scientific reports

Check for updates

OPEN

Temporal stability and lack of variance in microbiome composition and functionality in fit recreational athletes

Ciara M. O' Donovan[1,2,3], Sai Ravi Chandra Nori[1,2], Fergus Shanahan[1,4], Gerardina Celentano[5], Thomas Brendan Murphy[5,6], Paul D. Cotter[1,2,6] & Orla O' Sullivan[1,2,6]✉

Human gut microbiome composition and function is influenced by environmental and lifestyle factors, including exercise and fitness. We studied the composition and functionality of the faecal microbiome of recreational (non-elite) runners ($n = 62$) with serial shotgun metagenomics, at 4 time points over a 7-week period. Gut microbiome composition and function was stable over time. Grouping of samples on the basis of their fitness level (fair, good, excellent, and superior) or habitual training (low (4–6 h/week), medium (7–9 h/week), high (10–12 h/week), and extreme (13 + hours/week)) revealed no significant microbiome-related differences. Overall, the species *Faecalibacterium prausnitzii*, *Blautia wexlerae*, and *Prevotella copri* were the most abundant members of the gut microbiome. Analysis of co-abundance groups (CAGs) revealed no significant relationship between CAGs and fitness levels or training subgroups. Functional pathways were similar across all samples and timepoints with no clustering based on associated metadata. The most abundant genes identified within samples corresponded to pathways for nucleoside and nucleotide biosynthesis, amino acid biosynthesis, and cell wall biosynthesis. Collectively, these results describe the microbiome of active recreational runners and note temporal stability amongst participants.

Keywords Athletes, Fitness, Faecal microbiome, Running

The composition and functional capacity of a healthy individual's gut microbiome is relatively stable[1–4]. However, environmental and lifestyle factors, including exercise[2], continually shape the microbiome. We and others have reported that the gut microbiome of athletes differs from that of sedentary individuals[5–7]. An increased abundance of specific taxa, such as *Veillonella*, *Akkermansia*, *Faecalibacterium prausnitzii*, or *Ruminococcus*, has been noted among some athlete cohorts[5–7]. In addition, particular taxa, including *Coprococcus* and *Veillonella*, have been detected at higher relative abundance following specific sporting endeavours[8,9], while exercise load was noted to be positively associated with *Prevotella*abundance among a cycling cohort[10]. A recent population study have also found a positive association between *Prevotella copri*and moderate intensity activity but not with vigorous activity[11]. In predominantly sedentary individuals, exercise-based interventions have yielded varied results in relation to the gut microbiome[12–14], in part because the exercise intervention may have been of insufficient duration. Longer periods of increased exercise have been associated with changes in microbiome profile[15].

Most studies of exercise and the human gut microbiome have been completed in elite athletes, with less known about casual athletes, although the latter may have more relevance to the general population. Here, we assessed the gut microbiome profile of recreational runners to determine the temporal stability of microbiome composition and function and whether there is variance at different levels of fitness as established by VO2 max measurements or self-reported training load; allowing us to identify if intense training will alter the gut microbiome at an individual level.

[1]APC Microbiome Ireland, Cork, Ireland. [2]Teagasc Food Research Centre, Moorepark, Cork, Ireland. [3]School of Microbiology, University College Cork, Cork, Ireland. [4]Department of Medicine, University College Cork, Cork, Ireland. [5]School of Mathematics and Statistics, University College Dublin, Dublin, Ireland. [6]VistaMilk, Fermoy, Cork, Ireland. ✉email: orla.osullivan@teagasc.ie

Methods

Ethical approval and recruitment

The samples used in this study were collected in accordance with the ethical principles set forth in the current version of the Declaration of Helsinki and the International Conference on Harmonization E6 Good Clinical Practice (ICH-GCP). Approval was received from the Clinical Research Ethics Committee of the Cork Teaching Hospitals, Cork, Ireland and all volunteers provided written, informed consent.

Faecal samples were collected from 62 individuals (39 male, 23 female, mean age 34 ± 9 years), who participated in at least 4 h of endurance sports per week, with a minimum 1.5 h of the training being a running activity, over the course of 7 weeks (4 samples; supplementary Fig. 1). Participants were grouped according to their habitual training volume: low (4–6 h/week), medium (7–9 h/week), high (10–12 h/week), and extreme (13 + hours/week). Subjects were instructed to maintain their diet and lifestyle unchanged during the study, including maintenance of their usual training volume, and to refrain from consuming non-steroidal anti-inflammatory drugs or any products that may contain live microbes during the study (supplementary Table 1).

Maximal oxygen consumption test

An incremental test was performed to measure maximal oxygen consumption (VO_2 max). The test was performed on a motorised treadmill with an initial speed of 10 km h^{-1}, with an increase of 1.0 km h^{-1}every 2 min until subjects achieved volitional exhaustion[16]. During the entire test treadmill inclination was set at 1%[17]. Respiratory gases were measured with an online respiratory system and VO_2 max was determined as the highest oxygen uptake per minute measured during the test. Samples were classified based on their VO_2 max into 4 fitness level groups (fair, good, excellent, and superior) based on previous classifications (Table 1)[18,19].

