CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	A multi-omics perspective on the biology and evolution of the genus Macrococcus
Authors	Mazhar, Shahneela
Publication date	2020-04-05
Original Citation	Mazhar, S. 2020. A multi-omics perspective on the biology and evolution of the genus Macrococcus. PhD Thesis, University College Cork.
Type of publication	Doctoral thesis
Rights	© 2020, Shahneela Mazhar. - https://creativecommons.org/licenses/by-nc-nd/4.0/
Download date	2026-10-31 08:03:16
Item downloaded from	https://hdl.handle.net/10468/11413



UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

Ollscoil na hEireann
National University of Ireland

Coláiste na hOllscoile, Corcaigh
School of Microbiology



A Multi-Omics Perspective on the Biology and Evolution of the Genus *Macrococcus*

Thesis presented by
Shahneela Mazhar, B.Sc.

Teagasc Food Research Centre, Moorepark, Fermoy, Co. Cork, Ireland
School of Microbiology, University College Cork, Cork, Ireland

This thesis is submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy in Microbiology

Under the supervision of
Head of School/Department: Prof. Paul O' Toole
Research Supervisor: Dr Olivia McAuliffe, Prof. Colin Hill

April 2020

Declaration

I hereby declare that the work submitted for assessment on the programme of study, leading to the award of PhD is entirely my own and has not been submitted to any other university or higher education institute, or for any other academic award in University College Cork. All external references and sources are clearly acknowledged and identified within the contents. I have read and understood the regulations of University College Cork concerning plagiarism.

Signature:

_____ Date _____

Shahneela Mazhar

This thesis is dedicated to my parents, the reason for who I am today

&

To my brothers, who are the core of my strength

Acknowledgement

First and foremost, I would like to thank ALLAH S.W.T for giving me the opportunity, determination, and strength to pursue my Ph.D. Without his blessing, this achievement would not have been possible.

Now I would like to thank and express my deep and sincere gratitude to Dr. Olivia McAuliffe for her continuous guidance, encouragement, and support throughout these four years. It has been a privilege to be a student of yours and have had an opportunity to learn from you. You have provided me with invaluable advice, suggestions, and inspiration that has greatly impacted my academic growth. I cannot thank you enough for also giving me the freedom to pursue the areas of my research while also ensuring that I stay on course and do not deviate from the core of my research. Without all your help and counsel, this thesis would not have been possible.

I would also like to thank my UCC supervisor Prof. Colin Hill for his input in my project work. I have great pleasure in acknowledging my gratitude for Prof. Seamus Fanning and his team in UCD, specifically Dr. Scott and my Microarray partner Yujie Hu for giving me the opportunity to work with you and to conduct my vital experimental work in your exceptional lab. Thank you for your valuable time and encouragement and for the compelling ideas that have profoundly aided in the completion of this thesis, and yes Yujie, for making me a “real microbiologist”.

I am very thankful to the Teagasc Walsh Fellowship scheme that has funded my research.

I am also very thankful to Dr. Kieran Kilcawley for his guidance and support regarding volatile analysis and David Mannion for aiding me with R software and statistics.

It would be improper to not mention the names of my senior colleagues who have been there for me since I started. Dr. Ewelina, Dr. Vincenzo, and Dr. Clara, you all have had an immense contribution towards building and developing me into the scientist I am today. Also, I would like to thank my new lab members Edel, Desiree, and Dr. Ashok, for the little talks and check-ups on my progress with writing. Thank you all.

My acknowledgment will be incomplete without thanking my best friend Hassan, who has stood by me for a decade. Your friendship means the world to me.

A special thanks to Aateef, you excellent DevOps engineer. Thank you for being there for me whenever I was in need and for all that scripting support.

I would also like to thank Dr. Shoaib, for those surprise visits, long walks, and chicken burgers, which will always be remembered.

Lastly but most importantly, I would like to thank my primary teachers, Mum and Dad. You have both been the epitome of love, strength, courage, and patience. I would not be where I am today without the sacrifices and the tough choices that you have had to make. No words in this world are enough to describe what you both mean to me. I hope I made you a little proud. I would love to thank my brothers, my little superheroes. You both are the core of all my strength, and this is for you.

Table of Contents

Declaration ..... **I**

Acknowledgement ..... **III**

Abstract ..... **XI**

Chapter 1 ..... **1**

Literature Review **1**

The Genus *Macrococcus*: An Insight into its Biology, Evolution and Relationship with *Staphylococcus* **1**

1.1 Abstract **2**

1.2 Introduction **3**

1.3 Defining the Genus *Macrococcus* **5**

1.3.1 Taxonomic History 5

1.3.2 Phylogeny 8

1.3.3 General Properties of the Genus *Macrococcus* 11

1.3.4 Approaches used to differentiate *Macrococcus* species 12

1.3.4.1 Phenotypic analysis 12

1.3.4.2 Cell Wall Composition 15

1.3.4.3 Fatty Acid Methyl Ester (FAME) - Gas Chromatograph (GC) Analysis 15

1.3.4.4 Genotyping and DNA Base Composition 16

1.3.4.5 Genotypic Methods for Intergenus Differentiation 16

1.3.4.6 Species-Specific Typing Methods 17

1.3.4.7 Strain-Specific Typing Methods 19

1.4 General Properties, Nomenclature, and Ecology of *Macrococcus* Species **21**

1.4.1 General Properties and Nomenclature 21

1.4.1.1 *Macrococcus caseolyticus* 21

1.4.1.2 *Macrococcus canis* 21

1.4.1.3 *Macrococcus equipercicus* 22

1.4.1.4 *Macrococcus bovicus* 22

1.4.1.5 *Macrococcus carouselicus* 23

1.4.1.6 *Macrococcus brunensis* 23

1.4.1.7 *Macrococcus hajekii* 23

1.4.1.8 *Macrococcus lamae* 24

1.4.2 Ecology 26

1.5 Comparative Genomics and Evolution..... 28
1.5.1Genome Comparison with Other Closely Related Organisms..... 28
1.5.2 Interspecies Differences in Metabolic Pathways *Staphylococcus* versus *Macrococcus*..... 31
1.5.2.1 Terminal Oxidases of Respiratory Chain..... 31
1.5.2.2 Starch Degradation and Glycogen Biosynthesis..... 34
1.5.2.3 Sugar Transport..... 34
1.5.2.4 Iron Acquisition..... 35
1.5.2.5 Vitamins: Biotin and Pantothenate Synthesis..... 35
1.6 Genomic Analysis of Virulence Genes and Regulatory Systems..... 37
1.6.1 Exotoxin Genes and Regulators..... 37
1.6.2 Two-Component Regulatory System..... 37
1.7Antibiotic Resistance..... 39
1.7.1 Mechanisms of β-lactams and Methicillin Resistance in Staphylococci and Macrococci..... 39
1.7.2 The Origin of Methicillin Resistance Gene *mecA*..... 42
1.7.3 The *mec* Complex in *Staphylococcus* and *Macrococcus* Species..... 42
1.7.4 Multidrug Resistance Gene *cfr*..... 46
1.7.5 Antibiotic Resistance Phenotype in *Macrococcus* Species..... 46
1.8 Occurrence of *M. caseolyticus* in Food Systems..... 48
1.9 Concluding Remarks..... 50
1.10 Acknowledgements..... 52
References..... 53
Chapter 2..... 69
A Rapid PCR-based Method to Discriminate *Macrococcus caseolyticus* and *Macrococcus canis* from closely-related *Staphylococcus* species based on the *cta*C Gene Sequence..... 69
2.1 Abstract..... 70
2.2 Introduction..... 71
2.3 Materials and Methods..... 72
2.3.1 Bacterial Strains and Culture Conditions..... 72
2.3.2 *M. caseolyticus* Isolation from Dairy and Non-Dairy Sources..... 72
2.3.3 Design of *ctaC*-Specific Primers..... 72
2.3.4 *ctaC*-Specific PCR Analysis of Reference Strains..... 74

2.3.5 *ctaC*-Specific PCR Screening of Complex Samples 74
2.3.6 16S rRNA Analysis 74
2.3.7 Gel Electrophoresis and Image Analysis 75
2.3.8 Pulsed Field Gel Electrophoresis (PFGE) 75
2.3.9 Phylogenetic Analysis of the Targeted *ctaC* Sequence 75
2.4 Results and Discussion 77
2.4.1 Design of *ctaC* Primers and *in silico* Specificity 77
2.4.4 Nucleotide Sequence Divergence of *ctaC* Fragments 81
2.5 Conclusion 82
2.6 Acknowledgements 83
2.7 References 84
Chapter 3 86
***Macrococcus linguae* sp. nov., Isolated from Bovine Tongue 86**
3.1 Abstract 87
3.2 Introduction 88
3.3 Materials and Methods 90
3.3.1 Bacterial Strains and Culture Conditions 90
3.3.2 Morphological, Physiological and Biochemical Characterization 90
3.3.3 Chemotaxonomic analysis 91
3.3.4 Genome sequencing 92
3.3.5 Phylogenetic Analysis 92
3.3.6 Sequence similarity calculations 93
3.3.7 Nucleotide sequence accession number 93
3.4 Results and Discussion 94
3.4.1 Phylogenetic Analysis 94
3.4.2 Genomic Characteristics of the Novel Isolate 98
3.3.3 Morphological and Biochemical Features 101
3.3.4 Chemotaxonomic Analysis 107
3.5 Conclusion 109
3.6 Description of *Macrococcus linguae* sp. nov. 110
3.7 Ethical Statement 111
3.8 Acknoledgements 112
3.9 References 113

Chapter 4 ..... **116**

A Proposal of *Macrcoccus canis* subsp. *bovinus* subsp. nov. and *Macrocccus goetzii* subsp. *corkensis* subsp. nov on the basis of Whole-Genome Analysis ... **116**

4.1 Abstract **117**

4.2 Introduction **118**

4.3 Materials and Methods **119**

4.3.1 Bacterial Strains and Culture Conditions 119

4.3.2 Phenotypic Characterization 121

4.3.3 Phylogenetic Analysis 121

4.3.4 Sequence Similarity Calculations 122

4.4 Results and Discussion **123**

4.4.1 Phylogenetic Analysis Separates the Proposed Subspecies into Two Groups 123

4.4.2 Phenotypic Characteristics Differentiating Subspecies 127

4.4.3 WGS- based Analysis Confirm *Macrococcus canis* and *Macrococcus goetzii* are Composed of Two Distinct Subspecies 132

4.5 Conclusion **135**

4.6 Acknowledgements **136**

4.7 Description of *Macrococcus canis* subsp. *canis* subsp. nov. **137**

4.8 Description of *Macrococcus canis* subsp. *bovinus* subsp. nov. **138**

4.9 Description of *Macrococcus goetzii* subsp. *goetzii* subsp. nov **139**

4.10 Description of *Macrococcus goetzii* subsp. *corkensis* subsp. nov **140**

4.11 References **141**

Chapter 5 ..... **144**

A Systems-wide Analysis of Proteolytic and Lipolytic Pathways Uncovers the Flavour-forming Potential of The Gram-Positive Bacterium *Macrococcus caseolyticus* subsp. *caseolyticus* **144**

5.1 Abstract **145**

5.2 Introduction **146**

5.3 Materials and Methods **149**

5.3.1 Bacterial Strains and Culture Conditions 149

5.3.2 Comparative Genome Analysis and Orthologous Groups Identification ... 149

5.3.3 Qualitative Analysis of Proteolytic and Lipolytic Activity 150

5.3.4 Quantitative analysis of the enzymes involved in proteolysis 150

5.3.4.1 Determination of CEP Activity 150
5.3.4.2 Preparation of Cell Free Extract 151
5.3.4.3 Determination of Aminopeptidases, Aminotransferases and Glutamate-Dehydrogenase Activities 151
5.3.5 Milk Protein Hydrolysis and Free Amino Acid analysis 152
5.3.6 Nitrogen Substrate Utilisation with Biolog Phenotypic Microarray PM3 Plate 153
5.3.7 Quantitative Analysis of Enzymes Involved in Lipolysis 153
5.3.8 Semi-Quantitative Assay for Hydrolytic Activities (API ZYM) 154
5.3.9 Volatile Compounds Analysis by Gas Chromatography-Mass Spectrometry (GC-MS) 154
5.3.10 Satistical analysis 155
5.4 Results 156
5.4.1 High CEP activity in *M. caseolyticus* subsp. *caseolyticus* strains is associated with a dairy origin 156
5.4.2 Limited Downstream Proteolytic Enzyme Activity in *M. caseolyticus* subsps. *caseolyticus* strains in Comparison to LAB Flavour Formers 161
5.4.3 High Esterase Activity Observed in strains with High CEP Activity 168
5.4.4 Comparative Genomics Reveals the Conservation of the Proteolytic and Lipolytic System Components in *M. caseolyticus* subsp. *caseolyticus* strains 171
5.4.5 Metabolomic Analysis with GC-MS Separates *M. caseolyticus* subsp. *caseolyticus* in to Two Distinct Groups Based on High/Low CEP and Esterase Activity 177
5.5 Discussion 185
5.6 Conclusion 191
5.7 Acknowledgements 192
5.8 References 193
Chapter 6 222
Comparative Genomics, Phylogenetic, and Metabolic Evaluation Reveals Considerable Genomic Diversity within the Genus *Macrococcus* 222
6.1 Abstract 223
6.2 Introduction 224
6.3 Materials and Methods 226
6.3.1 Bacterial Strains 226
6.3.2 Metabolic Profiling 226

6.3.3 Assessment of Antibiotic Resistance 226
6.3.4 Phylogenomics Analysis 230
6.3.5 Computational Analysis 230
6.3.6 Construction of Pan/Core- genome Families and Unique Genes 230
6.3.7 Identification of Prophages, Genomic Islands, CRISPR-Cas and Secondary Metabolites 231
6.3.8 Identification of Antibiotic Resistance and Virulence Genes 231
6.4 Results and Discussion 234
6.4.1 Comparison of *Macrococcus* and *Staphylococcus* Reveal Genus-level Genomic Differences 234
6.4.2 The *Macrococcus* Genus has a Large Open Pangenome with a Small Core Genome 236
6.4.3 Species of the Genus *Macrococcus* species Fall into Two Distinct Groups based on GC Content 241
6.4.4 WGS-based Analysis of AMR Reveals Extrinsic Resistome of *Macrococcus* Species 248
6.4.5 Phenotypic Antibiotic Resistance correlates with the Presence of AMR Genes Identified in Specific Macrococcal Strains 253
6.4.6 Diverse Distribution of Virulence Factors Identified Across *Macrococcus* Species 261
6.4.7 Evidence of Horizontal Gene Transfer in *Macrococcus* species 264
6.4.8 Genomic Islands, Transposase and Prophages Constitute a Major Portion of the Accessory Genetic Material Existent within the *Macrococcus* Genus 264
6.4.9 A wide distribution of *Com* operon across the *Macrococcus* genomes 268
6.4.10 Biosynthetic gene clusters conferring selective advantage in macrococci 268
6.4.11 Genus-wide Distribution of CRISPR-Cas and Restriction Modification Systems Suggests a Need to Resist Foreign DNA 272
6.5 Conclusion 275
6.6 Acknowledgements 276
6.7 References 277
Chapter 7 284
General Discussion 284
7.1 General Discussion 285
7.2 References 293
Appendix 1 296

Draft Genome Sequences of the Type Strains of Six *Macrococcus* Species 296

8.1 Abstract 297

8.2 Genome Announcement 298

8.3 Acknowledgements 300

8.4 References 301

Appendix 2 304

Draft Genome Sequences of *Macrococcus caseolyticus*, *Macrococcus canis*, *Macrococcus bohemicus*, and *Macrococcus goetzii* 304

9.1Abstract 305

9.2 Genome Announcement 306

9.3 Acknowledgments 308

9.4 References 309

Abstract

The Gram-positive genus *Macrococcus* currently comprises eleven species; *Macrococcus bovicus*, *Macrococcus carouselicus*, *Macrococcus equipercicus*, *Macrococcus brunensis*, *Macrococcus hajekii*, *Macrococcus lamae*, *Macrococcus goetzii*, *Macrococcus bohemicus*, *Macrococcus epidermidis*, *Macrococcus canis*, and *Macrococcus caseolyticus* that are commonly isolated from animal skin and their meat and milk products. The species of this genus are evolutionarily closely related to the species of *Staphylococcus*. In contrast to staphylococcal species, macrococci are generally thought to be avirulent. However, new evidence suggests the emergence of the species *M. caseolyticus* as a potential new pathogen. Additionally, genomic insights into the genomes of some members of this genus demonstrate the potential for these organisms to serve as reservoirs of methicillin resistance determinants. Considering that strains of the *M. caseolyticus* species have been associated with flavour development in dairy and fermented meat products, further investigation of this fascinating genus, and the species within it, is warranted. Therefore, this thesis aimed to explore genus-wide genomic diversity and distribution of antimicrobial resistance along with examining the role of *M. caseolyticus* in the flavour development of fermented dairy products, paying particular attention to the safety of this organism in the food system.

During this thesis work, a rapid PCR-based method was developed to isolate and establish a strain collection of *M. caseolyticus* and *M. canis* from diverse sources. The 13 strains isolated in this study were investigated using phylogenetic analysis based on partial 16S rRNA and ctaC gene, revealed DPC7161^T strain to form a separate branch from all the known members of the genus. A polyphasic approach based on the combination of genomic, phenotypic, and chemotaxonomic characteristics were used to identify the taxonomic position of DPC7161^T. The results from this analysis suggested that strain DPC7161^T represents a novel species of the genus *Macrococcus* with the name *Macrococcus linguae* sp. nov, proposed. Further comparative genomics analysis suggested that *M. goetzii* and *M. canis* species are composed of two distinct subspecies: *M. canis* subsp. *canis* subsp. nov., *M. canis* subsp. *bovinus* subsp. nov., *M. goetzii* subsp. *goetzii* subsp. nov, and *M. goetzii* subsp. *corkensis* subsp. nov. After a thorough taxonomic characterisation of strain collection, six strains belonging to the *M. caseolyticus* species were further investigated using a systems wide approach

integrating the enzymatic, metabolic and genomic data to unravel their flavour forming potential. The enzymatic analysis has identified highly active cell wall bound proteases driving extensive casein hydrolysis associated with dairy-derived strains of *M. caseolyticus.* Finally, pan-genomic analysis revealed considerable genetic diversity within the genus *Macrococcus*. The presence of antimicrobial resistance genes on genomic islands associated with mobile genetic elements suggests horizontal gene transfer as a key driver in the spread of antimicrobial resistance genes across the members of this genus. Overall, this thesis presents the genomic and metabolic diversity within the genus *Macrococcus* and the potential application of *M. caseolyticus* in flavour diversification.

Chapter 1

Literature Review

The Genus *Macrococcus*: An Insight into its Biology, Evolution and Relationship with *Staphylococcus*

This entire chapter has been published in Advances in Applied Microbiology (2018). 105, 1-50.

DOI: 10.1016/bs.aambs.2018.05.002

1.1 Abstract

The Gram-positive genus *Macrococcus* is composed of eight species that are evolutionarily closely related to species of the *Staphylococcus* genus. In contrast to *Staphylococcus* species, species of *Macrococcus* are generally regarded to be avirulent in their animal hosts. Recent reports on *Macrococcus* have focused on the presence of novel methicillin resistance genes in *Macrococcus caseolyticus* and *Macrococcus canis*, with the discovery of the first plasmid-encoded methicillin resistance gene in clinical *Staphylococcus aureus* of probable macrococcal origin generating further interest in these organisms. Furthermore, *M. caseolyticus* has been associated with flavour development in certain fermented foods and its potential as a food bio-preservative has been documented. The potential application of these organisms in food seems at odds with the emerging information regarding antibiotic resistance and is prompting further examination of the potential safety issues associated with such strains, given the European Food Safety Authority framework for the safety evaluation of micro-organisms in the food chain. A comprehensive understanding of the genus would also contribute to understanding the evolution of staphylococci in terms of its acquisition of antibiotic resistance and pathogenic potential. In this review, we discuss the current knowledge on *Macrococcus* with regard to their phenotypic capabilities, genetic diversity, and evolutionary history with *Staphylococcus.* Comparative genomics of the sequenced *Macrococcus* species will be discussed, providing insight into their unique metabolic features and the genetic structures carrying methicillin resistance. An in-depth understanding of these antibiotic resistance determinants can open the possibilities for devising better preventative strategies for an unpredictable future.

Keywords: *Macrococcus,* methicillin resistance, evolution, *Staphylococcus*, comparative genomics.

1.2 Introduction

Macrococcus are Gram positive, coagulase negative, catalase and oxidase positive cocci, that belong to the family of *Staphylococcaceae* (Götz et al., 2006). At present, eight species exist in the *Macrococcus* genus: *Macrococcus caseolyticus, Macrococcus bovicus, Macrococcus carouselicus, Macrococcus equipercicus, Macrococcus brunensis, Macrococcus hajekii, Macrococcus lamae* and *Macrococcus canis* (Brawand et al., 2017). Early studies proposed various classifications for species belonging to this genus and it was not until 1998, when a polyphasic approach was used to propose the genus *Macrococcus* (Kloos et al., 1998a).

Members of the *Macrococcus* genus are disseminated in nature as animal commensals that are phylogenetically most closely related to the genus *Staphylococcus* (Baba et al., 2009; Hiramatsu et al., 2014). Phylogenetic analysis based on 16S rRNA sequences indicate that the two genera have the same ancestor and are estimated to have diverged from each other over 200 million years ago (Hiramatsu et al., 2014; Lory, 2014). At present, there is limited knowledge available regarding the ecological distribution of *Macrococcus* across mammalian animals, as only a few reports have identified host preference and isolated *Macrococcus* species from the skin samples of animal and their meat and milk products (Kloos et al., 1998a; Mannerová et al., 2003; Wu et al., 2009; Brawand et al., 2017; Schwendener et al., 2017). In contrast to *Staphylococcus* species, *Macrococcus* are not known to cause any human or animal diseases (Baba et al., 2009; Lory, 2014). However, there has been a case where *M. caseolyticus* was isolated from the abscesses of slaughtered lambs, and in a few cases, *M. canis* and *M. caseolyticus* species have been isolated from infection sites in dogs (De la Fuente et al., 1992; Gomez-Sanz et al., 2015; Brawand et al., 2017; Cotting et al., 2017).

The completed or draft genomes of three macrococci are available in the NCBI database (www.ncbi.nlm.nih.gov/genome). The genome analysis of one of the *M. caseolyticus* strains has focused on elucidating information regarding its genetic divergence from its sister genus *Staphylococcus*, along with the structural information of its unusual *mec* gene complex (Baba et al., 2009; Tsubakishita et al., 2010a; Schwendener et al., 2017). Recent genome analysis of *M. canis* KM45013^{T} has also indicated the ability of this species to acquire antibiotic resistance genes (Brawand et

al., 2018). Some studies have suggested that the methicillin resistance gene in pathogenic *Staphylococcus* species was acquired through horizontal gene transfer (HGT) from macrococcal species (Hiramatsu et al., 2013; Lory, 2014; Becker et al., 2018). In addition to methicillin resistance, the presence of other multi-drug resistant genes on plasmids in *M. caseolyticus* has also been reported (Wang et al., 2012b). *M. caseolyticus* has a documented presence in fermented foods which raises questions regarding its safety, as a potential transfer of these mobile elements to other commensal or pathogenic bacteria such as *Staphylococcus aureus*, may result in a non-negligible threat to human health (Randazzo et al., 2002; Wang et al., 2012b; Schwendener et al., 2017).

Given the potential application of these organisms in food and the emerging information regarding antibiotic resistance, a thorough review of the currently available information on *Macrococcus* is warranted. In this review, we will provide an overview of the genus *Macrococcus*, with a general description of the current species within the genus and a summary of the methodological approaches which contributed to classification of these species. Genomic features of the sequenced *Macrococcus* species are described, and comparisons to its sister genus *Staphylococcus* are discussed. The emerging information regarding the presence of methicillin resistance and other multidrug resistant genes in *M. caseolyticus* and *M. canis* from animal origin will be reviewed. Finally, the industrial exploitation of the *M. caseolyticus* and its presence and potential role in flavour development of fermented foods will also be reviewed.

1.3 Defining the Genus *Macrococcus*

1.3.1 Taxonomic History

Historically, there have been many classifications suggested for the Gram-positive and catalase-positive cocci that subsequently have been classified as the family *Staphylococcaceae*. In the 1870's, researchers first observed coccus-shaped bacteria associated with inflammation, abscesses and pus (Billroth, 1874; Ogston, 1881). Toxicogenic properties of these cocci were described in a classical paper by Ogston entitled "*Micrococcus* Poisoning" (Ogston, 1882). Ogston was first to introduce the name *Staphylococcus* by differentiating spherical bacteria into two types; those arranged in groups or clusters were called *Staphylococcus* and others in chains were referred to as *Streptococcus*. In 1884, Rosenbach proposed a formal taxonomical classification of the genus *Staphylococcus*, when he divided the genus into two species, *Staphylococcus aureus* and *Staphylococcus albus* (Rosenbach, 1884). Zopf then categorised mass-forming staphylococci and tetrad-forming micrococci in the genus *Micrococcus* (Zopf, 1885). In 1886, Flügge separated the two genera based on their action on gelatin. Staphylococci were observed to liquefy gelatin whereas micrococci were variable in their reaction (Flügge, 1886). *Staphylococcus* was then later placed into the family of *Micrococcaceae*, which also contained the genera *Micrococcus* and *Planococcus* (Stackebrandt and Woese, 1979). It was not until 1955, when tests for simple fermentation and oxidation activity were introduced to distinguish *Micrococcus* from *Staphylococcus*, and these approaches were applied by Evans, Bradford & Niven. Their results indicated that staphylococci produce acid anaerobically from glucose (Evans et al., 1955). However, these tests were not sufficient to differentiate weak anaerobic staphylococci from certain micrococci such as *Micrococcus varians*, that can also produce small amount of acid anaerobically from glucose (Schleifer et al., 1982a). With major advancements in DNA technology, base composition could be used to make a clear distinction between the two Gram-positive cocci. This demonstrated that members of the *Staphylococcus* genus have a lower GC content of 33–40%, whereas members of the *Micrococcus* genus had a higher GC content of DNA, around 63-73% (Silvestri and Hill, 1965). Further studies have demonstrated the differentiation of staphylococci from micrococci and other bacteria on the basis of their cell wall composition, cytochrome and menaquinone patterns, sensitivity to erythromycin, lysostaphin, bacitracin and furazolidone, polar

lipid composition, DNA-rRNA hybridization and comparative oligonucleotide cataloguing of 16S rRNA (Schleifer and Kandler, 1972a; Schleifer and Kloos, 1975; Faller et al., 1980; Kilpper et al., 1980; Collins and Jones, 1981; Ludwig et al., 1981; Falk and Guering, 1983; Baker, 1984; Nahaie et al., 1984).

The first strains of what is now known as the genus *Macrococcus* were isolated from raw milk in 1916 by Evan, and then placed in the *Micrococcus* genus, based on the morphological and carbohydrate decomposition characteristics (Evans, 1916). Later, comparative chemical, biochemical and nucleic acid hybridization analyses were performed on three *Micrococcus* strains including *Micrococcus caseolyticus* ATCC13548^T, which was then assigned to the genus *Staphylococcus* on the basis of its cell wall composition, low GC content of DNA and results of DNA-rRNA hybridization. However, strain ATCC13548^T differed from other known members of the staphylococci in terms of its production of class II fructose-1,6-diphosphate aldolase. DNA-DNA hybridization indicated that this strain was not closely related to any of the known *Staphylococcus* species. Therefore, this strain was proposed to represent a new species of the *Staphylococcus* genus and was referred to as *Staphylococcus caseolyticus* (Schleifer et al., 1982b). In 1998, *Staphylococcus caseolyticus* was reclassified into a newly proposed genus, *Macrococcus*, on the basis of its comparative 16S rRNA analysis, DNA-DNA hybridization, ribotype patterns, cell wall composition, and phenotypic characteristics (Kloos et al., 1998a).

At present, members of the *Macrococcus* genus can be distinguished from *Staphylococcus* on the basis of their larger cell size, higher GC content (38–45%), smaller genome size and significantly lower 16S rRNA sequence similarity. Macrococci are also oxidase positive whereas the majority of staphylococci are negative, with the exceptions of *Staphylococcus sciuri*, *Staphylococcus lentus*, and *Staphylococcus vitulinus* (Kloos et al., 1998a; Baba et al., 2009). A summary of the currently known *Macrococcus* species can be found in Table 1. The *Macrococcus* genus remains in the family of *Staphylococcaceae* and 16S rRNA sequence analysis indicates its closest relative to be *Staphylococcus*. In addition, *Macrococcus* genus also clusters with other Gram positive bacteria with low DNA GC content such as *Salinicoccus* and *Bacillus* (Kloos et al., 1998a; Götz et al., 2006).

.

Table 1: List of species in the genus *Macrococcus*

Species name	Source	Reference
Macrococcus caseolyticus	Whale skin, Cow's milk, Bovine tongue, Bovine heart, Food-processing factory	(Kloos et al., 1998a)[a] (Evans, 1916)
Macrococcus bovicus	Skin of Holstein cow, Irish thoroughbred horse, Morgan horse, Anglo-Trakehner horse	(Kloos et al., 1998a)[a]
Macrococcus equipercicus	Skin of Irish thoroughbred horse, Morgan horse and Shetland pony	(Kloos et al., 1998a)[a]
Macrococcus carouselicus	Skin of Irish thoroughbred horse, Morgan horse, Anglo-Trakehner horse, Shetland pony	(Kloos et al., 1998a)[a]
Macrococcus brunensis	Skin of llama's	(Mannerová et al., 2003)[a]
Macrococcus hajekii	Skin of llama's	(Mannerová et al., 2003)[a]
Macrococcus lamae	Skin of llama's	(Mannerová et al., 2003)[a]
Macrococcus canis	Canine Infection	(Brawand et al., 2017)[a]

[a] First description

1.3.2 Phylogeny

On the basis of comparative 16S rRNA gene sequence studies, the genus *Macrococcus* belongs to the phylum of *Firmicutes*. They are most closely related to *Staphylococcus* genus and other *Firmicutes* with low DNA GC content, including *Listeriaceae,* bacilli, enterococci, streptococci and lactobacilli (Fig. 1) (Vos et al., 2011). Taxonomically, the genus is placed in the family of *Staphylococcaceae*, which was proposed to combine the genera *Staphylococcus*, *Gemella*, *Macrococcus* and *Salinicoccus* (Garrity and Holt, 2001). Currently, the *Staphylococcaceae* family also includes species from the genera *Jeptgalicoccus* and *Nosocomiicoccus* (Kwok and Chow, 2003; Lory, 2014).

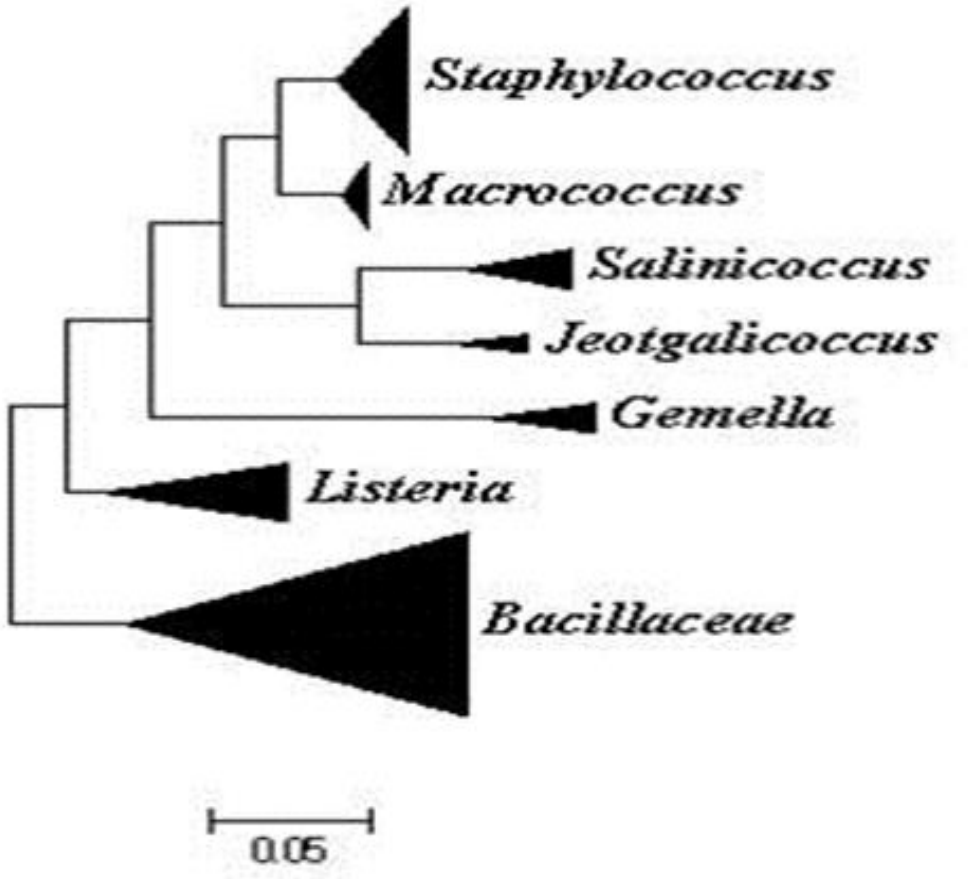


Fig.1. 16S rRNA-based tree reflecting the relationship of the *Staphylococcus*-*Macrococcus*- *Salinicoccus*- *Jeotgalicoccus* clusters and its sister groups amongst the bacilli. The tree was re-constructed using maximum likelihood method of all currently available almost complete small subunit rRNA sequences. Only sequence positions which shared identical nucleotides in at least 50% of all compared sequences from representative *Staphylococcaceae* and other bacilli were used to construct the tree. Length bar indicates 5% sequence divergence.

The genera *Staphylococcus* and *Macrococcus* are monophyletic (Fig. 2), with intergenera 16S rRNA sequence similarities of 93.4–95.3%. The intragenus 16S rRNA similarity of staphylococci is at least 96.5% and slightly higher for macrococci which is 97.7% (Vos et al., 2011). A phylogenetic study based on comparative sequence analysis of heat-shock proteins (*hsp60*) indicates clear separation of most *Staphylococcus* species from four of the eight *Macrococcus* species (*M. equipercicus, M. bovicus, M. carouselicus* and *M. caseolyticus*) However, these four *Macrococcus* species cluster closely to the *Staphylococcus sciuri* group, comprising of species that produce cytochrome c oxidase (*S. sciuri, S. lentus* and *S. vitulinus*). This clustering was also evident from the 16S rRNA sequence analysis performed in a number of studies (Kloos et al., 1998a; Lory, 2014). However, the *hsp60* sequence appeared to be more discriminatory than the 16S rRNA gene sequence for the identification and differentiation of staphylococcal and macrococcal species. Intragenus phylogenetic analysis performed using partial sequence of the *hsp60* gene resulted in sequence similarities of 82-87% for the four macrococcal species, while sequence similarities ranged from 74-98% amongst the 40 staphylococcal species examined. These values were significantly lower than the sequence similarities between the 16S rRNA gene sequence (Kwok and Chow, 2003).

DNA-DNA hybridization studies of the same four *Macrococcus* species indicated *M. equipercicus, M. bovicus, M. carouselicus* were more closely related to one another than *M. caseolyticus* (Kloos et al., 1998a). All of the studies based on DNA-DNA hybridization, partial *hsp60* gene and 16S rRNA sequence analysis indicate a closer relationship of the genus *Macrococcus* to *S. sciuri* species group than to staphylococcal species outside of this group. Particularly, *M. caseolyticus* and *M. canis* are more closely related to *Staphylococcus* species than the other *Macrococcus* species based on their 16S rRNA sequence similarities, and their significantly lower GC content (36.9-38%) that is shared by several staphylococcal species (Kloos et al., 1998b; Brawand et al., 2017).

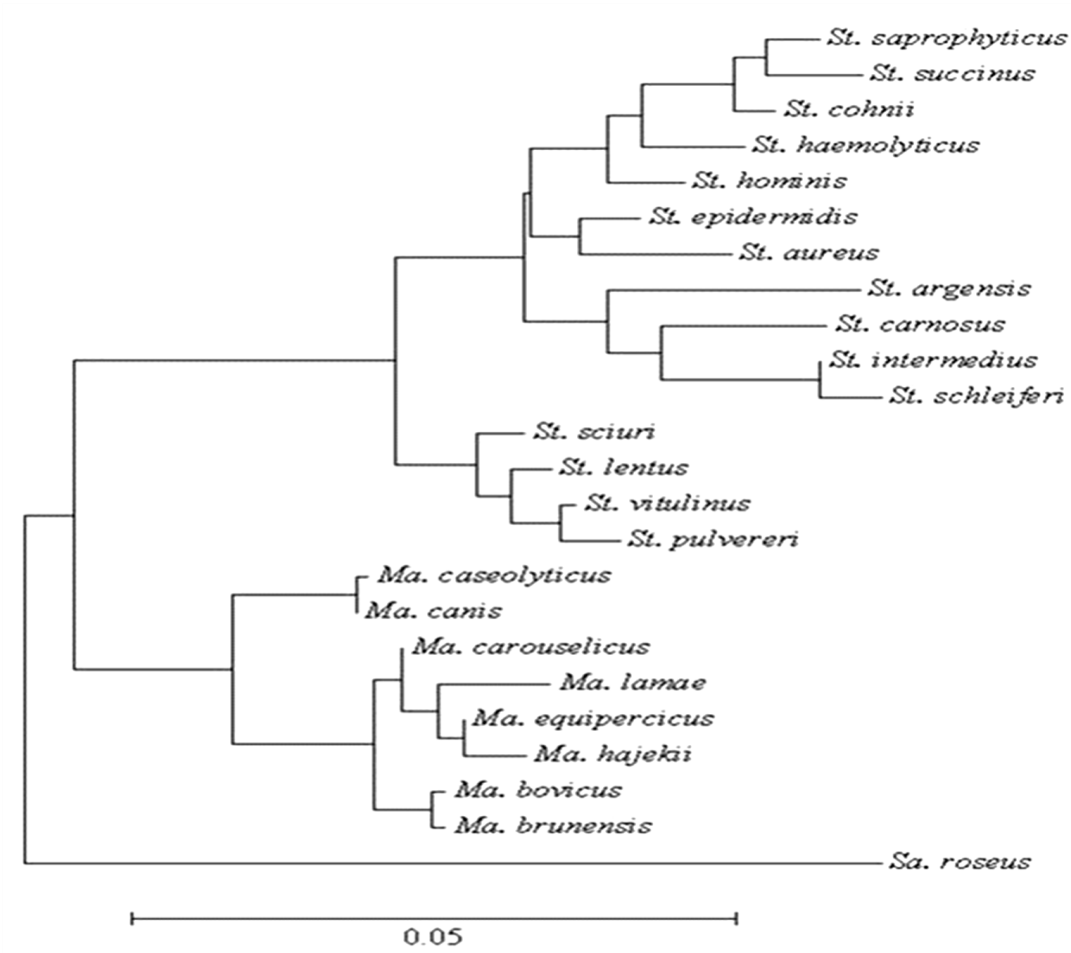


Fig.2. 16S rRNA-based tree reflecting the phylogenetic relationships of typed strains of *Macrococcus* species with the representative selection of typed strains of *Staphylococcus* and *Salinicoccus roseus* species. The latter species indicate the level of relatedness of the *Macrococcus*-*Staphylococcus* species. The tree is based on the results of maximum likelihood analysis from the output of multiple sequence alignment of the selected small subunit rRNA sequences. Substitution model Kaimura-2 applied to compute the distance formula. Only sequence positions which shared identical nucleotide in at least 50% of all included members of the staphylococci, macrococci and salinicocci were included for the tree calculation. The bar length indicates 5% estimated sequence divergence.

1.3.3 General Properties of the Genus *Macrococcus*

The name *Macrococcus* or "large coccus" was adopted, as the members of this genus have a large cell size compared to species of its sister genus *Staphylococcus* (cell size ≈ 2.5-4 times the diameter of *Staphylococcus aureus* cells when grown in Tryptic soy broth) (Kloos et al., 1998a). The bacteria are Gram positive, occurring mostly in pairs, tetrads or in irregular clusters, and occasionally as single cells or arranged in short chains (Kloos et al., 1998a; Götz et al., 2006). Cells are spherical or coccoid shaped (1.1-2.5 µm in diameter). They are generally unencapsulated, non-motile and non-sporeforming. With the exception of *M. caseolyticus*, macrococci do not contain teichoic acid in their cell wall. Lipoteichoic acid is present in *M. caseolyticus, M. carouselicus, M. equipercicus* and *M. boviscus*. Peptidoglycan types are Lys-Gly_{3-4}, L-Ser (*M. caseolyticus, M. carouselicus* and *M. equipercicus*) or Lys-Gly_3, L-Ser (*M. bovicus*). Macrococci are chemoorganotrophic and metabolism is predominantly respiratory. They can be marginally facultatively anaerobic, but do grow better under aerobic conditions. They are coagulase-negative and catalase- and cytochrome c oxidase-positive (Götz et al., 2006). They contain a-, b- and/or c type cytochromes whereas the majority of the staphylococci contain a- and b- type cytochromes, with the exception of the *S. sciuri* group (*S. sciuri, S. lentus* and *S. vitulinus*) that possess c type cytochromes. They are negative for ornithine decarboxylase, β-glucuronidase, arginine deiminase activities and for acid production from the fermentation of D-cellobiose, D-melezitose, D-raffinose, D-turanose, D-xylose, L-arabinose and salicin. Macrococcal species are generally more susceptible to a wide range of antibiotics in comparison to staphylococcal species, including penicillin G, erythromycin, clindamycin, tetracycline, ciprofloxacin, rifampin, trimethoprim-sulfamethoxazole, gentamicin, kanamycin, streptomycin, vancomycin and chloramphenicol (Kloos et al., 1998a). However, macrococci are resistant to bacitracin and lysozyme and susceptible to furazolidone (Schleifer, 2015). Genome size range from approximately 1.5-1.8 Mb (Kloos et al., 1998a). Whereas, the whole genome sequenced *Macrococcus* are reported in the range of 2.1-2.3 Mb, with average GC content of 36.9-45% (Kloos et al., 1998b; Mannerová et al., 2003; Baba et al., 2009; Brawand et al., 2017; Schwendener et al., 2017).

1.3.4 Approaches used to differentiate *Macrococcus* species

The classification and identification of species of the *Macrococcus* are based on a variety of comparative phenotypic, chemotaxonomic and genotypic analysis. DNA-DNA hybridization, DNA-rRNA hybridization, ribotyping and 16S rRNA sequencing methods have been exploited extensively to elucidate the phylogenetic position of *Macrococcus* and its differentiation from the *Staphylococcus* species (Schleifer et al., 1982a; Kloos et al., 1998a; Mannerová et al., 2003). Established conventional methods were utilised in the 1980's by Schleifer *et al*, to study characteristics at the cellular and population level (Schleifer et al., 1982a), including morphological and physiological properties (Baird-Parker, 1963), cell wall composition (Schleifer and Kandler, 1972b; Schleifer, 1973), menaquinone pattern (Jeffries et al., 1968), cytochrome pattern (Faller et al., 1980), enzyme reactions (Götz et al., 1979; Faller and Schleifer, 1981), sensitivity to lysostaphin (Schleifer and Kloos, 1975) among others. These phenotypic characteristics were combined with additional characteristics that were investigated by other studies and examined for their correlation with DNA-DNA hybridization, DNA base composition and 16S rRNA sequencing, fatty acid and methyl esters (FAME) analysis and ribotype and macrorestriction patterns (Schleifer et al., 1982b; Kloos et al., 1998a; Mannerová et al., 2003).

1.3.4.1 Phenotypic analysis

An organism's phenotype refers to observable traits that are a result of unique interactions between its genetic background and environmental factors (Cheng et al., 2016). Historically, taxonomic classification relied on phenotypic approaches for species delineation. According to Wayne *et al,* description of new species required a minimum number of phenotypic characteristics (Wayne et al., 1987). Today, delineating species using phenotype-based approach alone is not sufficient and, therefore, a polyphasic description including phenotypic, genotypic and chemotaxonomic properties are required to provide a comprehensive description and allow differentiation of species (Vandamme et al., 1996; George et al., 2005). In a number of studies, extensive conventional phenotypic analyses have been employed to identify and differentiate species of *Macrococcus* (Kloos et al., 1998a; Mannerová et al., 2003; Brawand et al., 2017). Some of these differential phenotypic characteristics of *Macrococcus* species are listed in Table 2. These include colony

morphology and pigmentation, acetoin production, nitrate reduction and production of acid from the fermentation of D-mannitol, D-cellobiose, sucrose, lactose and glycerol.

Table 2: Phenotypic characteristics employed for the differentiation of *Macrococcus* species and closely related *Staphylococcus* species.
Symbols: +, 90% or more of the strains positive; -, 90% or more of the strains negative; V, variable reaction; ND, not determined.

	M. caseolyticus	*M. equipercicus*	*M. bovicus*	*M. carouselicus*	*M. brunensis*	*M. hajekii*	*M. lamae*	*M. canis*	*S. sciuri* subsp. *sciuri*	*S. vitulinus*	*S.lentus*
Character											
Pigment production	V	+	V	V	-	-	+	ND	V	V	V
Growth in: 4% NaCl	+	+	+	+	+	-	-	ND	+	+	+
7.5% NaCl	+	+	+	+	-	-	-	ND	+	+	+
Haemolysis	-	-	-	-	-	-	-	V	ND	ND	ND
Acetoin production	+	-	-	-	-	-	-	ND	-	-	-
Esculin hydrolysis	V	V	V	+	-	-	-	ND	+	V	+
Nitrate reduction	+	-	V	-	+	+	-	ND	+	+	+
DNase	V	-	V	+	-	-	-	ND	ND	ND	+
Urease	-	V	V	-	-	-	-	ND	-	-	-
Acid production from:											
Glycerol	V	+	+	V	-	-	-	ND	+	+	+
Sucrose	V	V	V	V	-	-	-	V	+	+	+
D-Mannitol	-	+	+	V	+	+	+	+	+	+	+
D-Cellobiose	-	-	-	-	-	-	-	ND	+	V	+
Lactose	V	-	-	-	-	-	-	-	V	-	V
D-Sorbitol	-	-	-	-	-	-	-	V	V	V	V
D-Mannose	-	-	-	-	-	-	-	V	V	-	+
Resistance to:											
Novobiocin	+	+	+	+	+	+	-	-	+	+	+
Esterase-lipase production	ND	ND	ND	ND	-	-	+	ND	ND	ND	ND

Adapted from Kloos *et al* (1976)(Kloos, Schleifer, & Smith, 1976); Kloos *et al* (1997)(Kloos et al., 1997); Kloos *et al* (1998) (W.E.Kloos et al., 1998a); Mannerová et al(2003) (Mannerová et al., 2003); Götz et al (2006)(Friedrich Götz et al., 2006) and Brawand *et al* (2017)(Brawand et al., 2017).

1.3.4.2 Cell Wall Composition

Due to the diversity in the structure of the peptidoglycan layer in Gram-positive bacteria, cell wall structure analysis and its chemical composition can be exploited for the differentiation of various Gram-positive species. The peptidoglycan of *M. caseolyticus, M. equipercicus, M. bovicus* and *M. carouselicus* species was determined to be of the A3α (L-Lys-Gly_{3-4}, L-Ser or Lys-Gly_3, L-Ser) type, which is a common type also found in several staphylococcal species. Variations in the composition of the cell wall of macrococcal species reported were in the levels of L-serine in the peptidoglycans of certain *Macrococcus* species, with the highest levels reported to be present in strains of *M. caseolyticus* [1-2-1.3 mol (mol glutamic acid)$^{-1}$] and the lowest levels were identified amongst the strains of *M. bovicus* [0.44-0.58 mol (mol glutamic acid)$^{-1}$] (Kloos et al., 1998a). Analysis of strain KM 45013^T of the recently-identified *M. canis* species also indicated an A3α type peptidoglycan (L-Lys–Gly_2–L-Ser)(Brawand et al., 2017).

Apart from the peptidoglycan layer, other components of the multi-component cell wall structure of Gram-positive bacteria involve anionic polymers such as teichoic acids and lipoteichoic acids. Cell wall teichoic acids are documented to be present only in the *M. caseolyticus* species from the *Macrococcus* genus (Kloos et al., 1998a). These co-polymers are generally found to be present in all of the staphylococcal species. Teichoic acids in most staphylococci consists of glycerol or ribitol (Götz et al., 2006). The teichoic acid present in *M. caseolyticus* species is described as atypical of the type poly (N-acetylglucosaminylphosphate) polymer, prior reports identified this unusual teichoic acid type in *Staphylococcus auricularis* (Endl et al., 1983). However, the level of teichoic acid found in the cell wall of *M. caseolyticus* was described to be an amount much lower than that found in the cell walls of staphylococci (Kloos et al., 1998a). Lipoteichoic acids are identified to be present at the cell surface of *M. caseolyticus, M. equipercicus, M. bovicus* and *M. carouselicus* species. The glycolipid in the cell membrane contains glycosyl residue of gentiobiosyl, which is the same glycosyl that is found in staphylococci (Kloos et al., 1998a).

1.3.4.3 Fatty Acid Methyl Ester (FAME) - Gas Chromatograph (GC) Analysis

Fatty acid profiling is a well-established method used for many years in the identification of bacteria, yeasts and fungi (Graham et al., 1995; Moser et al., 1996; Santos et al., 2018). FAME analysis involves the methylation of the fatty acids found

in bacterial cells, leading to the formation of methyl esters, which can then be identified by gas chromatography (GC) (Sasser, 1990; 2006). FAME analysis exploits the composition of highly conserved fatty acids to generate fatty acid profiles based on both qualitative difference (genus level) and quantitative difference (species level). This method has been exploited in the past to quantitatively differentiate all of the currently known macrococcal species from one another. The most predominant fatty acids of whole-cell hydrolysates of *M. canis* KM 45013^{T} strain are $C_{14:0}$, $C_{18:3}$ ω6c and $C_{16:0}$ n alcohol while in *M. caseolyticus* the predominant fatty acids are $C_{14:0,}$ $C_{18:1}$ ω9c and $C_{16:1}$ ω9c (Brawand et al., 2017). Mannerová *et al*, have reported abundant fatty acids for, *M. bovicus, M. carouselicus, M. equipercicus, M. brunensis, M. hajekii* and *M. lamae* strains (Mannerová et al., 2003).

1.3.4.4 Genotyping and DNA Base Composition

In the 1980's, major advancements in the development of the polymerase chain reaction (PCR) and DNA fingerprinting led to the development of the DNA-based typing that we exploit today for the identification and discrimination of microorganisms from genus to strain level, depending on the DNA-based method applied (Mullis et al., 1986; Jeffreys et al., 1988). DNA base composition and DNA-DNA hybridization were reported to be the first genotypic methods to be used in bacterial systematics (Priest and Goodfellow, 2012). Molecular typing methods were employed to investigate the inter-species and inter-strain differentiation of macrococci. These included strain-specific typing methods such as pulsed-field gel electrophoresis (PFGE) and ribotyping, species-specific typing including 16S rRNA sequencing and DNA-DNA hybridizations (DDH) and intragenus classification and separation of macrococci from staphylococci using genomic DNA base composition (Kloos et al., 1998a; Mannerová et al., 2003).

1.3.4.5 Genotypic Methods for Intergenus Differentiation

DNA base composition (GC content) varies amongst prokaryotes and is a key characteristic used in bacterial taxonomy to complete the description of novel genera (Parker, 2001). In 1965, Silvestri & Hill used this method to differentiate aerobic Gram-positive cocci from the genus of *Staphylococcus* (cocci with low GC content in their DNA) and the genus of *Micrococcus* (cocci high GC content in their DNA) (Silvestri and Hill, 1965). There are various studies that have investigated methods of determining the GC content of bacterial genomes and its applicability in intrageneric

classification (Boháček et al., 1967; Mesbah et al., 2011). Early investigation of DNA base composition of three *Micrococcus* species including *Micrococcus caseolyticus* (now referred to as *Macrococcus caseolyticus*) led to the reclassification of these species to the genus of *Staphylococcus* (Schleifer et al., 1982b). In 1998, Kloos *et al* reported the GC content of the DNA of the four *Macrococcus* species including *M. caseolyticus* to be 38-45% using the thermal denaturing method (Marmur and Doty, 1962). This was higher than staphylococci (30-39%) and lower than other Gram-positive, oxidase and catalase positive cocci belonging to the genera of *Kocuria, Dermacoccus, Kytococcus, Micrococcus* and *Deinococcus* (GC content of the DNA of these is reported to be 62–76%) (Kloos et al., 1998a). The novel genus *Macrococcus* was proposed, as a result of this genomic feature of the four species, which correlated with their distinctive phenotypic profile and subsequent analysis of their DNA-DNA hybridization and 16S rRNA sequencing further supported their separation from staphylococci. In 2003, Mannerová *et al* isolated additional members of this genus, *M. brunensis*, *M. hajekii* and *M. lamae* from llama skin, and reported the GC content of their DNA to be in the range of 40-42 mol% (Mannerová et al., 2003). However, GC content from the whole genome sequence of *M. caseolyticus* and *M. canis* are reported to be slightly below the previously defined range (36.5-36.9%) (Baba et al., 2009; Brawand et al., 2018).

1.3.4.6 Species-Specific Typing Methods

DNA–DNA hybridization (DDH) techniques have been used by taxonomists since the 1960's for the classification of prokaryotes. It has been considered a gold standard method for determining the extent of relatedness between a set of strains, through hybridization of their respective genomic sequences and evaluation of the resulting hybrids for their degree of association or their thermal stability (Rossello-Mora et al., 2011). The recommended cut-off point for DDH similarity to define a new species is ≤70% which corresponds to the 95% average nucleotide identity (ANI) cut-off value (Colston et al., 2014). ANI is an *in silico* tool that is an alternative to DDH as a means to measure genomic relatedness of prokaryotic strains (Arahal, 2014). In 1998, *in vitro* DNA hybridization analysis was performed on selected strains of *M. equipercicus, M. bovicus, M. carouselicus* and *M. caseolyticus* along with type strains from selected species of *Staphylococcus* and type strains of *Salinicoccus roseus* and *Bacilus subtilis* (Kloos et al., 1998a). The DNA relatedness of these strains was evaluated under

optimal (55°C) and stringent (77°C) hybridization conditions. Under both optimal and stringent conditions, the four *Macrococcus* species tested formed a well-defined, separate DNA similarity group. Strains from the same species of *Macrococcus* had DDH values higher than the cut-off point, whereas strains from different species of *Macrococcus* had significantly lower DDH values. According to the DDH similarity, the *M. caseolyticus* species was more distant from the three other *Macrococcus* species analysed (*M. equipercicus, M. bovicus* and *M. carouselicus*), which were more closely related to each other, forming a species group which was designated as the *Macrococcus equipercicus* species group (Kloos et al., 1998a). In fact, the *M. caseolyticus* species was the most closely related *Macrococcus* species to the *Staphylococcus sciuri* species group which correlates with the species low GC content values and its 16S rRNA sequence (Kloos et al., 1998a). Another study, proposing three novel additional species to the *Macrococcus* genus, *i.e. M. brunensis, M. hajekii* and *M. lamae,* performed DDH analysis on strains of these new species along with type strains of the previously defined *Macrococcus* species from Kloos *et al* (Kloos et al., 1998a; Mannerová et al., 2003). The resulting DDH values were below 54%, confirmed that the three novel species were relatively distant from the four previously defined species of *Macrococcus* (Mannerová et al., 2003). A recent study defined a new species, *Macrococcus canis*, based on DDH analysis of a former strain of *M. caseolyticus*; KM45013. A DDH value of 53.7% led to a reclassification of strain KM45013 as *M. canis*, and KM45013^{T} as a type strain (Brawand et al., 2017).

16S rRNA gene is the most common housekeeping genetic marker used in bacterial phylogeny and taxonomy (Janda and Abbott, 2007). The 16S rRNA gene is around 1500 bp consisting of well conserved and divergent regions, where conserved areas reflect the phylogenetic relationship amongst species and variable regions reflect differences between species (Mitreva, 2017). In contrast, to DDH, there are no defined sequence similarity threshold values reported for 16S rRNA, above which a universal agreement exists of what constitutes a definitive and conclusive identification to the rank of species (Janda and Abbott, 2007). This gene has been used in a number of reports to infer phylogenetic relationships amongst the currently known macrococci and their sister genus *Staphylococcus* (Kloos et al., 1998a; Götz et al., 2006). A recent study has proposed a percentage similarity matrix obtained from the alignment of the

16S rRNA gene sequences of the eight *Macrococcus* species indicating their independence from each other (Brawand et al., 2017).

1.3.4.7 Strain-Specific Typing Methods

The principal of PFGE was first described by Schwartz and Cantor (Schwartz and Cantor, 1984). This method requires the preparation of unsheared DNA, digestion of the DNA with infrequent cutting restriction enzyme, separation of fragments by employing alternately pulsed electric field, followed by visualization and interpretation of banding patterns (Kaufmann, 1998). PFGE provides vital information on the degree of genetic relatedness between strains and is referred to as the "gold standard" fingerprinting method due to its high discriminatory power (Hata, 2010). This method is also described as a superior typing method in comparison to most of the other molecular typing methods available (Olive and Bean, 1999). Tenover *et al*, have reported the criteria for interpreting PFGE fingerprints in relation to determining strain relatedness (Tenover et al., 1995). Kloos *et al*, used restriction endonuclease *Sma*I to digest the chromosome of *M. caseolyticus* (Kloos et al., 1998a). *Sma*I is also commonly used to generate genome fingerprints of staphylococcal species and cleaves the *Staphylococcus* chromosome into ~7-30 fragments (George and Kloos, 1994). In the *M. caseolyticus* strains tested, *Sma*I cleaved the genome into 29-34 fragments and conserved sizes of these fragments amongst *M. caseolyticus* strains have been reported in this study (Kloos et al., 1998a). *Not*I was used to generate banding patterns for the members of the *M. equipercicus* species, producing 33-37 fragments. Most of the generated fragments were identified to be highly conserved amongst the members of this species. *Not*I also cleaved *M. bovicus* chromosome (14-17 fragments) and *M. carouselicus* strains (12-16 fragments) (Kloos et al., 1998a). Mannovera *et al*, utilised a *Xba*I to generate fingerprints of *M. brunensis, M. hajekii* and *M. lamae* strains and provided cluster analysis of the macrorestriction fragments of the typed strains of all of the then known macrococcal species (except *M. canis*) indicating their phylogenetic relationship (Mannerová et al., 2003).

The knowledge of the intraspecies conservation of the 16S rRNA gene sequence (Fox et al., 1980), and basic structure of the ribosomal operon consisting of three ribosomal RNA genes 16S-23S-5S (Doolittle and Pace, 1971) led to the development of ribosomal RNA gene restriction pattern analysis for bacterial classification (Grimont and Grimont, 1986). Conventional ribotyping requires restriction endonuclease

digestion of entire chromosomal DNA followed by electrophoretic separation of the resulting DNA fragments, southern blot transfer, and hybridization of transferred DNA fragments with a radiolabeled ribosomal operon probe (Southern, 1975; Bouchet et al., 2008). After probing, only those bands containing a portion of the ribosomal operon are highlighted, creating a fingerprint. The number of fragments generated using this method corresponds to the multiplicity of ribosomal RNA operons present in that bacterial species. Studies have reported the diversity of ribotype patterns found amongst the members of the *Macrococcus* genus. Comparison of these patterns indicated that strains of four *Macrococcus* species i.e. *M. equipercicus, M. bovicus, M. carouselicus* and *M. caseolyticus*, appeared to be closer to one another, than to any of the *Staphylococcus* species (Kloos et al., 1998a). Typed strains from these species were then combined and examined with three additional *Macrococcus* species (*M. brunensis, M. hajekii* and *M. lamae*). The pattern analysis identified a clear separation of the macrococcal species from staphylococcal species, with a low similarity of 17% (Mannerová et al., 2003).

1.4 General Properties, Nomenclature, and Ecology of *Micrococcus* Species

The general characteristics of species from this genus have been defined in a number of studies (Kloos et al., 1998a; Kloos et al., 1998b; Mannerová et al., 2003; Brawand et al., 2017), and some of these characteristics of the type strains are listed in Table 3.

1.4.1 General Properties and Nomenclature

1.4.1.1 *Micrococcus caseolyticus*

The first description of this species was made in 1916 by Evans, who isolated this bacterium from cow's milk. The species was characterised by its ability to rapidly and completely peptonise the milk and therefore, the name *caseolyticus* (casein-dissolving) was suggested (Evans, 1916). Early screening studies investigated the presence of *M. caseolyticus*, at the time referred to as *Staphylococcus caseolyticus*, on the skin of cattles, goats, horses, ponies, whales, dolphins and meat products (Ballard et al., 1995). However, the investigation only isolated *M. caseolyticus* in three samples of beef, and from the skin of a pilot whale. It was therefore believed to be a relatively uncommon species in contrast to other species of *Micrococcus*, which were identified in larger numbers. This species was described as the most distinguishable micrococcal species phenotypically, displaying less phenotype similarity with other known micrococcal species. This correlated with the *M. caseolyticus* genotypic data, and therefore the species was separated from the *M. equipercicus* species group (Kloos et al., 1998a). 16S rRNA sequence analysis indicate it is most closely related to the newly-defined *M. canis* species with sequence similarities between the type strains of both species (*M. caseolyticus* 13548^T and *M. canis* KM45013^T) of 99.7% (Brawand et al., 2017).

1.4.1.2 *Micrococcus canis*

The majority of the strains, including the type strain of this species have been isolated from canine sources and therefore, the name *M. canis* has been proposed (Brawand et al., 2017). The type strain of this species, KM45013^T,was isolated from the nares of a dog with rhinitis and was previously recognised as a strain of *M. caseolyticus* (Gomez-Sanz et al., 2015). Later, with extensive genotypic and chemotaxonomic analysis, KM45013^T and other strains isolated from healthy dogs as well as infection sites were assigned to this novel species (Brawand et al., 2017). Phenotypic characteristics that

differentiate this species from its most closely related *M. caseolyticus* species include the ability of *M. canis* species to ferment mannitol and the inability to produce acid from lactose. The presence of complete haemolysis is also a specific trait of this species (Brawand et al., 2017).

1.4.1.3 *Macrococcus equipercicus*

This species, along with *M. bovicus* and *M. carouselicus*, were fortuitously isolated during a study investigating the presence of *M. caseolyticus* (at the time referred to as *S. caseolyticus*) on the skin of various mammals (cattle, goats, horses, ponies, whales and dolphins) and in meat products (Ballard et al., 1995). The first descriptions of the species and its phenotypic capabilities were reported by Kloos *et al*, based on the characteristics of 22 strains isolated from the skin of horses and ponies. The name *M. equipercicus* was suggested, after the horse named Percy from which it was first isolated. This species differed from other species of macrococci, due to its larger cell size and the presence of small piliform projections on its cell surface. *M. equipercicus* strains were identified to be resistant to novobiocin and susceptible to lysostaphin (Kloos et al., 1998a). Recent analysis 16S rRNA sequence analysis indicate it is most closely related to *M. carouselicus* with sequence similarities between type strains of both species (*M. equipercicus* ATCC51831^T and *M. carouselicus* ATCC 51828^T) of 99.6% (Brawand et al., 2017).

1.4.1.4 *Macrococcus bovicus*

This species was first isolated from bovine skin, and therefore the species name *bovicus* was suggested. The first descriptions of the species and its phenotypic capabilities were reported by Kloos *et al*, based on the characteristics of 10 strains isolated from the skin of a cow, a pony and various horses. According to scanning electron microscopy analysis (SEM), the cell surface of this organism is irregular. Culture growth causes alpha haemolysis (greening) of horse and bovine blood, and all strains were shown to be resistant to novobiocin and susceptible to lysostaphin and oxacillin (Kloos et al., 1998a). 16S rRNA sequence analysis indicate it is most closely related to *M. brunensis* with sequence similarities between the type strains of both species (*M. brunensis* CCM4811^T and *M. bovicus* ATCC 51825^T) of 99.8% (Brawand et al., 2017).

1.4.1.5 *Macrococcus carouselicus*

The name of this species pertains to carousel horses, as strains of this organism were identified on the skin of horses and ponies in larger numbers than other *Macrococcus* species. The phenotypic properties associated with this species were established through examination of the characteristics of 18 strains of *M. carouselicus* isolated from the skin of the above mammals (Kloos et al., 1998a). All of these strains were resistant to novobiocin and susceptible to lysostaphin and oxacillin. The species is also positive for DNase activity. 16S rRNA sequence analysis indicate it is most closely related to *M. equipercicus* with sequence similarities between the type strains of both species (*M. equipercicus* ATCC51831^T and *M. carouselicus* ATCC 51828^T) of 99.6% (Brawand et al., 2017).

1.4.1.6 *Macrococcus brunensis*

Mannerova *et al*, used a polyphasic approach to investigate and establish phenotypic and genotypic characteristics of *M. brunensis* along with two other macrococcal species (*M. hajekii* and *M. lamae*) (Mannerová et al., 2003). According to this study, *M. brunensis* was isolated from llama skin and the species name pertains to the city of Brno (Roman name of this city is Bruna) from where the type strain of this species was first isolated. All strains of this species were reported to hydrolyse casein, gelatin and were negative for the hydrolysis of Tween 80, starch, lecithin, aesculin and tyrosine. The species was defined to be resistant to novobiocin and susceptible to furazolidone. 16S rRNA sequence analysis indicate it is most closely related to *M. bovicus* with sequence similarities between the type strains of both species (*M. brunensis* CCM4811^T and *M. bovicus* ATCC 51825^T) of 99.8% (Brawand et al., 2017).

1.4.1.7 *Macrococcus hajekii*

This species is named after a Czech microbial taxonomist, Václav Hájek and was isolated from the skin of a llama (Mannerová et al., 2003). The species is capable of hydrolysing casein and gelatin but incapable of hydrolysis of Tween 80, starch, lecithin, aesculin and tyrosine. All strains of this species are urease and haemolysis negative and resistant to bacitracin and susceptible to furazolidone. 16S rRNA sequence analysis indicate it is most closely related to *M. equipercicus* with sequence similarities between type strains of both species (*M. equipercicus* ATCC51831^T and *M. hajekii* CCM4809^T) of 99.4% (Brawand et al., 2017).

1.4.1.8 *Macrococcus lamae*

The name of this species pertains to the zoological genus name of the llama, from which it is isolated (Mannerová et al., 2003). All members of the species are capable of hydrolysing casein, gelatin, and are negative for the hydrolysis of Tween 80, starch, lecithin, aesculin and tyrosine. The species was defined to be susceptible to furazolidone and moderately susceptible to novobiocin (1.6 mg). 16S rRNA sequence analysis indicate it is most closely related to *M. equipercicus* with sequence similarities between type strains from both species (*M. equipercicus* ATCC51831^{T} and *M. lamae* CCM4815^{T}) of 99.1% (Brawand et al., 2017).

Table 3: General properties of the type strains of *Micrococcus* species.

Character	*M. caseolyticus* ATCC13548^T	*M. equipercicus* ATCC51831^T	*M. bovicus* ATCC51825^T	*M. carouselicus* ATCC5128^T	*M. brunensis* CCM4811^T	*M. hajekii* CCM4809^T	*M. lamae* CCM4815^T	*M. canis* KM45013^T
Cell size (µm)	1.1-2	1.3-2.3	1.2-2.1	1.4-2.5	0.89–1.2	0.89	0.74-0.92	0.8
Cell wall teichoic acids	+	-	-	-	-	-	-	ND
Colony diameter (mm)	4-5	8	3-5	4-5	2-4	2-3	2-5	2
Colony pigmentation: orange	-	+	+	+	-	-	+	-
Colony lustre: glistening	+	-	+	+	+	+	+	+
DNA G+C content (mol%)	38-39	45	44	41	42	40	41	36.9
Haemolysis(β-type)	-	-	-	-	-	-	-	+
Peptidoglycan type	Lys-Gly_4, L-Ser	Lys-Gly_{3-4}, L-Ser	Lys-Gly_3, L-Ser	Lys-Gly_{3-4}, L-Ser	ND	ND	ND	L-Lys-Gly_2, L-Ser
Proteolytic activity on casein	+	ND	ND	ND	+	+	+	ND
Clumping factor	+	-	+	+	-	-	-	ND

Symbols: +, positive; -, negative; ND, not determined.

Adapted from Evans (1916), (Evans, 1916); Kloos *et al* (1998), (W.E.Kloos et al., 1998a); Mannerová et al(2003) (Mannerová et al., 2003); Götz et al (2006)(Friedrich Götz et al., 2006) and Brawand et al (2017)(Brawand et al., 2017).

1.4.2 Ecology

Extensive ecological studies have reported the colonisation of the skin and mucous membranes of various birds and mammals by staphylococci (Götz et al., 2006). In contrast, limited numbers of reports have identified patterns of host preference of the *Macrococcus* species. In some cases qualitative patterns of similar host preferences between species of the two genera have been highlighted (Wang et al., 2012b; Cotting et al., 2017).

The first described species of this genus, *M. caseolyticus* was primarily isolated as a contaminant of milk and is thought to be a member of the bacterial flora of cow's udders (Evans, 1916). This species is found in smaller population numbers in comparison to other macrococcal species that share similar habitats. Studies have identified its preference for artiodactyl (cattle, sheep and goats) and cetacean (whale) hosts. This species has been found predominantly on the skin of these mammals, with a few strains identified in their milk and meat products (Kloos et al., 1998a). Other studies have isolated strains of this species from pigs (nasal area) and canine sources (Wang et al., 2012a; Cotting et al., 2017). *M. caseolyticus* strains have been found living on the skin of dogs and colonising the groin and oral mucosa in these animals. In contrast, the occurrence of *M. canis* in these canine sources is more frequent than *M. caseolyticus* species. *M. canis* and *M. caseolyticus* strains have been isolated from animals with rhinitis, otitis externa, dermatitis and mastitis (Cotting et al., 2017; Schwendener et al., 2017). *M. canis* was originally isolated from the nares of a dog with rhinitis and its presence on the healthy body sites as well as infection sites in the two breeds, West Highland white terriers and Newfoundland dogs, have been reported (Gomez-Sanz et al., 2015). Investigations have identified the occurrence of *M. canis* in healthy body sites as well as infected sites in dogs, resembled to what is observed with *Staphylococcus pseudintermedius* (opportunistic pathogen). This staphylococcal species is a part of the normal microflora of healthy dogs however, the occurrence of *M. canis* on the skin of the above breed of dogs were reported to be less frequent than *S. pseudintermedius* (colonises 46-92% dogs) (Cotting et al., 2017). *M. canis* strains are reported to colonise areas such as axilla, ear, the oral mucosa and the groin in these dogs. The presence of this bacterium, along with other *Macrococcus* species, in other breeds of dogs has also been highlighted (Cotting et al., 2017).

M. equipercicus and *M. carouselicus* species have a preference for perissodactyl hosts (such as horses and ponies) and these species have been found in large populations on the skin of these mammals. *M. bovicus* species was first reported to be isolated from the skin of cattle and the species also has preference for perissodactyls (horses and ponies) and artiodactyls (cattle) hosts (Kloos et al., 1998a). *M. brunensis*, *M. hajekii* and *M. lamae* have been defined to be present as part of the surface microflora of llamas and strains of these species have been isolated from specific sites of the llama such as the nostrils, groin, abdomen, area around the jaws and under the tail (Mannerová et al., 2003).

1.5 Comparative Genomics and Evolution

1.5.1Genome Comparison with Other Closely Related Organisms

There are currently three macrococcal strains that have been completely sequenced, two of which belong to *M. caseolyticus* species (strains JCSC5402 and IMD8019) and one from the latest member of *Macrococcus* genus, which is *M. canis* (KM45013^T strain). The general features of these sequenced strains, along with similar features of the closely related *S. aureus* N315 and *B. subtilis* 168 genomes are listed in Table 4. The GC content of *M. caseolyticus* JCSC5402, IMD8019 and *M. canis* KM45013^T is 36.7-37.1%, values which fall between those of *S. aureus* and *B. subtilis* (Baba et al., 2009). BLAST analysis of *M. caseolyticus* JCSC5402 revealed that more than half (64.9%) of its 1,957 open reading frames (ORFs) were found to have genes that were orthologous to the genes present in staphylococcal species. In particular, core genes such as those involved in DNA replication, RNA transcription, translation, glycolysis and the TCA cycle were the most akin to those in the family of *Staphylococcaceae*. This demonstrated that the organism is more closely related to *Staphylococcus* than to any other genus (Baba et al., 2009). However, half of the reported 63 ORFs which belong to class 1-4 membrane bioenergetics including the genes involved in cellular respiration are more similar to the orthologs in the family of *Bacillaceae*. The taxonomic distribution of genes akin to the 113 ORF's on the three prophages and the genes found on the eight plasmids identified in strain JCSC5402 were homologous to a wide range of Gram-positive bacteria including *Staphylococcus, Bacillus, Streptococcus, Lactobacillus, Enterococcus, Listeria* and *Mycoplasma*, which suggested that these mobile elements are omnipresent across a great range of bacteria belonging to the phylum *Firmicutes* (Baba et al., 2009).

Further investigations of the JCSC5402 genome revealed conserved chromosomal structures which were also present in the genera *Staphylococcus* and *Bacillus*. The phylogenetic distance calculated between these three genera suggests that they have been derived from a common Gram-positive bacterium. The *Macrococcus* species are disseminated in nature as animal commensals and are indicated to be the immediate antecedent of the *Staphylococcus* species. According to the investigations carried out by Hiramatsu *et al*, the appearance of staphylococci in the phylogenetic time tree was estimated to be about 250 million years ago whereas macrococcal species were described to be present before this time (Hiramatsu et al., 2014). In both studies, the

phylogenetic analysis has suggested that the ancestral bacterium of the family *Staphylococcaceae* occurred from the divergence of the family *Bacillaceae* (Baba et al., 2009; Hiramatsu et al., 2014). The small genome size of the *Staphylococcaceae* ancestral bacterium in contrast to *Bacillus* species (*S. aureus* 2.8 Mb versus 4.2 Mb *B. subtilis*) may have been a result of co-evolution with mammals and the impact of the new environment, causing the downsizing of the chromosome through the loss of non-essential genes. Following colonisation of the animal body with the ancestral bacterium, speciation of macrococcal and staphylococcal species may have started. This speciation event, indicated to have occurred about 225 million years ago, may have size through the acquisition of an array of virulence genes, whereas *Macrococcus* remained avirulent to the animal host (Baba et al., 2009; Hiramatsu et al., 2014).

Table 4: Comparison of the genomics features of sequenced *Macrococcus* species; *M. caseolyticus* JCSC5402, *M. caseolyticus* IMD0819 and *M. canis* KM45013^{T} to the closely related *S. aureus* N315 and *B. subtilis* 168.

Parameters	*M. caseolyticus* JCSC5402[a]	*M. caseolyticus* IMD0819[b]	*M. canis* KM45013$^{(T)}$ [c]	*S. aureus* N315[d]	*B.* subtilis 168[e]
Chromosome Length of sequence (Mb)	2.1	2.27	2.36	2.81	4.21
G + C content (total genome %)	36.9	36.7	37.1	32.8	43.5
ORF's					
No. of protein coding regions	1,957	2,297	2,438	2,595	4,106
No. of rRNAs					
16S	4	5	5	5	10
23S	4	5	5	5	10
5S	5	6	6	6	10
No. of transfer RNAs	48	59	59	62	86
No. of Plasmids	8	0	0	1	0

[a] Genome sequence reported by Baba et al. (2009).

[b] Genome sequence reported by Schwendener et al. (2017).

[c] Genome sequence reported by Brawand et al. (2017).

[d] Sequenced at the University of Oklahoma Health Sciences Centre.

[e] Genome sequence reported by Barbe *et al*. (2009).

1.5.2 Interspecies Differences in Metabolic Pathways *Staphylococcus* versus *Macrococcus*

1.5.2.1 Terminal Oxidases of Respiratory Chain

Oxidative phosphorylation is a vital process that generates cellular energy through the transfer of electrons from oxidisable substances to final acceptor molecules, a process that is accomplished by the stepwise transport of electrons through the electron transport chain composed of specific dehydrogenases and respiratory enzyme complexes (Gel'man et al., 1967). Despite the close evolutionary relationship outlined between the *Macrococcus* and *Staphylococcus* genus, the electron transport components involved in the oxidative phosphorylation pathway differ. In 2009, Baba *et al*, presented a three-way comparison of the genes involved in the electron transport chain of *M. caseolyticus* JCSC5402 with those present in the *Staphylococcus* and *Bacillus* species. This investigation indicated that, with the exceptions of *S. sciuri* group of species, *Staphylococcus* species do not possess the genetic machinery for the terminal electron transducer component cytochrome c oxidase complex, which is present in both *Macrococcus* and *Bacillus* species (Baba et al., 2009).

Cytochrome c oxidase (COX) is a terminal oxidase, composed of four catalytic subunits (SUI, SUII SUIII and SUIV) that functions to transport electrons from reduced cytochrome c to the final electron acceptor oxygen molecule (Thöny-Meyer, 1997). The cytochrome c present in *M. caseolyticus* 13548^{T} is described to be of aa3-type cytochrome c (Kloos et al., 1998a). This type of cytochrome c is described to be highly conserved in different organisms. The COX genes in *M. caseolyticus* JCSC5402 are organised in the operon cta*ABCDEFG* (GenBank accession number; NC_011999.1). The structural subunits of cytochrome c encoded by *ctaCDEF*, are described in *Bacillus* species (Liu and Taber, 1998; Barbe et al., 2009). Fig. 3a illustrates the gene organisation of cytochrome c oxidase operon for *M. caseolyticus* JCSC5402.

The component of cytochrome *bd* oxidase (encoded by Cyd*AB*), an alternative terminal oxidase which functions under microaerobic conditions identified in staphylococcal species, including *S. aureus* are also present in *M. caseolyticus* and *B. subtilis* (Baba et al., 2009). Other two alternative and menaquinol-dependent terminal oxidases involved in the branched respiratory system of *S. aureus* are illustrated in Fig

.3b. These terminal oxidases are capable of accepting electrons directly from quinol and using them for reducing molecular oxygen (Thöny-Meyer, 1997). In *S. aureus*, the main terminal proton pumping oxidases that function under aerophilic conditions are quinol oxidase and cytochrome *bo* oxidase. The genes which encode for quinol oxidase system (Qox*ABCD*) have been identified in *S. aureus*, composed of cytochromes a-602 and b-561. Cytochromes a-605, which is identified in all staphylococcal species is described to be a part of the Qox system and is associated with a ubiquitous cytochrome *b*-552 (Faller et al., 1980). Cytochrome *bo* oxidase is composed of cytochrome *b*-555 that is preconnected to cytochrome *b*-557 (Götz and Mayer, 2013).

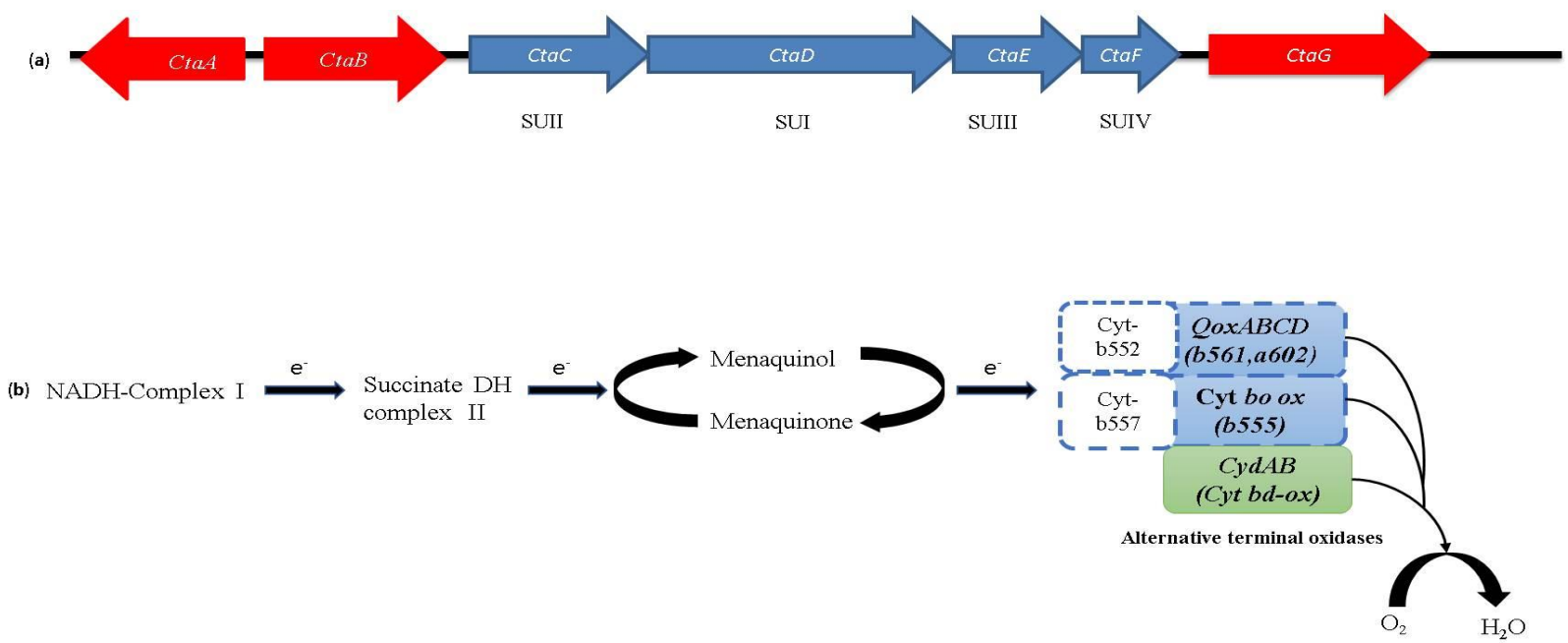


Fig. 3(**a**) Organisation of genes encoding caa3 cytochrome c oxidase (COX) in *M. caseolyticus* JCSC5402 strain (NC_011999.1). Below the structural genes are indicated their respective subunits (SU) (*ctaCDEF*; shown in blue). Genes that are involved in the production of: heme A (*ctaA)* heme O (*ctaB*) and caa3 assembly (*ctaG)* are shown in red.

(**b**) Illustrating the proposed respiratory system of *S. aureus*. Electrons (e^-) transferred from the complex I and II lead to the reduction of menaquinone to menaquinol which then transfers its electrons to several terminal oxidases, that results in reducing oxygen to H_2O. Proton pumping terminal oxidases are shaded blue, these include quinol oxidase (*QoxABCD;* composed of cytochromes *a*-602, *b*-561 and pre-connected to b-552) and cytochrome *bo* oxidase (consisting of b-555 and connected to b-557). Non-proton pumping terminal oxidase cytochrome *bd* oxidase (*CydAB*), is shaded green.

1.5.2.2 Starch Degradation and Glycogen Biosynthesis

Glycogen is an intracellular glucose polymer that has been reported in over 40 different species which includes Gram negative, Gram positive, photosynthetic and archaebacteria (Iglesias and Preiss, 1992). The main role of glycogen biosynthesis in bacteria is to serve as a process by which organisms accumulate energy reserves, which can be accessed under starvation condition (Iglesias and Preiss, 1992). The starch digesting alpha amylase gene that is absent in staphylococcal genomes is present in *M. caseolyticus* JCSC5402 genome. In addition, the genes for glycogen biosynthesis (*glg*) are organised in an operon *glgBCDAP* present in *M. caseolyticus* JCSC5402, along with the starch digesting amylase gene were identified to be akin to those in *Bacillus* species. This suggests that *M. caseolyticus* can store polymerized forms of glucose and digest starch. The ability of this organism to carry out such activities indicates its ability to utilise glucose efficiently in an environment where glucose shortages are of a great concern. These genes along with those genes involved in the respiratory chain are reported to be conserved amongst the species of *M. caseolyticus* and *Bacillus,* suggesting that the genome of *M. caseolyticus* retains a part of the genome of the common ancestor for the genera *Bacillus* and *Macrococcus* (Baba et al., 2009).

1.5.2.3 Sugar Transport

Limited sugar utilisation capability has been reported for *M. caseolyticus* JCSC5402 due to the lack of a complete set of genes required for the metabolism of mannitol, trehalose, maltose and lactose (Baba et al., 2009). Whereas, the transport and metabolism of the majority of these carbohydrates have been well studied in *S. aureus*, *Staphylococcus xylosus* and *Staphylococcus carnosus* species (Egeter and Brückner, 1995; Götz et al., 2006). The lactose (*lac*) operon required for the lactose metabolism has been reported in *S. aureus*. The lac (*lacABCDFEG*) operon consists of structural genes (*lacABCD*) encoding for the enzymes required for the tagatose-6-Phosphatase pathway, where *lacA* and *lacB* codes for galactose-6- phosphatase isomerase, *lacC* for tagatose-6-Phosphatase kinase, and *lacD* for tagatose-1,6-diphoasphatase aldolase (Rosey et al., 1991). The genes *lacF and lacE* encode for the galactoside-specific PTS permease, enzyme IIA, and enzyme IICB and the last gene in the lac operon *lacG*, encodes for the phospho-β-galactosidase enzyme (Breidt and Stewart, 1986). The *lacR* encoding the lactose phophotransferase system repressor is located upstream of the

lacABCDFEG operon (Oskouian and Stewart, 1990). However, *M. caseolyticus* IMD0819 (GenBank accession number; NZ_CP021058.1) genome confirms the presence of complete set of genes (*lacABCDFEG*) required for lactose metabolism. The mannitol phosphotransferase (*mtl*) system have been studied in *S. carnosus*, this system consists of an EIICB enzyme specified by the gene *mtlA* and EIIA encoded by *mtlF* gene, together these components form mannitol-specific PTS permease. These genes are clustered along with *mtlD* gene which encodes for mannitol-1-phosphate-5 dehydrogenase which produces fructose 6-Phosphate (Fischer et al., 1989; Fischer and Hengstenberg, 1992). The *mtl* system had also been identified in *S. aureus* genome (Reiche et al., 1988). In contrast to *M. caseolyticus* strains, *M. canis* KM45013^T strain has indicated a mannitol positive phenotype, which is confirmed by the presence of the *mtl* system in *M. canis* KM45013^T genome (GenBank accession number; CP021059.1) (Brawand et al., 2017).

1.5.2.4 Iron Acquisition

Iron is an important micronutrient and microrganisims obtain iron from sources such as the host proteins (transferrin and lactoferrin), heme and through the production of low molecular weight iron chelators, known as siderophores (Krewulak and Vogel, 2008). *S. aureus* requires iron for host colonization and successive pathogenesis. Siderophore biosynthesis has been reported in *S. aureus* species. *S. aureus* produces two distinct sidephores known as staphyloferrin A and staphyloferrin B, through two different pathways (a) the non-ribosomal peptide synthetase pathway and (b) the non-ribosomal peptide synthetase independent pathway (Hammer and Skaar, 2011). The small genome of *M. caseolyticus* JCSC5402 strain lacks the siderophore biosynthesis genes which are involved in the iron acquisition pathway. The fact that *M. caseolyticus* lacks the majority of these genes along with other genes involved in iron transport indicate that this organism colonises such environments where it does not need to be aggressive in iron import (Baba et al., 2009).

1.5.2.5 Vitamins: Biotin and Pantothenate Synthesis

Biotin (also known as vitamin H or B7) is an essential vitamin for all life forms, which functions as an active cofactor for biotin-dependent enzyme reactions that includes carboxylases, decarboxylases and transcarboxylases (Streit and Entcheva, 2003). The clustering of the genes that encode for biotin biosynthetic enzymes and the role of bifunctional BirA protein as an effective regulator of the biotin operon has been

investigated in *S. aureus* species (Satiaputra et al., 2016; Satiaputra et al., 2018). Pantothenate (also known as vitamin B5) is indicated to be a key precursor for coenzyme A (CoA), which is an essential enzyme cofactor that functions as an acyl group carrier and carbonyl activating group in various reactions involved in cellular metabolism such as the synthesis of phospholipids, synthesis and degradation of fatty acids, and the operation of the tricarboxylic acid cycle (Leonardi and Jackowski, 2007; Spry et al., 2007). The genes encoding for the enzymes required in the pantothenate biosynthesis pathway, such as pantothenate synthetase (*panC*) has also been identified in *S. aureus* species (Satoh et al., 2010). In contrast, *M. caseolyticus* strain JCSC5402 lacks a full set of genes required for synthesis of these essential vitamins, indicating the species must obtain these from the environment that it colonises (Baba et al., 2009).

1.6 Genomic Analysis of Virulence Genes and Regulatory Systems

1.6.1 Exotoxin Genes and Regulators

The genome analysis of *M. caseolyticus* JCSC5402 has defined this organism to lack all of the virulent determinants identified in *S. aureus* (Baba et al., 2003). However, a 54% identity match of ORF (MCCL_1166) to the hemolysin A gene of *Bacillus cereus* was reported (Baba et al., 2009). Hemolysin A is one of the exotoxins identified in *S. aureus* and is described to play a part in pathogenesis caused by *S. aureus* species (Berube and Wardenburg, 2013). The genome study of *M. canis* KM45013^{T} have identified the presence of *hlgC* and *hlgB,* biocomponents of another exotoxin, γ-Hemolysin. (Brawand et al., 2018). The γ-Hemolysin is also one of the five cytolytic toxins identified in *S. aureus* and the locus of this toxin has been reported to be composed of three linked genes *hlgA*, *hlgB* and *hlgC.* These genes encode for the components required in the production of the gamma haemolysin toxin (Cooney et al., 1993). Studies have indicated the production of these and other virulence factors by *S. aureus*, which are expressed under the control of at least two regulator genes, accessory gene regulator (*agr*) and staphylococcal accessory regulator (*sar*) (Abdelnour et al., 1993; Manna et al., 1998). These regulatory systems are absent in *M. caseolyticus* JCSC5402 and therefore, the organism does not appear to harbour toxin genes whose expression is under the control of *agr* and *sar*. Thus, it has been suggested that after the divergence of the genera *Staphylococcus* and *Macrococcus*, *S. aureus* acquired virulence genes along with its regulatory genes (Baba et al., 2009).

1.6.2 Two-Component Regulatory System

Two-component regulatory systems are widespread amongst Gram positive bacteria and these systems constitute a major part of bacterial signal transduction pathways (Zwir et al., 2007). These systems allows the host organism to adapt to changing environmental conditions through modifying cellular physiology, which includes initiating gene expression, catalyzing reactions, or modifying protein–protein interactions (Skerker et al., 2005; Monedero et al., 2017). A total of eleven sets of two-component systems were identified on the genome of *M. caseolyticus* JCSC5402, ten of which are most similar to those in *Staphylococcus* species. Six of these ten were also present in *Bacillus* species, whereas BLAST analysis of another set, a sensor kinase and a response regulator, are most similar to those of *Clostridium tetani* and *Clostridium perfringens* (Altschul et al., 1997; Baba et al., 2009). In *S. aureus* a total

of sixteen conserved two-component regulators have been identified (Kolar et al., 2011). Orthologs of *vraSR*, *phoPR*, *nreBC*, *lysSR*, *srrAB*, and *arlSR*, two-component regulatory systems in *S. aureus*, were also identified in *M. caseolyticus* JCSC5402 genome (Baba et al., 2009). In *S. aureus*, the *vraSR* two-component system is associated with cell wall biosynthesis. Studies have defined this system as "sentinel" as it has indicated to be capable of rapidly sensing and coordinating a response to cell wall peptidoglycan damage (Belcheva and Golemi-Kotra, 2008). The *phoPR* two component systems participate in the cellular response to conditions of phosphate availability (Hulett, 2002; Allenby et al., 2005; Howell et al., 2006). The *nreBC* two-component regulatory system encoded by the *nre*ABC operon, is associated with the nitrate/nitrite reduction in response to oxygen (Fedtke et al., 2002; Kamps et al., 2004). The *srr*AB two-component is associated with mediating aerobic to anaerobic switch in response to changes in respiratory flux. (Mashruwala and Boyd, 2017). Finally, the *arlSR* two-component regulatory system composed of a sensor protein (ArlS) and a response regulator (ArlR) is recognised as an autolysis related locus. In addition to controlling the rate of autolysis, this system also demonstrated to control attachment to polymer surfaces by influencing the secreted peptidoglycan hydrolase activity (Fournier and Hooper, 2000). Other studies have also associated this system with the regulation of exoproteins as reports have identified a decrease in the transcription and production of β-haemolysin, lipase, coagulase and cell wall bound protein A in *S. aureus* regulated under the control of the *arlSR* system (Fournier et al., 2001). The physiological functions of the other four two-component regulatory system were not established (Baba et al., 2009).

1.7Antibiotic Resistance

1.7.1 Mechanisms of β-lactams and Methicillin Resistance in Staphylococci and Macrococci

Penicillin has been used in humans since the 1940's and within a year of its first administration, the first case of penicillin resistance in *S. aureus* emerged (Rammelkamp and Maxon, 1942). The resistance was directed by an enzyme called penicillinase (β-lactamase), which inactivates the penicillin by hydrolysing its β-lactam ring structure. The β-lactamase is encoded by the *blaZ* gene, located on a large transposon that was identified on a plasmid (Ambler, 1980; Peacock and Paterson, 2015). Fig. 4a illustrates this enzymatic inactivation of β-lactams present in *S. aureus* species. Numerous studies were then undertaken to chemically alter the molecule of penicillin thus prevent its cleavage by β-lactamase (Davies and Davies, 2010). Later, in 1959, a landmark discovery was made with the introduction of methicillin, envisaged to be a guaranteed defence against the action of penicillinases. However, the protection was short-lived, as within three years methicillin-resistant *S. aureus* (MRSA) strains began to emerge and after five decades, more than a half of all *S. aureus* clinical isolates were found to be methicillin resistant. Methicillin resistance initially emerged in *S. aureus* through the acquisition of the *mecA* gene in susceptible strains, as a result of horizontal gene transfer mediated by staphylococcal cassette chromosome (SCC*mec*) from an ancestral bacterium of *Staphylococcus* species (Hiramatsu et al., 2014). This *mecA* gene encodes for a modified version of the penicillin-binding proteins (PBP) referred as PBP2a (Fig. 4b). The PBP are essential proteins required in the synthesis of new peptidoglycan, employed at the cell division site during the bacterial cell cycle (Scheffers and Pinho, 2005). The alternative PBP2a has an unusually low affinity for all β-lactams and therefore, conserves the transpeptidation activity of the PBP's even in the presence of high concentrations of these antibiotics, which then allows the cell wall biosynthesis to continue (Macheboeuf et al., 2006). Recently, methicillin resistance has been identified in the two closely related species of *M. caseolyticus* and *M. canis* isolated from chicken skin, bovine and canine sources (Tsubakishita et al., 2010a; Cotting et al., 2017; Schwendener et al., 2017). Methicillin resistance in *Macrococcus* is associated with the acquisition of *mecA* homolog, *mecB* which also encodes for an alternative PBP

(PBP2a also called PBP2′). Fig. 4b illustrates the mechanism of β-lactam resistance in *M. caseolyticus* and *M. canis* species.

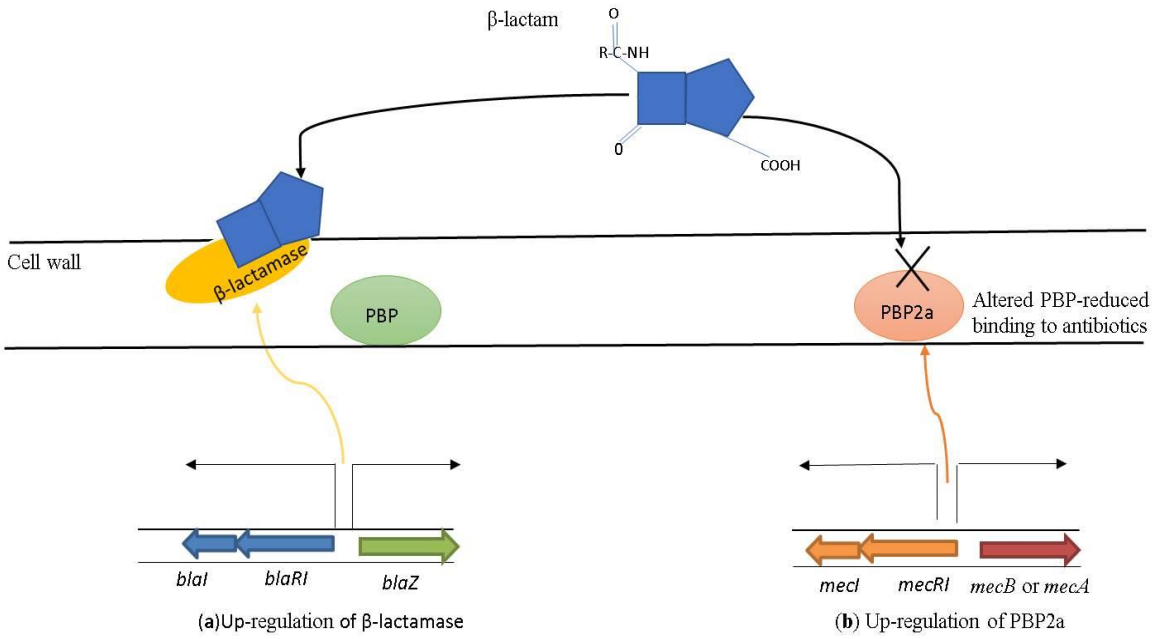


Fig. 4 Mechanisms of β-lactam resistance (**a**) through enzymatic inactivation of antibiotic trough production of β-lactamase achieved with the induction of blaZ gene reported in *S. aureus* (**b**) modification of the antibiotic target site by expression of mecB (*M. caseolyticus* and *M. canis*) or mecA gene (*S. aureus* species) in the presence β-lactam encoding for PBP2a.

1.7.2 The Origin of Methicillin Resistance Gene *mecA*

The first *mecA* gene homolog was encountered in the chromosome of *S. sciuri* strains. (Wu et al., 1996). Later reports identified the *mecA* gene on the chromosome of one of the oldest staphylococcal species, *Staphylococcus fleurettii*, that had an identical nucleotide sequence (99.8%) match with *mecA* gene carried by SCC*mec* on the MRSA chromosome (Tsubakishita et al., 2010b). Therefore, *S. fleurettii mecA* was regarded as the original *mecA*, which was acquired as the methicillin-resistance determinant of the SCC*mec* that converted methicillin susceptible *S. aureus* (MSSA) into MRSA. Whole genome sequencing of the *S. fleurettii* strain SFMP07 indicated the *mecA* gene complex located in the absence of mobile elements like SCC around it, suggested *mecA* as an intrinsic component of the chromosome of *S. fleurettii* species. This may have been an important penicillin binding protein for the ancestral staphylococcal species enabling them to survive under conditions of antibiotic pressure inferred by β-lactams producing environmental microorganisms (Hiramatsu et al., 2013). Phylogenetic distribution of the *mecA* homologs in another study reported *mecA* had been vertically transmitted as an ortholog for some time during the period of initial speciation of the *S. sciuri* group of staphylococcal species (Hiramatsu et al., 2014). However, a loss of methicillin resistance was observed in the descendants. This could be a result of deletion or mutations, which was represented by the emergence of methicillin susceptible *S. aureus*. A phylogenetic time tree has calculated the divergence between *S. sciuri* group of species from the major staphylococcal clade (including *S. aureus* speciss) to be 200-300 million years ago. This corresponds to the geological age of the emergence of mammals. It is suggested that the descendants of the staphylococcal species have lost the *mecA* gene after they had successfully adapted to mammalian hosts, as they became protected from the threat of β-lactam-producing microorganisms, owing to the host's immune system. Later, the introduction of methicillin caused *S. aureus* to regain *mecA* gene from *S. fleurettii* via the SCC*mec* (Hiramatsu et al., 2014).

1.7.3 The *mec* Complex in *Staphylococcus* and *Macrococcus* Species

The SCC*mec*-like elements discovered in *Macrococcus* species were described to differ from the SCC*mec* of the MRSA in terms of their nucleotide sequence and genomic organization (Tsubakishita et al., 2010b). In *Staphylococcus*, the conventional *mec*-gene complex is composed of a gene encoding for methicillin

resistance *mecA* and its regulatory genes *mecR1* (sensor/signal transducer) and *mecI* (transcriptional repressor) (Elements, 2009). Another gene homolog of *mecA, mecC* is also identified in staphylococcal species (García-Álvarez et al., 2011; Harrison et al., 2013a; Harrison et al., 2013b). Unlike the *mecA*, *mecC* is accompanied by *blaZ* homolog encoding for β-lactamase. The *mecC* gene complex (*mecI-mecR1-mecC1-blaZ*) of *S. xylosus* shared structural similarity with the *mec* gene complex found in *M. caseolyticus*. However, this *mecC* complex is not associated with the transposase genes which were initially identified in *M. caseolyticus* species (Baba et al., 2009; Tsubakishita et al., 2010a). Resistance to β-lactams in *S. aureus* is mediated by *blaZ* and *mecA* or *mecC* expression controlled by an inducible proteolytic signal transduction pathway (Zhang et al., 2001; Arêde and Oliveira, 2013).

In contrast, the gene encoding for methicillin resistance in *Macrococcus* is *mecB*, a gene homolog of *mecA*. This was initially identified in *M. caseolyticus* JCSC5402, a strain isolated from animal meat in 2009. The *mecB* complex was designated as $mecA_m$ which was associated with transposon Tn60465 harbored on a 80-kb plasmid (PMCCL2). This plasmid carries multiple resistance genes against β-lactams, macrolides and aminoglycosides (Baba et al., 2009). Curiously, the $mecA_m$ complex structure reported ($mecI_m$-$mecr1_m$-$meca_m$-$blaz_m$) was unique compared to the usual *mec* complex identified in methicillin resistance staphylococci and was probably part of a primordial form of methicillin resistance gene complex. This proposed $mecA_m$ gene complex structure correlates with Matsuhashi's speculation on the historic makeup of the *mecA* gene complex, formed as a result of recombination of *blaZ-blaR1-blaI* operon with *mecA* gene encoding PBP (Song et al., 1987). This plasmid carrying the *mecB* transposon also carries other diverse genes that have blast top-hit entries across several other bacterial genera besides *Macrococcus* and *Staphylococcus*. This seems to have a broad host range and it is speculated that *mecB* is disseminated to diverse bacterial genera via the plasmid (Hiramatsu et al., 2013).

Later, another strain of *M. caseolyticus*, JCSC7096, was reported to contain the same structure of $mecA_m$ gene complex that was accompanied with transposon (Tn60465), harbored on the chromosome. This transposon was bracketed by direct repeats (DR1 and DR2) and right next to this region an SCC element carrying cassette chromosome recombinase (*ccrAB)* genes were also located. A single copy of DRs of the transposon

separated these two elements (Tn60465 and the SCC). This study concluded that a mere deletion of the DRs that were separating the two elements or a mutational inactivation of Tn60465 would be sufficient to combine the two elements forming a new SCC*mec* element (Tsubakishita et al., 2010a). In 2005, a *mecB* carrying SCC*mec* element was identified in the type strain of *M. canis* (KM45013^{T}). This complex was presumed to have integrated as an exogeneous element, which was demarcated at both extremities by direct repeats (DRs) that functioned as integration site sequences (ISSs) (Gomez-Sanz et al., 2015). Like *mecC*, the *mecB* gene was accompanied by a *blaZ* homolog coding for β-lactamase. The phylogenetic analysis of the *mecB* gene suggests it is distantly related to *mecA* with a 61.7% nucleotide identity match. Comparison of the *mecB* complex of macrococci with the class E *mec* gene complex in staphylococci, indicated 57.4% nucleotide identity with *S. xylosus* S04009 and 56.8% to that of *S. aureus* LGA251 strain (García-Álvarez et al., 2011; Harrison et al., 2013a). The *ccr* genes detected in the SCCmec from KM45013^{T} demonstrated closest identity to the ccr gene present in MRSA strain HDE288 (51.6% to *ccrA*4 and 47.3% to *ccrB*4) (Gomez-Sanz et al., 2015). The most recent report has also identified *mecB*-carrying plasmid in *S. aureus* UKM4229 strain sourced from a human nasal-throat swab. This *mecB* gene present in UKM4229 strain is distantly related to that of *M. caseolyticus* JCSC5402 which suggested a possibility of a gene transfer between the two genera (Becker et al., 2018).

Another gene homolog *mecD* has also been identified in *M. caseolyticus* isolates from bovine and canine sources. The *mecD* complex (*mecD-mecRI*$_m$*-mecr1*$_m$) was integrated at the 3′ end of the *rspl* gene. This novel methicillin resistance gene illustrates antimicrobial resistance to all classes of β-lactams which includes anti-MRSA cephalosporins. The *mecD* gene is associated with a site-specific integrase indicating a potential for dissemination (Schwendener et al., 2017). Unlike the *mecB* gene, *mecD* was not accompanied with a *blaZ* homolog. The *mecD* gene of IMD8019 shares 69% nucleotide identity match with *mecB* gene of JCSC5402 strain and 61% nucleotide identity match with *mecC* gene of *S. aureus* LGA251. Fig. 5 illustrates comparison of the *mec* complex structures of *S. aureus* N315 (*mecA* complex), *M. caseolyticus*, JCSC5402, JCSC7096 (*mecB* transmitted as a transposn TN6045), *M. canis* KM45013^{T} (*mecB* transmitted as SCC$_{mec}$ element) and *M. caseolyticus* IMD0819 (*mecD* complex) strains.

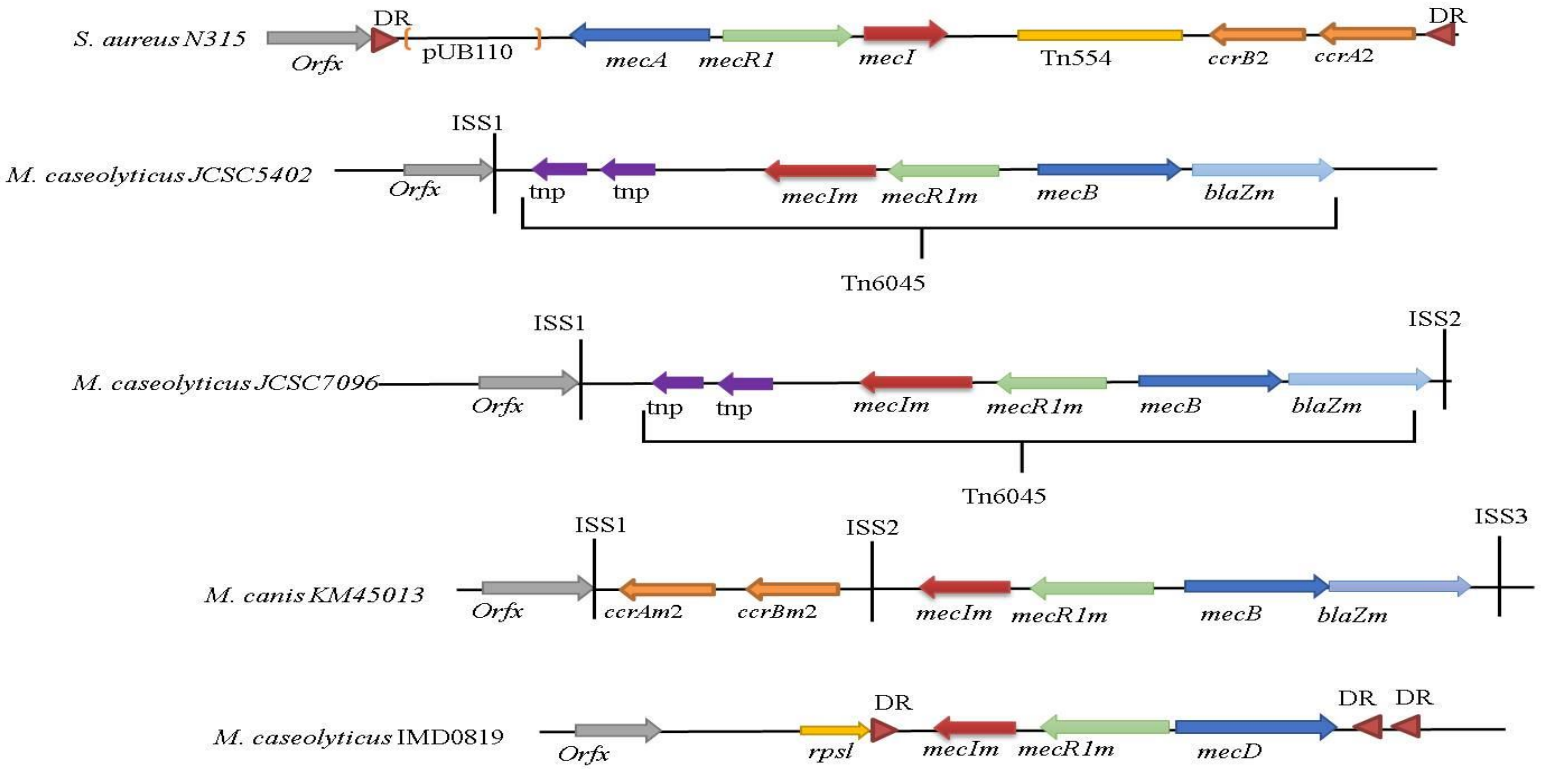


Fig. 5 Genomic *mec* complex structures reported amongst the representative *Staphylococcus* and *Macrococcus* strains. Arrows indicate the genes and their directions of transcription. The *mecA*, *mecB* and *mecD* are illustrated in blue arrows and blazm are illustrated in light blue. The mec regulators *mecR1* and *mecR1m* are represented in green and *mecI* and *mecIm* are represented in red. Open reading frames (orf's) encoding transposase (tnp) are represented in purple or encoding transposon Tn554 is represented in yellow box or ccr gene complex are represented in orange. Direct repeats (DR) are indicated in red (triangles) and Insertion sequences (ISS) are represented by vertical black lines. The *mec* gene complex in *M. caseolyticus* JCSC5402, JCSC7096 is associated with transposases and transmitted as tranposon TN6045. *S. aureus* N315 integrated copy of plasmid pUB110 (indicated in orange brackets) in the proximity of the *mecA* gene. The *mecD* operon in *M. caseolyticus* IMD0819 is integrated into the 3´ end of the *rpsl* gene represented in yellow arrow and grey arrows represent orfx.

1.7.4 Multidrug Resistance Gene *cfr*

The *cfr* gene encodes for methyltransferase that catalyses the methylation of A2503 base in bacterial 23S rRNA (Kehrenberg et al., 2005). This multiresistance gene is described to confers resistance to five chemically unrelated classes of antimicrobials, including phenicols, lincosamides, oxazolidinones, pleuromutilins, and streptogramin A (Long et al., 2006). In addition, decreased susceptibility towards 16-membered macrolides, josamycin and spiramycin is also reported (Smith and Mankin, 2008). The *cfr* gene was initially identified as a chloramphenicol-florfenicol resistant determinant carried on the multiresistance plasmid pSCFS1 isolated from *S. sciuri*, in 2000 (Schwarz et al., 2000). Since then several studies have reported the presence of this gene in staphylococcal species isolated from both animal and human origin. In majority of the cases, the *cfr* gene was found to be located on various plasmids (Mendes et al., 2008; Kehrenberg et al., 2009), whereas in a other cases, it was identified to be present on the chromosome (Kehrenberg et al., 2007; Toh et al., 2007; Locke et al., 2012). The *cfr* gene is not restricted to *Staphylococcus* species, as its presence in other bacteria such as in *Bacillus* species, *Enterococcus faecalis*, *Proteus vulgaris* and *Escherichia coli*, from porcine and bovine origin have also been documented (Dai et al., 2010; Wang et al., 2011; Liu et al., 2012; Wang et al., 2012a). The first report of the presence of plasmids harbouring the *cfr* gene in two strains of *M. caseolyticus* K3 and 207 from porcine origin was in 2012 (Wang et al., 2012b). This report identified the presence of ~7 kb plasmid in the K3 strain and a ~53 kb plasmid present in both 207 and K3 strain. Further investigation of the plasmids with restriction and sequence analysis revealed that the small plasmid of the K3 strain was in fact plasmid pSS-03 that was initially identified in coagulase-negative *Staphylococcus*. The ~53 kb plasmid designated as pJP1 also demonstrated indistinguishable restriction patterns with pJP1 plasmid in *Jeotgalicoccus pinnipedialis* 102. The presence of plasmid pJP1 in strains of *M. caseolyticus* and *J. pinnipedialis* suggested that this plasmid carrying multi drug resistance *cfr* gene can replicate within, and disseminate between different genera of Gram-positive bacteria.

1.7.5 Antibiotic Resistance Phenotype in *Macrococcus* Species

The presence of an alternative penicillin binding protein (PBP2a) can be determined phenotypically through screening for cefotoxin or oxcacillin resistance (Schwendener et al., 2017). The minimum inhibitory concentration for these and a number of other

drugs has been obtained for strains belonging to *M. caseolyticus* and *M. canis* species investigated by Cotting *et* al. The authors identified the occurrence of antibiotic resistance in macrococci from dogs and confirmed the antibiotic resistance by investigating the corresponding resistance mechanisms. These investigations identified the presence of multidrug resistance in a strain of *M. canis* (IMD0218) that was isolated from otitis externa, which exhibited resistance to β-lactams, tetracyclines, aminoglycosides, macrolides and lincosamides. Moreover, the *mecB* was identified in large numbers of strains of *M. canis* and *M. caseolyticus* obtained from healthy dogs as well as in strains isolated from infected sites (rhinitis, otitits externa, dermatitis and mastitis)(Cotting et al., 2017). Another study has also investigated the correlation between the phenotype and genotype of resistance found in *M. caseolyticus* strains (Schwendener et al., 2017).

1.8 Occurrence of *M. caseolyticus* in Food Systems

In early reports of *M. caseolyticus,* then referred to as *Micrococcus caseolyticus,* the organism was documented to be present as part of the secondary flora of cheese and its role in ripening and flavour development was documented (Bhowmik and Marth, 1988; 1990). Flavour development of fermented foods involves a complex network of metabolic reactions, of which proteolysis plays a major role. In dairy fermentation, the proteolytic cascade begins with casein degradation by extracellular proteinases. The protease and peptidase activity of certain strains of this organism were investigated by Moreno and Kosikowski. The activity of enzymes from *M. caseolyticus* and other micrococci on β-casein that lead to the production of short peptides and amino acids suggested that *M. caseolyticus* was capable of producing the substrates for flavour compound production during cheese ripening (Moreno and Kosikowski, 1973). The ability of *M. caseolyticus* to produce an extracellular caseinolytic enzyme and isolation of this extracellular enzyme has been described by Desmazeaud and Hermier (Desmazeaud and Hermier, 1968). These extracellular enzymes were later exploited by a French company, Roussel-Uclaf, who incorporated this protease in to liposomes for the production of a commercial enzyme called Rulactine (Yoovidhya et al., 1986). This metalloproteinase from *M. caseolyticus* was then used in the production of Saint Pauli cheese, where its activity intensified proteolysis of milk and altered the texture of the curd (Alkalaf et al., 1987). Further investigations of adding Rulactine to the starter culture indicated significant acceleration of the cheese ripening process (Piard et al., 1986; Alkalaf et al., 1987; Alkhalaf et al., 2006). The technological applications of enzymes extracted from this organism have been described in a number of patents. A method for producing a polymeric enzyme from a culture of *M. caseolyticus,* used in the production of aspartame, is described by the inventor Paul and Duchiron (Paul et al., 1990). Further exploitation of enzymes from *M. caseolyticus* and addition to milk had increased cheese capacity of milk used for the production of uncooked or half-cooked pressed paste cheeses (Barthelemy and Desmazeaud, 1986). *M. caseolyticus* was also used to produce novel cheese products and patents describe the technical contribution of the organism in the overall development of the desirable body and flavour of elastic cheese and low fat ripened skim milk cheese products (Hargrove and Mcdonough, 1964; Kasik and Luksas, 1971). The more conventional role of *M.*

caseolyticus in the flavour development of Cantonese sausage has also been investigated (Wu et al., 2009).

The ability of this organism to produce bioactive peptides was also investigated in one study where *M. caseolyticus* was screened for its inhibitory potential against Shiga toxin-producing *Escherichia coli* (STEC) on a model cheese curd. Results indicated this species to be amongst the most inhibitory, indicating it's potential role in future bio-preservation (Callon et al., 2016). Another species from this genus, *M. bovicus* was also investigated for the development of an antibacterial technology (Abdel-Aziz et al., 2015). In this study, *M. bovicus* was used in the biosynthesis of nano-scaling silver (NSAg). This nanocomposite was shown to have antimicrobial activity towards Gram-negative and Gram-positive bacteria as well as certain fungi.

1.9 Concluding Remarks

The studies described in this review provide a basic understanding of the *Macrococcus* species in terms of their phenotypic capabilities, genetic diversity, evolutionary history with *Staphylococcus,* and antibiotic resistance. Despite their close relationship with *Staphylococcus* from an evolutionary perspective, the species from the *Macrococcus* genus have largely been described as avirulent. Despite this, a few cases describing infections associated with the hemolytic *M. canis* species have been reported (Cotting et al., 2017). However, due to the lack of current knowledge on virulence factors, the ability of this species to play a role in opportunistic infection cannot be fully determined and thus, merits further investigations. With the advancement in sequencing technology, it has become easier to monitor current and historic events regarding the emergence and spread of antimicrobial resistance. In recent publications, researchers have used comparative genomics to trace back the conventional *mecA* gene to *S. fleuretti* species, emphasizing that antibiotic resistance in this species pre-dates the introduction of methicillin, whereas reports have discovered a primordial form of the methicillin resistance gene complex in *M. caseolyticus* (Baba et al., 2009; Hiramatsu et al., 2013; Hiramatsu et al., 2014). The presence of novel homologs of methicillin resistance genes in *Macrococcus* from animal origin are emphasizing that adaptations conferring antibiotic resistance towards broad spectrum β-lactams, and in some cases all classes of β-lactams that are widely used in human and veterinary medicine. Novel levels of risk associated with the transfer of these methicillin resistance genes to pathogenic staphylococci can hinder future therapeutic options to treat infections caused by MRSA (Becker et al., 2018). The limited knowledge available in terms of host preference of the known *Macrococcus* species and its presence in food, indicates similar patterns to that of staphylococci (Evans, 1916; Kümmel et al., 2016; Cotting et al., 2017). It is likely that these evolutionary close relatives share similar ecological niches and therefore a possibility of transmission of such mobile genetic elements to staphylococci is possible. Indeed, reports have indicated a possible transfer of the *mecB* resistance plasmid of *M. caseolyticus* to *S. aureus,* where other reports have identified a transfer of a staphylococcal multidrug resistant *cfr* gene to *M. caseolyticus* species (Wang et al., 2011). The possible transfer of these resistance plasmids through the food chain, to commensal and pathogenic bacteria of humans, is a cause for concern. Further studies are required to understand

the transfer mechanisms (inter- and intra-species) and detailed nature of the *mec* gene complex in *Macrococcus* species, and its prevalence in humans and animals. This information could enable better understanding of the vicious cycle of dissemination of antimicrobial resistance via these mobile elements, as well as in implementing better control measures for the future.

1.10 Acknowledgements

Thanks to my supervisor Dr. Olivia McAuliffe for the suggestions for writing this review. This research is funded by Teagasc (ref. 6697) and the Teagasc Walsh Fellowship programme (ref. 2015055).

References

Abdel-Aziz, M.S., Abou-El-Sherbini, K.S., Hamzawy, E.M.A., Amr, M.H.A., and El-Dafrawy, S. (2015). Green Synthesis of Silver Nano-particles by *Macrococcus bovicus* and Its Immobilization onto Montmorillonite Clay for Antimicrobial Functionality. *Applied Biochemistry and Biotechnology* 176**,** 2225-2241.

Abdelnour, A., Arvidson, S., Bremell, T., Ryden, C., and Tarkowski, A. (1993). The accessory gene regulator (*agr*) controls *Staphylococcus aureus* virulence in a murine arthritis model. *Infection and Immunity* 61**,** 3879-3885.

Alkalaf, W., Vassal, L., Desmazeaud, M., and Gripon, J. (1987). Use of Rulactine as ripening agent in semi-hard cheese. *Lait (France)*.

Alkhalaf, W., Piard, J.-C., El Soda, M., Gripon, J.C., Desmazeaud, M., and Vassal, L. (2006). *Liposomes as Proteinase Carriers for the Accelerated Ripening of Saint-Paulin Type Cheese.*

Allenby, N.E., O'connor, N., Prágai, Z., Ward, A.C., Wipat, A., and Harwood, C.R. (2005). Genome-wide transcriptional analysis of the phosphate starvation stimulon of *Bacillus subtilis*. *Journal of Bacteriology* 187**,** 8063-8080.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Resources* 25**,** 3389-3402.

Ambler, R. (1980). The structure of β-lactamases. *Phil. Trans. R. Soc. Lond. B* 289**,** 321-331.

Arahal, D.R. (2014). "Chapter 6 - Whole-Genome Analyses: Average Nucleotide Identity," in *Methods in Microbiology,* eds. M. Goodfellow, I. Sutcliffe & J. Chun. Academic Press), 103-122.

Arêde, P., and Oliveira, D.C. (2013). Proteolysis of *mec*A repressor is essential for expression of methicillin resistance by *Staphylococcus aureus*. *Antimicrobial agents and chemotherapy* 57**,** 2001-2002.

Baba, T., Kuwahara-Arai, K., Uchiyama, I., Takeuchi, F., Ito, T., and Hiramatsu, K. (2009). Complete genome sequence of *Macrococcus caseolyticus* strain JSCS5402, reflecting the ancestral genome of the human-pathogenic staphylococci. *Journal of Bacteriology* 191**,** 1180-1190.

Baba, T., Takeuchi, F., Kuroda, M., Ito, T., Yuzawa, H., and Hiramatsu, K. (2003). The genome of Staphylococcus aureus. *The Staphylococcus aureus: Molecular and Clinical Aspects***,** 66-153.

Baird-Parker, A. (1963). A classification of micrococci and staphylococci based on physiological and biochemical tests. *Microbiology* 30**,** 409-427.

Baker, J.S. (1984). Comparison of various methods for differentiation of staphylococci and micrococci. *Journal of clinical microbiology* 19**,** 875-879.

Ballard, D., Kloos, W.E., Mcdowell, C.I., Cole, E.M., L., B., and & Webster, J.A. (1995). Rediscovery of *Staphylococcus caseolyticus* and description of the related species group, new species isolated from cattle, horses, and food. *In Abstracts of the 95th General Meeting of the American Society for Microbiology 1995***,** abstract R-12, pp. 480.

Barbe, V., Cruveiller, S., Kunst, F., Lenoble, P., Meurice, G., Sekowska, A., Vallenet, D., Wang, T., Moszer, I., Medigue, C., and Danchin, A. (2009). From a consortium sequence to a unified sequence: the *Bacillus subtilis* 168 reference genome a decade later. *Microbiology* 155**,** 1758-1775.

Barthelemy, P., and Desmazeaud, M. (1986). "Process for preparing cheese and product produced". Google Patents).

Becker, K., Van Alen, S., Idelevich, E.A., Schleimer, N., Seggewiß, J., Mellmann, A., Kaspar, U., and Peters, G. (2018). Plasmid-Encoded Transferable mecB-Mediated Methicillin Resistance in *Staphylococcus aureus*. *Emerging infectious diseases* 24**,** 242.

Belcheva, A., and Golemi-Kotra, D. (2008). A Close-up View of the VraSR Two-component System A mediator of *Staphylococcus aureus* Response to Cell Wall Damage. *Journal of Biological Chemistry* 283**,** 12354-12364.

Berube, B.J., and Wardenburg, J.B. (2013). *Staphylococcus aureus* α-toxin: nearly a century of intrigue. *Toxins* 5**,** 1140-1166.

Bhowmik, T., and Marth, E.H. (1988). Protease and peptidase activity of *Micrococcus* species. *Journal of Dairy Science* 71**,** 2358-2365.

Bhowmik, T., and Marth, E.H. (1990). Rote of *Micrococcus* and *Pediococcus* species in cheese ripening: A Review1. *Journal of Dairy Science* 73**,** 859-866.

Billroth, T. (1874). Coccobacteria septica. *Georg Reimer, Berlin* 240.

Boháček, J., Kocur, M., and Martinec, T. (1967). DNA base composition and taxonomy of some micrococci. *Microbiology* 46**,** 369-376.

Bouchet, V., Huot, H., and Goldstein, R. (2008). Molecular genetic basis of ribotyping. *Clinical microbiology reviews* 21**,** 262-273.

Brawand, S.G., Cotting, K., Gómez-Sanz, E., Collaud, A., Thomann, A., Brodard, I., Rodriguez-Campos, S., Strauss, C., and Perreten, V. (2017). *Macrococcus canis* sp. nov., a skin bacterium associated with infections in dogs. *International Journal of Systematic and Evolutionary Microbiology* 67**,** 621-626.

Brawand, S.G., Rychener, L., Schwendener, S., Pantůček, R., and Perreten, V. (2018). Complete Genome Sequence of the Type Strain of *Macrococcus canis*. *Genome Announcements* 6**,** e01507-01517.

Breidt, F., and Stewart, G.C. (1986). Cloning and expression of the phospho-beta-galactosidase gene of *Staphylococcus aureus* in *Escherichia coli*. *Journal of Bacteriology* 166**,** 1061-1066.

Callon, C., Arliguie, C., and Montel, M.C. (2016). Control of Shigatoxin-producing Escherichia coli in cheese by dairy bacterial strains. *Food Microbiology* 53**,** 63-70.

Cheng, K.C., Katz, S.R., Lin, A.Y., Xin, X., and Ding, Y. (2016). "Chapter Four - Whole-Organism Cellular Pathology: A Systems Approach to Phenomics," in *Advances in Genetics*, ed. N.S. Foulkes. Academic Press), 89-115.

Collins, M., and Jones, D. (1981). Distribution of isoprenoid quinone structural types in bacteria and their taxonomic implication. *Microbiological reviews* 45**,** 316.

Colston, S.M., Fullmer, M.S., Beka, L., Lamy, B., Gogarten, J.P., and Graf, J. (2014). Bioinformatic genome comparisons for taxonomic and phylogenetic assignments using Aeromonas as a test case. *MBio* 5**,** e02136-02114.

Cooney, J., Kienle, Z., Foster, T., and O'toole, P. (1993). The gamma-hemolysin locus of *Staphylococcus aureus* comprises three linked genes, two of which are identical to the genes for the F and S components of leukocidin. *Infection and Immunity* 61**,** 768-771.

Cotting, K., Strauss, C., Rodriguez-Campos, S., Rostaher, A., Fischer, N.M., Roosje, P.J., Favrot, C., and Perreten, V. (2017). *Macrococcus canis* and *M. caseolyticus* in dogs: occurrence, genetic diversity and antibiotic resistance. *Veterinary Dermatology*.

Dai, L., Wu, C.-M., Wang, M.-G., Wang, Y., Wang, Y., Huang, S.-Y., Xia, L.-N., Li, B.-B., and Shen, J.-Z. (2010). First report of the multidrug resistance gene *cfr* and the phenicol resistance gene fexA in a *Bacillus* strain from swine feces. *Antimicrobial agents and chemotherapy* 54**,** 3953-3955.

Davies, J., and Davies, D. (2010). Origins and evolution of antibiotic resistance. *Microbiology and molecular biology reviews* 74**,** 417-433.

De La Fuente, R., Suarez, G., Santa Quiteria, J.R., Meugnier, H., Bes, M., Freney, J., and Fleurette, J. (1992). Identification of coagulase negative staphylococci isolated from lambs as *Staphylococcus caseolyticus*. *Comparative immunology, microbiology and infectious diseases* 15**,** 47-52.

Desmazeaud, M., and Hermier, J. (Year). "Isolement, purification et propriétés d'une protéase exocellulaire de *Micrococcus caseolyticus*", in: *Annales de Biologie Animale Biochimie Biophysique*: EDP Sciences), 565-577.

Doolittle, W.F., and Pace, N.R. (1971). Transcriptional organization of the ribosomal RNA cistrons in *Escherichia coli*. *Proceedings of the National Academy of Sciences* 68**,** 1786-1790.

Egeter, O., and Brückner, R. (1995). Characterization of a genetic locus essential for maltose-maltotriose utilization in *Staphylococcus xylosus. Journal of Bacteriology* 177**,** 2408-2415.

Elements, I.W.G.O.T.C.O.S.C.C. (2009). Classification of staphylococcal cassette chromosome mec (SCCmec): guidelines for reporting novel SCCmec elements. *Antimicrobial Agents and Chemotherapy* 53**,** 4961-4967.

Endl, J., Seidl, H., Fiedler, F., and Schleider, K. (1983). Chemical composition and structure of cell wall teichoic acids of staphylococci. *Archives of microbiology* 135**,** 215-223.

Evans, A.C. (1916). The bacteria of milk freshly drawn from normal udders. *The Journal of Infectious Diseases***,** 437-476.

Evans, J.B., Bradford Jr, W., and Niven Jr, C. (1955). Comments concerning the taxonomy of the genera *Micrococcus* and *Staphylococcus*. *International Journal of Systematic and Evolutionary Microbiology* 5**,** 61-66.

Falk, D., and Guering, S. (1983). Differentiation of *Staphylococcus* and *Micrococcus* spp. with the Taxo A bacitracin disk. *Journal of clinical microbiology* 18**,** 719-721.

Faller, A., and Schleifer, K.-H. (1981). Modified oxidase and benzidine tests for separation of staphylococci from micrococci. *Journal of Clinical Microbiology* 13**,** 1031-1035.

Faller, A.H., Götz, F., and Schleifer, K.-H. (1980). Cytochrome-patterns of staphylococci and micrococci and their taxonomic implications. *Zentralblatt für Bakteriologie: I. Abt. Originale C: Allgemeine, angewandte und ökologische Mikrobiologie* 1**,** 26-39.

Fedtke, I., Kamps, A., Krismer, B., and Götz, F. (2002). The nitrate reductase and nitrite reductase operons and the narT gene of *Staphylococcus carnosus* are positively controlled by the novel two-component system NreBC. *Journal of Bacteriology* 184**,** 6624-6634.

Fischer, R., Eisermann, R., Reiche, B., and Hengstenberg, W. (1989). Cloning, sequencing and overexpression of the mannitol-specific enzyme-III-encoding gene of *Staphylococcus carnosus*. *Gene* 82**,** 249-257.

Fischer, R., and Hengstenberg, W. (1992). Mannitol-specific enzyme II of the phosphoenolpyruvate-dependent phosphotransferase system of *Staphylococcus carnosus*. *The FEBS Journal* 204**,** 963-969.

Flügge, C. (1886). "Die Mikroorganismen Leipzig". Germany).

Fournier, B., and Hooper, D.C. (2000). A new two-component regulatory system involved in adhesion, autolysis, and extracellular proteolytic activity of *Staphylococcus aureus*. *Journal of Bacteriology* 182**,** 3955-3964.

Fournier, B., Klier, A., and Rapoport, G. (2001). The two-component system ArlS–ArlR is a regulator of virulence gene expression in *Staphylococcus aureus*. *Molecular Microbiology* 41**,** 247-261.

Fox, G.C.-A., Stackebrandt, E., Hespell, R., Gibson, J., Maniloff, J., Dyer, T., Wolfe, R., Balch, W., Tanner, R., and Magrum, L. (1980). The phylogeny of prokaryotes. *Science* 209**,** 457-463.

García-Álvarez, L., Holden, M.T., Lindsay, H., Webb, C.R., Brown, D.F., Curran, M.D., Walpole, E., Brooks, K., Pickard, D.J., and Teale, C. (2011). Meticillin-resistant *Staphylococcus aureus* with a novel *mec*A homologue in human and bovine populations in the UK and Denmark: a descriptive study. *The Lancet infectious diseases* 11**,** 595-603.

Garrity, G.M., and Holt, J.G. (2001). "The road map to the manual," in *Bergey's Manual® of Systematic Bacteriology*. Springer), 119-166.

Gel'man, N.S., Lukoyanova, M.A., and Ostrovskii, D.N. (1967). "Oxidative Phosphorylation in Bacteria," in *Respiration and Phosphorylation of Bacteria*. (Boston, MA: Springer US), 161-192.

George, C.G., and Kloos, W.E. (1994). Comparison of the Sma I-digested chromosomes of *Staphylococcus epidermidis* and the closely related species *Staphylococcus capitis* and *Staphylococcus caprae*. *International Journal of Systematic and Evolutionary Microbiology* 44**,** 404-409.

George, M.G., Julia, A., and Timothy, G. (2005). "Bergey's manual of systematic bacteriology," in *Volume 2: The Proteobacteria*. Springer-Verlag, Berlin), 735-769.

Gomez-Sanz, E., Schwendener, S., Thomann, A., Gobeli Brawand, S., and Perreten, V. (2015). First Staphylococcal Cassette Chromosome *mec* Containing a *mecB*-Carrying Gene Complex Independent of Transposon Tn6045 in a *Macrococcus caseolyticus* Isolate from a Canine Infection. *Antimicrobial Agents and Chemotherapy* 59**,** 4577-4583.

Götz, F., Bannerman, T., and Schleifer, K.-H. (2006). "The Genera *Staphylococcus* and *Macrococcus*," in *The Prokaryotes: Volume 4: Bacteria: Firmicutes, Cyanobacteria*, eds. M. Dworkin, S. Falkow, E. Rosenberg, K.-H. Schleifer & E. Stackebrandt. (New York, NY: Springer US), 5-75.

Götz, F., and Mayer, S. (2013). Both terminal oxidases contribute to fitness and virulence during organ-specific *Staphylococcus aureus* colonization. *MBio* 4**,** e00976-00913.

Götz, F., Nürnberger, E., and Schleifer, K. (1979). Distribution of class-I and class-II D-fructose 1, 6-biphosphate aldolases in various staphylococci, peptococci and micrococci. *FEMS Microbiology Letters* 5**,** 253-257.

Graham, J.H., Hodge, N.C., and Morton, J.B. (1995). Fatty Acid methyl ester profiles for characterization of glomalean fungi and their endomycorrhizae. *Applied and Environmental Microbiology* 61**,** 58-64.

Grimont, F., and Grimont, P. (Year). "Ribosomal ribonucleic acid gene restriction patterns as potential taxonomic tools", in: *Annales de l'Institut Pasteur/Microbiologie*: Elsevier), 165-175.

Hammer, N.D., and Skaar, E.P. (2011). Molecular mechanisms of *Staphylococcus aureus* iron acquisition. *Annual review of microbiology* 65**,** 129-147.

Hargrove, R.E., and Mcdonough, F.E. (1964). "Process of making low-fat ripened skim milk cheese, US3156568A". Google Patents).

Harrison, E.M., Paterson, G.K., Holden, M.T., Ba, X., Rolo, J., Morgan, F.J., Pichon, B., Kearns, A., Zadoks, R.N., and Peacock, S.J. (2013a). A novel hybrid SCC mec-mecC region in *Staphylococcus sciuri*. *Journal of Antimicrobial Chemotherapy* 69**,** 911-918.

Harrison, E.M., Paterson, G.K., Holden, M.T., Morgan, F.J., Larsen, A.R., Petersen, A., