
    }
  }
  return(falph.out)
}
```

BD.PCoA.GG (Wrapper function to generate GGplot2 ordination plots)

```
BD.PCoA.gg <- function(gg.Kin){
  plot.out.PCoA <-
  {
    {
      if (exists("col.fill") == FALSE | exists("grouping.var.main") == FALSE){
        if (exists("grouping.var.main")==FALSE){
          print("Bummer! 'grouping.var.main' was not set, attempting to guess wha
t you want...

'GG.data.in$grp/GRP.var/Group' being used")
          if ("grp"%in%names(GG.data.in) == FALSE & "GRP.var"%in%names(GG.data.in
) == FALSE &

"Group"%in%names(GG.data.in) == FALSE){
            print("Oh shit, the grouping variable wasn't detected in the GG.data.
in object!!

 GGplot cannot proceed.
                  Check the status of the variables!")
          } else {
            if("grp"%in%names(GG.data.in) == TRUE){
              grouping.var.main = GG.data.in$grp; print("GG.data.in$grp used to d
efine groups")}
            if("GRP.var"%in%names(GG.data.in) == TRUE){
              grouping.var.main = GG.data.in$GRP.var; print("GG.data.in$GRP.var u
sed to define

groups")}
            if("Group"%in%names(GG.data.in) == TRUE){
              grouping.var.main = GG.data.in$Group; print("GG.data.in$Group used
to define groups")}
          }
          if (exists("col.fill") == FALSE) {
            col.fill <- distinctColorPalette(k = length(levels(grouping.var.main)
))
            print(c("Palette Generated: ", col.fill))
          }
      }
        }
      if (exists("axis.data.X") == FALSE) {
        if ("V1"%in%names(GG.data.in) == TRUE) {
          axis.data.X <- GG.data.in$V1
```

```
          print("Using GG.data.in$V1 as X variable")}
        else if ("Dim1"%in%names(GG.data.in) == TRUE) {
          axis.data.X <- GG.data.in$Dim1
          print("Using GG.data.in$Dim1 as X variable")}
        else if ("V1"%in%names(GG.data.in) == FALSE & "Dim1"%in%names(GG.data.in)
 == FALSE)

{axis.data.X <- GG.data.in[1]
        print("X axis variable not supplied, using GG.data.in[1]")}
      }
      if (exists("axis.data.Y") == FALSE) {
        if ("V2"%in%names(GG.data.in) == TRUE) {
          axis.data.Y <- GG.data.in$V2
          print("Using GG.data.in$V2 as Y variable")}
        else if ("Dim2"%in%names(GG.data.in) == TRUE) {
          axis.data.Y <- GG.data.in$Dim2
          print("Using GG.data.in$Dim2 as Y variable")}
        else if ("V2"%in%names(GG.data.in) == FALSE & "Dim2"%in%names(GG.data.in)
 == FALSE)

{axis.data.Y <- GG.data.in[2]
        print("Y axis variable not supplied, using GG.data.in[2]")}
      }
      if (exists("alpha.col.var") == FALSE) {alpha.col.var <- 1.0}
      if (exists("alpha.fil.var") == FALSE) {alpha.fil.var <- 1.0}
      if (exists("lab.var") == FALSE) {lab.var <- levels(grouping.var.main)}
      if (exists("siz.dot.opt") == FALSE) {siz.dot.opt <- 2}
      if (exists("guide.var") == FALSE) {
        guide.var <-
          guides(alpha = FALSE,
                 size = FALSE,
                 color = guide_legend(override.aes = list(size = 5)),
                 fill = guide_legend(keywidth = 3, keyheight = 2)
          )
      }
      if (exists("theme.var") == FALSE) {
        theme.var <-
          theme(plot.title = element_text(hjust = 0.5),
                plot.margin = margin(10,10,10,10, "pt"),
                axis.text.x = element_text(size = rel(1.0), face = "bold"),
                axis.title.x = element_text(size = rel(1.0), face = "bold",
```

```
margin = margin(10,10,10,10, "pt")),
                axis.text.y = element_text(size = rel(1.0), face = "bold"),
                axis.title.y = element_text(size = rel(1.0), face = "bold",

margin = margin(10,10,10,10, "pt")),
                legend.text = element_text(size = rel(0.8), colour = "black",

angle = 0, face = "bold"),
                strip.text.x = element_text(size = rel(1.0), face = "bold",

margin = margin(1.5,1.5,1.5,1.5, "pt")),
                axis.ticks.length = unit(5, "pt"),
                axis.ticks.y = element_line(size = rel(.8)),
                axis.ticks.x = element_line(size = rel(.8)),
                panel.grid.minor = element_blank(),
                panel.grid.major = element_line(size = rel(.3)),
                legend.background = element_rect(color = "black"),
                legend.key.size = unit(2, "cm"),
                legend.position = "right",
                legend.margin = margin(-10.0,10.0,10.0,10.0))}
      if (is.list(guide.var) == FALSE){
        if (guide.var == "blank") {
          guide.var = NULL
          theme.var + theme(legend.position = "none")
        }
      }
      {
        if (exists("themeBW.var") == FALSE) {themeBW.var <-
          theme_bw(base_size = 18,
                   if (Sys.info()['sysname']=="Windows"){loadfonts(device = "win"
);base_family = "sans"}
                   else {base_family = "Helvetica"})
        }
        if (exists("hex.mode.on")==FALSE) {hex.mode.on = FALSE}
        if (hex.mode.on == TRUE) {
          ggplot(NULL, aes(axis.data.X, axis.data.Y))
        } else {
          ggplot()
        }
      }
  } +
      themeBW.var +
```

```
      theme.var +
      {
        if (exists("elip.opt") == FALSE) {elip.opt = FALSE}
        if (hex.mode.on == FALSE & elip.opt == TRUE) {
          if (exists("sub.scale.opt") == FALSE) {sub.scale.opt = FALSE}
          elip.pout <- list()
          elip.pout[[1]] <- stat_ellipse(aes(axis.data.X, axis.data.Y, fill = eli
p.fill.var), show.legend = FALSE,
                                         type = "t", level = elip.fill.CI, geom =
 "polygon"
          )
          elip.pout[[2]] <- stat_ellipse(level = elip.fill.CI,
                                         #if (identical(elip.fill.var,elip.colo.v
ar) ==TRUE){}
                                         if (sub.scale.opt == TRUE){
                                           aes(axis.data.X, axis.data.Y, fill = e
lip.colo.var)
                                         } else {
                                           aes(axis.data.X, axis.data.Y, color =
elip.colo.var)
                                         } )
          elip.pout
        }
      } +
      geom_vline(xintercept = 0) +
      geom_hline(yintercept = 0) +
      {
        if (exists("hex.mode.dot") == FALSE) {hex.mode.dot = TRUE}
        if (exists("dot.outline.opt") == FALSE) {dot.outline.opt = FALSE}
        if (hex.mode.dot == TRUE){
          if (dot.outline.opt == TRUE) {
            geom_point(data = GG.data.in,
                       aes(axis.data.X, axis.data.Y,
                           color = "black"),
                       size = siz.dot.opt+1)
          }
        }
      } + {
        if (hex.mode.dot == TRUE){
        geom_point(data = GG.data.in,
                   aes(axis.data.X, axis.data.Y,
```

```
                      color = grouping.var.main),
                  size = siz.dot.opt)
        }
      } +
      {
        if (exists("hex.mode.on") == FALSE) {hex.mode.on = FALSE} # NEEDDS WORK!!
!!!!!
        if (hex.mode.on == TRUE){
          hex.mode.pout <- list()
          for (h in 1 : 3){#length(hex.mode.dat)){
            hex.mode.pout[[h]] <- stat_binhex(data = hex.mode.dat[[h]],
                                              bins = hex.mode.bin, position = "id
entity",
                                              aes(
                                                x = hex.mode.dat[[h]]$Dim1,
                                                y = hex.mode.dat[[h]]$Dim2,
                                                alpha = ..count..), fill = hex.mo
de.col[h])
          }
          hex.mode.pout
        }
      } + {
        if(exists("sub.ID.opt") == FALSE) {sub.ID.opt = FALSE}
        if(sub.ID.opt == TRUE) {
          if (sub.ID.all == TRUE) {
            data.tmp = GG.data.in}
          else if (sub.ID.all == FALSE) {
            data.tmp = sub.ID.repel}
          geom_label_repel(
            data = data.tmp,
            size = rel(5),
            fontface = "bold",
            force = 2,
            segment.alpha = 0.5,
            box.padding = unit(.5, "lines"),
            color = "black",
            if (sub.ID.all == TRUE) {
              aes(axis.data.X, axis.data.Y,
                  label = sub.ID.var)}
            else {
              aes(axis.data.X, axis.data.Y,
```

```
                  fill = factor(sub.ID.repel$grp), #NEEDS WORK
                  label = sub.ID.var)}
          )}
      } + {
        if(exists("tit.var") == FALSE) {tit.var = "Woah, such graph!"}
        if(exists("Xax.var") == FALSE) {Xax.var = ""}
        if(exists("Yax.var") == FALSE) {Yax.var = ""}
        if(exists("tit.var")==TRUE&exists("Xax.var")==TRUE&exists("Yax.var")==TRU
E){
          labs(title = tit.var,x = Xax.var,y = Yax.var)
        }
      } + {
        if(exists("annote.opt") == FALSE) annote.opt = FALSE
        if(annote.opt)annote.obj
      } + {
        if(exists("elip.fill.leg") == FALSE) {elip.fill.leg = "legend"}
        if(hex.mode.on == FALSE) {
          scale_fill_manual(
            name = "",
            labels = c(lab.var),
            guide = elip.fill.leg,
            values = alpha(c(col.fill),alpha.fil.var))
        }
      } + {
        if (exists("sub.scale.opt") == FALSE) {sub.scale.opt = FALSE}
        if (sub.scale.opt == TRUE) {
          scale_color_viridis(aes(axis.data.X, axis.data.Y, color = sub.scale.var
), option = "D", alpha = 1, discrete = sub.scale.dis)
        } else {
          scale_colour_manual(
            name = "",
            guide = "legend",
            labels = c(lab.var),
            values = alpha(c(col.fill),alpha.col.var))
        }
      }
  }
  # STRESS LABEL - ADD OPT. FOR LOCATION + COLLISION DETECTION
  if (exists("stress.opt") == FALSE) {stress.opt = FALSE}
  if (stress.opt == TRUE) {
    range.x <- ggplot_build(plot.out.PCoA)$layout$panel_ranges[[1]]$x.range
```

```
    range.y <- ggplot_build(plot.out.PCoA)$layout$panel_ranges[[1]]$y.range
    sts.anot <- list()
    sts.anot[[1]] <-  annotate("label",
                               x = min(range.x),
                               y = max(range.y),
                               hjust = 0.1,
                               alpha = .4,
                               fontface = "bold",
                               fill = "white",
                               label = paste0("Stress = ", stress.var))
    plot.out.PCoA <- plot.out.PCoA + sts.anot
  }
  #LIMIT EXTRACTION FOR DENSITY PLOT CONFIGURATION
  #DENSITY PLOTS
  if (exists("axis.density.opt") == FALSE) {axis.density.opt = FALSE}
  if (axis.density.opt == TRUE){
    range.x <- ggplot_build(plot.out.PCoA)$layout$panel_ranges[[1]]$x.range
    range.y <- ggplot_build(plot.out.PCoA)$layout$panel_ranges[[1]]$y.range
    if (exists("axis.density.grp") == FALSE) {axis.density.grp <- grouping.var.ma
in}
    if (exists("axis.density.col") == FALSE) {axis.density.col <- col.fill}
    if (exists("axis.density.X") == FALSE) {axis.density.X <- axis.data.X}
    if (exists("axis.density.Y") == FALSE) {axis.density.Y <- axis.data.Y}
    xdensity <- ggplot(GG.data.in, aes(axis.density.X, y = ..scaled.., fill = axi
s.density.grp)) +
      geom_density(alpha=.5) +
      scale_fill_manual(values = axis.density.col) +
      scale_x_continuous(limits = range.x, expand = c(0,0)) + themeBW.var + theme
.var +
      scale_y_continuous(breaks = c(0,0.5,1.0),
                         labels = c("     ","   0.5","   1.0")) +
      theme(legend.position = "none",
            axis.text.x = element_blank(),
            axis.title.y = element_text(size = rel(0.8)),
            panel.grid.minor = element_blank(), panel.grid.major = element_line(s
ize = rel(.5))) +
      labs(x = NULL, y = "Density")
    ydensity <- ggplot(GG.data.in, aes(axis.density.Y, y = ..scaled.., fill = axi
s.density.grp)) +
      geom_density(alpha=.5) +
      scale_fill_manual(values = c(axis.density.col)) +
```

```
      scale_x_continuous(limits = range.y, expand = c(0,0)) + themeBW.var + theme
.var +
      scale_y_continuous(breaks = c(0.0,0.5,1.0),
                         labels = c("","0.5","1.0")) +
      labs(x = NULL, y = "Density") +
      theme(legend.position = "none",
            #TMP
            axis.text.y = element_blank(),
            axis.title.x = element_text(size = rel(0.8)),
            panel.grid.minor = element_blank(), panel.grid.major = element_line(s
ize = rel(.5))) +
      coord_flip()
    blankPlot <- ggplot()+geom_blank(aes(1,1))+
      theme(plot.background = element_blank(),
            panel.grid.major = element_blank(),
            panel.grid.minor = element_blank(),
            panel.border = element_blank(),
            panel.background = element_blank(),
            axis.title.x = element_blank(),
            axis.title.y = element_blank(),
            axis.text.x = element_blank(),
            axis.text.y = element_blank(),
            axis.ticks = element_blank(),
            axis.line = element_blank())
    if (is.list(guide.var) == TRUE){
      if (is.null(guide.var) == FALSE){
        main.legend <- get_legend(plot.out.PCoA)
      }
    }
    main.plot <- plot.out.PCoA + theme(legend.position = "none")
    main.title <- ggdraw() + draw_label(tit.var, fontface = "bold", size = rel(20
))
    if (exists("fig.config") == FALSE){fig.config <- "traditional"}
    if (fig.config == "traditional") {
      panels.internal <- plot_grid(ncol = 2,
                                   nrow = 2,
                                   main.plot + theme(plot.title = element_blank()
),
                                   ydensity, xdensity, NULL, rel_heights = c(1,0.
25,1), rel_widths = c(1,0.25,1))
      if (is.list(guide.var) == FALSE){
```

```
        if (is.null(guide.var) == TRUE) {
          panels.external <- panels.internal
        }
      }
      if (is.list(guide.var) == TRUE){
        if (is.null(guide.var) == FALSE) {
          panels.external <- plot_grid(main.legend, panels.internal, rel_widths =
 c(.1,1))
        }
      }
    }
    if (fig.config == "neat") {
      if (is.list(guide.var) == FALSE){
        if (is.null(guide.var) == TRUE) {
          panels.internal <- plot_grid(ncol = 2,
                                       nrow = 2,
                                       xdensity,
                                       blankPlot,
                                       main.plot + theme(plot.title = element_bla
nk()),
                                       ydensity,
                                       rel_heights = c(0.9,0.2), rel_widths = c(0
.9,0.2),
 scale = c(1,1,1,20))
          panels.external <- panels.internal
        }
      }
      if (is.list(guide.var) == TRUE) {
        if (is.null(guide.var) == FALSE) {
          panels.internal <- plot_grid(ncol = 2,
                                       nrow = 2,
                                       xdensity + theme(plot.margin = unit(c(0.5,
0,0,0),"lines")),
                                       main.legend,
                                       main.plot + theme(plot.title = element_bla
nk(),
                                                         plot.margin = unit(c(0,0
,0,0),"cm")),
                                       ydensity + theme(plot.margin = unit(c(0,0.
5,0,0),"lines")),
```

```
                                        rel_heights = c(0.2,1.0), rel_widths = c(1
.0,0.2), scale = c(1,1,1,1))
          panels.external <- panels.internal
        }
      }
    }
    plot.out.d <- plot_grid(main.title, panels.external, ncol = 1, rel_heights =
c(0.1, 1))
    plot.out.PCoA <- plot.out.d
  }
  return(plot.out.PCoA)
}fdatish
```