Faecal sample collection and DNA extraction

Faecal samples were collected by participants in storage tubes. Samples were stored at −80°C before DNA extraction. Samples were homogenised and DNA extraction was completed using the repeated bead beating plus column (RBBC) method, as previously described[20]. Briefly, 0.25 g of faecal sample was added into a sterile 2 mL screw-cap tube containing 0.25 g of zirconia beads (0.25 g of 0.1 mm, 0.125 g of 1.5 mm, and a single 2.5 mm bead). Following the addition of 1 mL of lysis buffer (500 mM NaCl, 50 mM Tris-HCl, 50 mM EDTA, and 4% sodium dodecyl sulphate (SDS)) samples were homogenized for 3 min on a Mini Beadbeater (BioSpec, Bartlesville, USA). Samples were incubated at 70 °C for 15 min to lyse the cells. Following centrifugation (16,000 xg; 5 min; 4 °C) supernatant was removed and the bead beating steps were repeated, this time using 200 µL of lysis buffer. Supernatant from both bead beating steps was pooled and 10 M ammonium acetate was added. Samples were centrifuged (16,000 xg; 10 min; 4 °C) and resulting supernatant was treated with one volume of isopropanol. DNA was pelleted and washed with 70% ethanol before being dissolved in Tris-EDTA. DNA was RNase and proteinase K treated. DNA was finally washed with 100% ethanol and buffers AW1 and AW2 (QIAamp DNA stool mini kit, Qiagen, England) before being eluted in 200 µL AE buffer (Qiagen). DNA was stored at −20 °C before library preparation.

DNA library preparation and sequencing

Samples were prepared for Illumina NextSeq shotgun sequencing as per modified Nextera XT DNA library preparation protocol (Illumina, California, USA). Briefly, 0.2 ng μL^{-1} of sample DNA was tagmented for 7.5 min at 55 °C. Tagmented DNA was amplified and index adapters (i5 and i7) were added. Agencourt AMPure beads (Beckman Coulter, Clare, Ireland) were used to clean up prepared library DNA. DNA samples were diluted to equimolar concentrations before pooling. Prepared library was sequenced on the NextSeq platform using standard Illumina protocols at the Teagasc sequencing facility. The datasets generated during the current study are available in the ENA repository under accession number PRJEB46398.

	Age	Fair	Good	Excellent	Superior
Male	< 20	38.4–45.1	45.2–50.9	51.0–55.9	> 55.9
	20–29	36.5–42.4	42.5–46.4	46.5–52.4	> 52.4
	30–39	35.5–40.9	41.0–44.9	45.0–49.4	> 49.4
	40–49	33.6–38.9	39.0–43.7	43.8–48.0	> 48.0
	50–59	31.0–35.7	35.8–40.9	41.0–45.3	> 45.3
Female	< 20	31.0–34.9	35.0–38.9	39.0–41.9	> 41.9
	20–29	29.0–32.9	33.0–36.9	37.0–41.0	> 41.0
	30–39	27.0–31.4	31.5–35.6	35.7–40.0	> 40.0
	40–49	24.5–28.9	29.0–32.8	32.9–36.9	> 36.9
	50–59	22.8–26.9	27.0–31.4	31.5–35.7	> 35.7

Table 1. Classification of VO2 max groupings. Classification of VO2max based on sex and age. Adapted from Heyward and The Cooper Institute for Aerobics Research[17,18]. Only those ages and classifications which were used in this study are shown.

Bioinformatic analysis

Sequence read quality was summarized using the FastQC package to check quality of reads before quality filtering and trimming. Sequences were quality checked by first removing contaminating human derived reads using NCBI Best Match Tagger (BMTagger)[21]. Resulting reads were trimmed and poor quality and duplicate reads removed using KneadData[22]. After quality trimming and host removal samples were reduced to an average of 5 million reads per sample. Taxonomic classifications of trimmed reads were determined with Kraken 2 (version 2.1.3) using the standard database and abundances were then calculated using Bracken[23,24]. Functional profiling was completed using the Human Microbiome Project Unified Metabolic Analysis Network 3 (HUMAnN3; http://huttenhower.sph.harvard.edu/humann3)[25].

Analysis of resulting taxonomic and functional tables was completed in R (R Foundation for Statistical Computing, Vienna, 2012) (version 4.4.1). Alpha diversity indices, including Shannon and Inverse Simpson indices, and beta diversity, measured by the Bray-Curtis distance, were calculated in R using the vegan package. (version 2.6.8)[26]. All data was assessed for normality prior to any statistical analysis. Analysis of significant differences between groups was only completed for species, which had a relative abundance ≥ 0.1% in at least one sample, using the F1-LD-F1 non-parametric test[27]. All reported p values were adjusted for multiple comparisons using the Benjamini-Hochberg method. The significance threshold was set at 0.05. Data visualisations were completed with ggplot2 (version 3.5.1) in R. Species co-abundance groups (CAGs) were established through first completing pairwise correlations of species followed by the generation of a heatmap using the complexHeatmap package in R.

Results

Gut microbiome composition remains stable in fit individuals over the investigation period with species co-abundance groups identified within samples

The gut microbiome of 62 individuals was investigated at four time points over the course of seven weeks (Supplementary Fig. 1). Intra-individual alpha diversity measures were not significantly different over time at both a compositional and functional level. In addition, beta diversity did not vary significantly over time (Fig. 1, Supplementary Fig. 2).

The impact of training load and fitness, as determined by VO2 max classification (Table 1) was of particular interest within this cohort. Analysis of the first sample collected from all individuals (visit 2 or 3) revealed no variance in the gut microbiome on the basis of either of these factors. Additional analysis investigating all other metadata (sex, BMI, age, etc.) collected within this investigation (supplementary Table 2) did not identify any significant variation in taxa present as a function of the metadata measures available.

Variation was evident in microbial composition between samples at all time points with species composition tending to be individual specific (Supplementary Fig. 3). The sum of the top 20 most abundant species identified (of 1839 species classified to at least one sample in a relative abundance > 0.01%) accounted for a mean of 54% of the microbiome composition across all samples. Variation in the dominant species by individual was evident at all time points (PERMANOVA, $R^2 = 0.74, p = 0.001$) with *Faecalibacterium prausnitzii*, *Blautia wexlerae*, *Prevotella copri*, *Phocaeicola vulgatus*, *Bacteroides uniformis*, and *Roseburia intestinalis* being dominant in at least some samples (supplementary Fig. 4). Over the seven week period some shifts in the identity of the dominant species occurred in the majority (50/62) of individuals. Among individuals who had a change in the dominant species, 30% had exactly one shift, whereas the remainder had two (32%) or more shifts (38%). The analysis did not reveal a consistent pattern among individuals experiencing shifts in dominant species. Furthermore, metadata factors were not associated with these shifts, nor was there a specific species where such changes were more prevalent. *F. prausnitzii* was the only species which was identified across all samples at all time points.

Co-abundant groups (CAGs) of species were identified by correlating all species from the first samples collected from all individuals (supplementary Fig. 5). Species present within CAGs accounted for 82–95% of the relative abundance of all assigned species (Fig. 2.A). CAG 1 contained 31 species (supplementary Table 3); 15 of which belong to the *Bacteroides* genus. CAG 2 contained 25 species; 7 of which belong to the *Bifidobacterium* genus. CAG 3 contained 12 species; 10 of which belong to the *Faecalibacterium* genus. CAG 4 and 5 contained 15 and 19 species, respectively, with no discernible pattern in terms of the associated species. The proportion of individual CAGs were found to vary across samples within this investigation although some CAGs were more evenly dispersed across samples than others. Although CAGs 1,2 and 3 showed similar abundance across all samples, great variability was observed among CAGs 4 and 5. Clustering of individual samples by the abundance of summed CAGs was observed in a heatmap (Fig. 2.B), which seemed to be driven in particular by CAGs 1 and 2. Furthermore, we clustered the samples based on the CAG 1 and CAG 2 groups, which revealed a visual distinction between the two CAGs (Supplementary Fig. 6). The CAG 2-dominant samples appeared to be more likely from overweight individuals, but we failed to find a significant difference ($p = 0.2$) between BMI and the CAG 1–2 groups. Overall, No obvious clustering was observed based on the metadata available within this analysis (*PERMANOVA*, $p = 0.54$).

Gut microbiome functionality was identified as remarkably similar between active individuals

The functional profiles of all samples identified 464 individual pathways (not stratified by species of origin) that appeared in at least one sample at any one timepoint. Overall, the gut microbiome functionality across participants appeared highly similar, with no prominent clustering by training load, fitness, sex, age, or other metadata variables (*PERMANOVA*, $p = 0.59$). However, when we performed hierarchical clustering on the functional profiles, two main clusters (Cluster 1 and Cluster 2) were identified (Fig. 3.A). A chi-square test revealed a significant difference in the distribution of BMI categories between these clusters (X-squared = 9.35, $p = 0.02$), driven in part by a higher proportion of overweight individuals in Cluster 1 ($n = 17$) compared

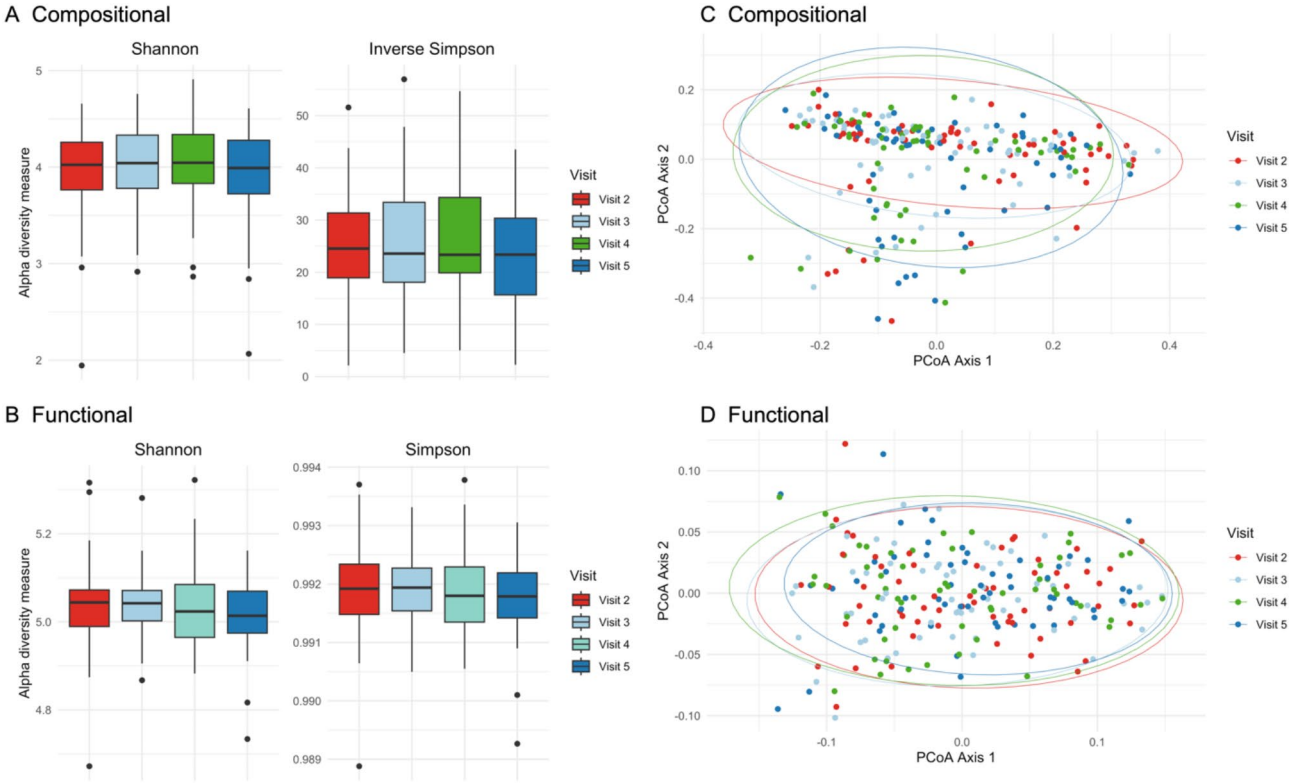

Fig. 1. Alpha and beta diversity over time. Diversity measures of gut microbiome (**A**) composition and (**B**) functional potential as determined using vegan in R coloured by visit (visit 2 ($N=61$), visit 3 ($N=58$), visit 4 ($N=59$), and visit 5 ($N=59$)). Species level abundance from Bracken following taxonomic classification with Kraken 2 were analysed with vegan to determine Shannon and Inverse Simpson diversities, for compositional data. Unstratified pathway abundance (reads per million) assignments as determined using HUMAnN3 were analysed with vegan to determine Shannon and Simpson diversities, for functional data. No significant differences were established in alpha diversity by visit. PCoA plots of beta diversity at (**C**) species level compositional profiles from Bracken following taxonomic classification with Kraken 2 and (**D**) HUMAnN3 profiles of functional potential, as determined using vegan in R, show the differences between faecal samples based on visit. Clustering of samples was not observed based on visit.

to Cluster 2 ($n=3$) (Supplementary Fig. 7). In contrast, we observed no significant differences in VO_2 max distributions between the two clusters.

We next examined whether specific functional pathways could help explain these two clusters. Among the 464 pathways, 150 showed significantly different abundances between Cluster 1 and Cluster 2 (Wilcoxon rank-sum test with Benjamini-Hochberg correction; Supplementary Table 4). Notably, however, none of these 150 pathways were significantly associated with the available metadata parameters (BMI category, VO_2 max, or other factors) when tested by Kruskal-Wallis or pairwise comparisons (all FDR-adjusted). Thus, although we detected a functional distinction between Cluster 1 and Cluster 2, no single metadata variable fully accounted for the pathways driving this separation.

Despite these cluster-level differences, the overall gut microbiome functionality across participants remained broadly similar. The top 10 most abundant pathways were PWY-1042, PWY-5941, PWY-5686, PWY-7790, PWY-7791, PWY-7221, PWY-7221, GLYCOGENSYNTH-PWY, VALSYN-PWY, and DTDPRHAMSYN-PWY. These pathways were predicted to be involved in nucleoside/nucleotide biosynthesis, amino acid biosynthesis, cell wall biosynthesis, energy production, and carbohydrate degradation (Fig. 3.B). Their relative proportions were consistent across individuals at the first sample collection time point and remained stable over multiple timepoints. Taken together, these results demonstrate that while the functional gut microbiome profiles can be grouped into two main clusters, one of which is enriched for overweight individuals but no individual functional pathway showed a strong correlation with BMI, VO_2 max, or other measured parameters.

Discussion

The results showed that the gut microbiome of fit recreational athletes is stable over a 7 week period and no significant differences in composition or function were observed in relation to fitness, training volume, or other metadata.

While a link between exercise and the gut microbiome in humans has been reported by many groups, many of the previous analyses identified differences in the microbiomes of athlete groups relative to controls, but studies

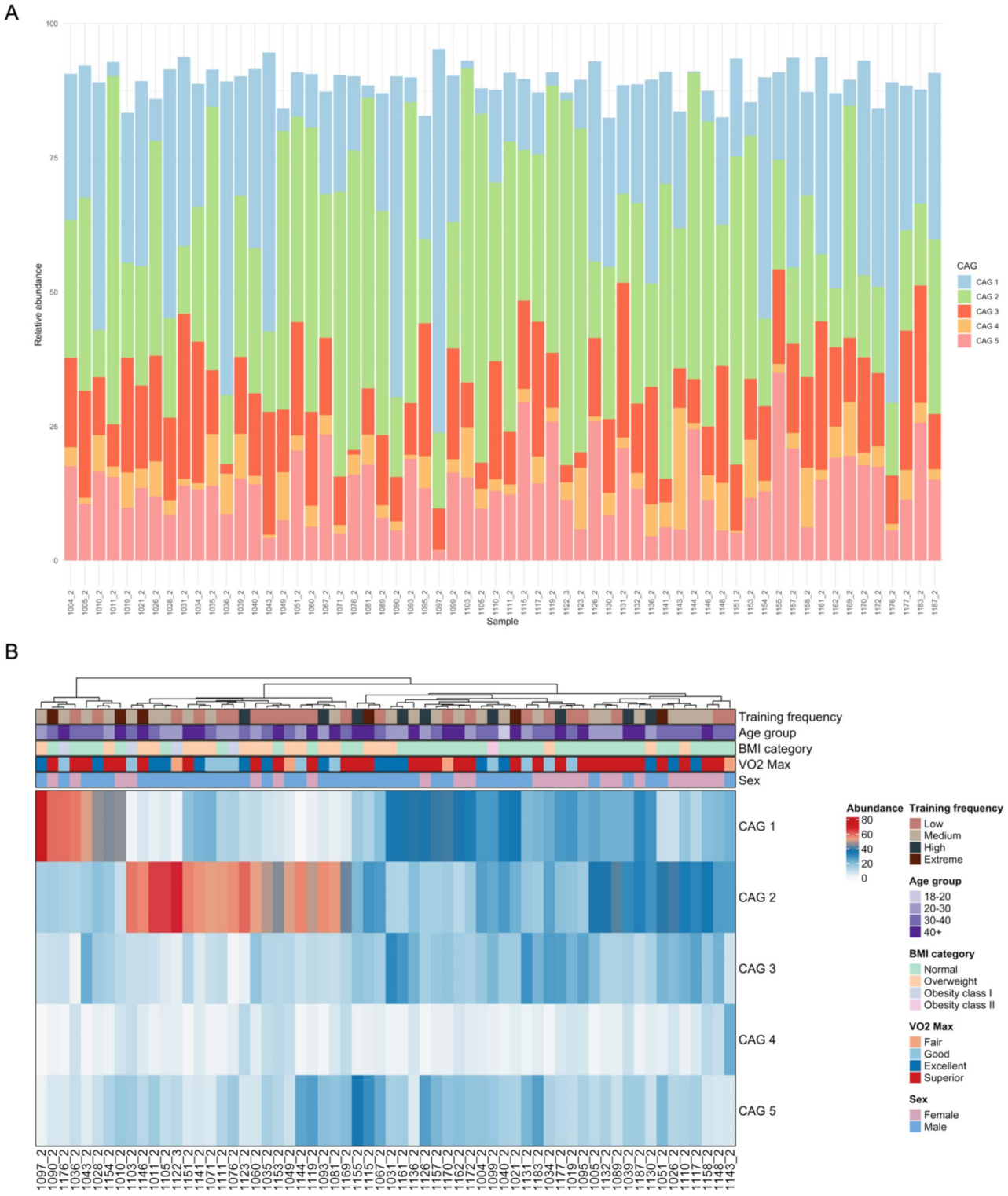

Fig. 2. Samples group by the presence of co-abundant species. Co-abundant species were identified as species whose relative abundance positively correlated using the first sample collected for all individuals (supplementary Fig. 5). (**A**) Relative abundance of summed species within each CAG by individual at the first sampling visit (visit 2 for all participants except 1122 with visit 3 being the first visit for this participant). Each colour represents a different CAG. (**B**) Heatmap of the abundances within each CAG at the first sampling visit showed no clustering by metadata available for samples.

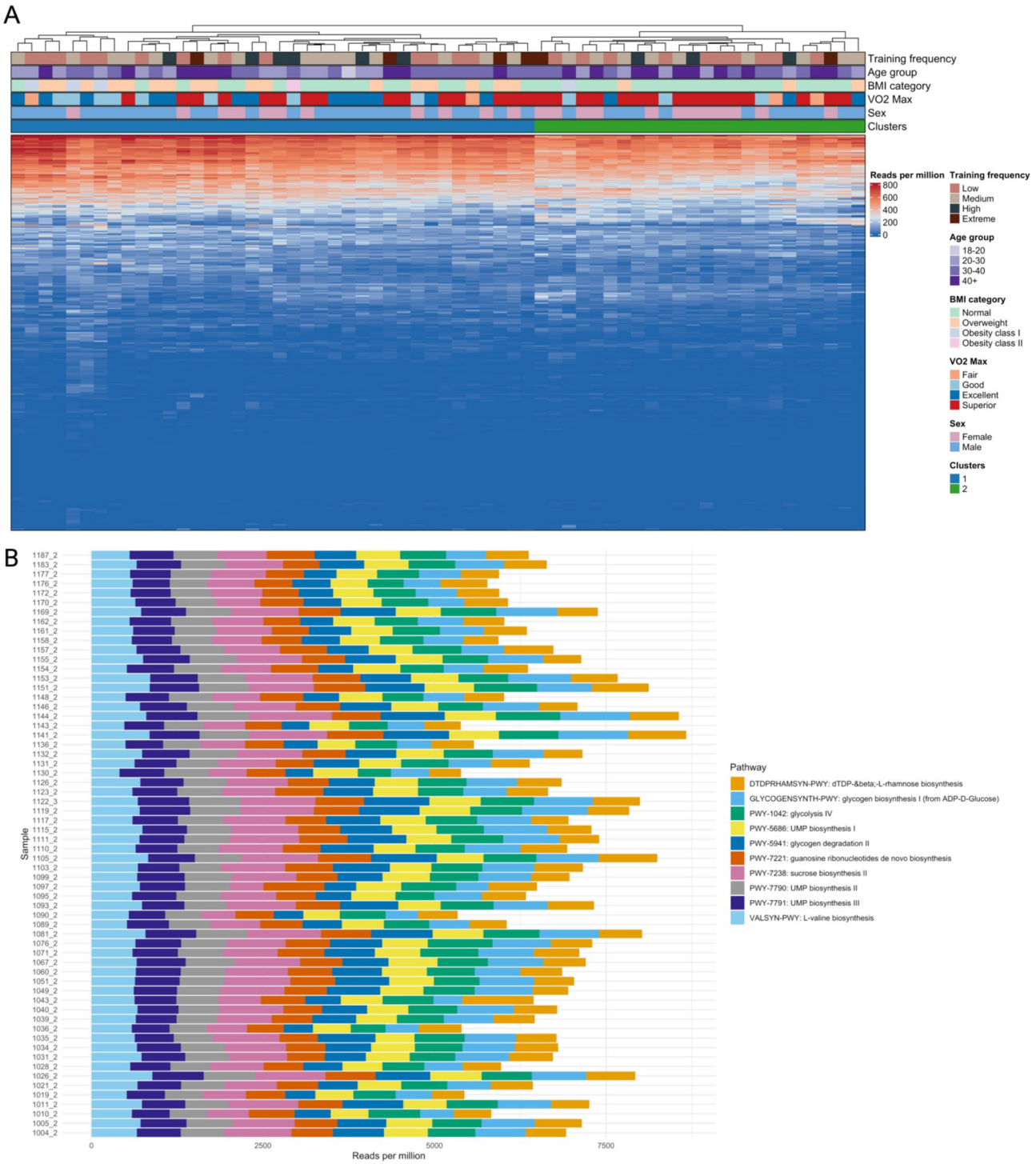

Fig. 3. Only subtle variance in microbial functional capacity was observed across samples. (**A**) Heatmap of all unstratified pathways (reads per million) as determined with HUMAnN3 within all samples identified little variation amongst samples and no obvious clustering by metadata available. **(B)** Top 10 most abundant pathways across samples at the first sample collection (Visit 2 for all samples except where visit 3 was used). Pathways were determined using HUMAnN3. No significant differences were determined between groups.

have varied widely with respect to duration, and intensity of the exercise/training[5,6,7,8,9,10,12,13,28,29]. Moreover, these investigations have largely been performed with elite athlete cohorts, where lifestyle is often highly controlled. Here, we investigated amateur athletes, i.e., fit individuals who are regularly active but whose lifestyle is not as tightly controlled as that of an elite athlete and are representative of a larger proportion of society. The findings of

our study underscore the consistent stability of the gut microbiome in active individuals over time, irrespective of variations in BMI, age, sex, or whether they engage in low or high training frequency.

The gut microbiome of these athletes collectively or when grouped according to fitness, training volume or other metadata, did not significantly change over the duration of the study, although some changes in the dominance of specific taxa within individuals over time were identified. Since many taxa within the gut microbiome share similar functional capacity the functional stability of the microbiota of the runners was very stable suggests that this change in dominant taxa over time is unlikely to have a major impact on overall microbiome activity.

It was established that the species that formed CAGs corresponded to a large proportion of the relative abundance each individual's microbiota. It has been previously proposed that CAGs may represent functional groups with similar functions[30]. In this investigation 5 CAGs were identified, 3 of which had dominant genera (*Bacteroides*, *Faecalibacterium*, and *Bifidobacterium*). Individuals were identified as clustering by the abundance of these CAGs, which is consistent with other investigations of the gut microbiome where in particular clusters of *Bacteroides* versus *Prevotella*dominant gut microbiomes have been identified[31,32], with additional clusters, including in some instances a *Bifidobacterium*cluster, being observed on some occasions[32]. These clusters have often been highlighted as reflective of diet. In this instance, these CAG groupings, which separated participants, was not associated with the dietary information collected for these individuals.

We were particularly interested in assessing the stability of participants gut microbiome over time within the context of training volume and fitness and we did not detect any significant shifts in microbial composition attributable to these variables. It is important to note that this was an observational rather than interventional study and so, in relation to training volume or other factors, studies that introduce a change over time may yield different results. Craven et al. found subtle changes to individual taxa with a change to training volume were noted[33]. A previous investigation has highlighted differences in gut microbiome in line with fitness[10], but among a quite different, cycling-based cohort. It is noteworthy that *Faecalibacterium prausnitzii*, the abundance of which has previously been positively associated with exercise levels[7,13,28], was found in all samples in the present study, but did not vary with fitness levels or training volume, perhaps because the participants in this study exceeded a minimum required level of physical activity.

From a functional perspective, samples were remarkably similar across individuals, which is consistent with the previous observations of functional redundancy within some human gut microbiomes[34]. Those functions that were identified as most abundant within the gut microbiome were those involved with nucleoside and nucleotide biosynthesis, amino acid biosynthesis, energy production, and carbohydrate degradation. Hierarchical clustering of the pathway profiles did reveal two main clusters differing in BMI distribution, although no single functional pathway showed a clear association with any metadata parameter, underscoring the complexity and redundancy of these active individuals gut microbiomes.

This study has limitations, including the absence of longitudinal dietary or data relating to dietary or exercise supporting supplements and intestinal transit time data. Future studies should focus on the investigation of these factors along with differences in the host and microbial metabolites produced by these individuals. The future longitudinal investigation of exercise implementation and continuation has the potential to address many of the questions left unanswered here. Nonetheless, the insights obtained here are important as they reflect an important cohort of society that are neither elite athletes nor those who do not exercise and indicate that the gut microbiome of these individuals is relatively stable over a somewhat acute period (seven weeks). By extension, it does not appear that tailored dietary approaches to improve or maintain the gut microbiome are required for individuals within this group, regardless of their fitness level or training intensity, at least in the absence of the severe impact of broad spectrum antibiotics or other events know to have a negative impact. Finally, the data presented here is a valuable addition to the growing wealth of information on the gut microbiome of fit individuals.

Data availability

The datasets generated during the current study are available in the ENA repository under accession number PRJEB46398.

Received: 10 October 2024; Accepted: 30 January 2025
Published online: 15 February 2025

References

1. Lozupone, C. A., Stombaugh, J. I., Gordon, J. I., Jansson, J. K. & Knight, R. Diversity, stability and resilience of the human gut microbiota. *Nature* **489**, 220–230 (2012).
2. Conlon, M. A. & Bird, A. R. The impact of Diet and Lifestyle on Gut Microbiota and Human Health. *Nutrients* **7**, 17–44 (2014).
3. David, L. A. et al. Host lifestyle affects human microbiota on daily timescales. *Genome Biol.* **15**, R89 (2014).
4. Shanahan, F., Ghosh, T. S. & O'Toole, P. W. The healthy Microbiome—what is the definition of a healthy gut Microbiome? *Gastroenterology* **160**, 483–494 (2021).
5. Barton, W. et al. 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 (2018).
6. Clarke, S. F. et al. Exercise and associated dietary extremes impact on gut microbial diversity. *Gut* **63**, 1913–1920 (2014).
7. Bressa, C. et al. Differences in gut microbiota profile between women with active lifestyle and sedentary women. *PLOS ONE*. **12**, e0171352 (2017).
8. Zhao, X. et al. Response of gut microbiota to Metabolite Changes Induced by endurance Exercise. *Front. Microbiol.* **9**, 765 (2018).
9. Scheiman, J. et al. Meta-omics analysis of elite athletes identifies a performance-enhancing microbe that functions via lactate metabolism. *Nat. Med.* **25**, 1104–1109 (2019).
10. Petersen, L. M. et al. Community characteristics of the gut microbiomes of competitive cyclists. *Microbiome* **5**, 98 (2017).

11. Baldanzi, G. et al. Accelerometer-based physical activity is associated with the gut microbiota in 8416 individuals in SCAPIS. *eBioMedicine* **100**, 104989 (2024).
12. Munukka, E. et al. Six-week endurance Exercise alters gut metagenome that is not reflected in systemic metabolism in over-weight women. *Front. Microbiol.* **9**, (2018).
13. Allen, J. M. et al. Exercise alters gut microbiota composition and function in lean and obese humans. *Med. Sci. Sports Exerc.* **50**, 747–757 (2018).
14. Cronin, O. et al. 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–e00018 (2018).
15. Barton, W. et al. The effects of sustained fitness improvement on the gut microbiome: A longitudinal, repeated measures case-study approach. 06.04.20046292 Preprint at (2020). https://doi.org/10.1101/2020.06.04.20046292 (2020).
16. Billat, L. V. & Koralsztein, J. P. Significance of the velocity at VO2max and time to exhaustion at this velocity. *Sports Med.* **22**, 90–108 (1996).
17. Jones, A. M. & Doust, J. H. A 1% treadmill grade most accurately reflects the energetic cost of outdoor running. *J. Sports Sci.* **14**, 321–327 (1996).
18. The Cooper Institute for Aerobics Research. The Physical Fitness Specialist Manual. (1997).
19. Heyward, V. H. *Advanced Fitness Assessment & Exercise Prescription* (Human Kinetics, 1998).
20. Yu, Z. & Morrison, M. Improved extraction of PCR-quality community DNA from digesta and fecal samples. *Biotechniques* **36**, 808–812 (2004).
21. Agarwala, R. & BMTagger Best Match Tagger for removing human reads from metagenomics datasets. (2011). https://ftp.ncbi.nlm.nih.gov/pub/agarwala/bmtagger/
22. Hsu, T. et al. Urban Transit System Microbial communities Differ by Surface Type and Interaction with humans and the Environment. *mSystems* **1** https://doi.org/10.1128/msystems.00018 (2016).
23. Wood, D. E. & Salzberg, S. L. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol.* **15**, R46 (2014).
24. Lu, J., Breitwieser, F. P., Thielen, P. & Salzberg, S. L. Bracken: estimating species abundance in metagenomics data. *PeerJ Comput. Sci.* **3**, e104 (2017).
25. Abubucker, S. et al. Metabolic reconstruction for metagenomic data and its application to the human microbiome. *PLoS Comput. Biol.* **8**, e1002358 (2012).
26. Oksanen, J. et al. vegan: Community Ecology Package. (2024).
27. Noguchi, K., Gel, Y. R., Brunner, E., Konietschke, F. & nparLD An R Software Package for the Nonparametric Analysis of Longitudinal Data in Factorial experiments. *J. Stat. Softw.* **50**, 1–23 (2012).
28. Jang, L. G. et al. The combination of sport and sport-specific diet is associated with characteristics of gut microbiota: an observational study. *J. Int. Soc. Sports Nutr.* **16**, 21 (2019).
29. Taniguchi, H. et al. Effects of short-term endurance exercise on gut microbiota in elderly men. *Physiol. Rep.* **6**, e13935 (2018).
30. Wu, G., Zhao, N., Zhang, C., Lam, Y. Y. & Zhao, L. Guild-based analysis for understanding gut microbiome in human health and diseases. *Genome Med.* **13**, 22 (2021).
31. Rinninella, E. et al. What is the healthy gut microbiota composition? A changing ecosystem across Age, Environment, Diet, and diseases. *Microorganisms* **7**, 14 (2019).
32. Gorvitovskaia, A., Holmes, S. P. & Huse, S. M. Interpreting Prevotella and Bacteroides as biomarkers of diet and lifestyle. *Microbiome* **4**, 15 (2016).
33. Craven, J. et al. The influence of exercise training volume alterations on the gut microbiome in highly-trained middle-distance runners. *Eur. J. Sport Sci.* **22**, 1222–1230 (2022).
34. Tian, L. et al. Deciphering functional redundancy in the human microbiome. *Nat. Commun.* **11**, 6217 (2020).

Acknowledgements

The authors thank Fiona Crispie and Laura Finnegan for their contributions to DNA sequencing. The authors are grateful to all those who donated their time, effort, and samples to this study. This research was conducted with the financial support of Science Foundation Ireland (SFI) under Grant Numbers SFI/12/RC/2273, 13/SIRG/2160, and 11/PI/1137 and Research Ireland under Grant number 22/FFP-A/10364.

Author contributions

F.S, P.C, and O.O.S conceived the study; C.O.D, S.R.C.N, G.C, T.B.M carried out sequencing and statistical analysis. All authors contributed to the writing of the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Data sharing statement

The datasets generated during the current study are available in the ENA repository under accession number PRJEB46398.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-88723-9.

Correspondence and requests for materials should be addressed to O.O.S.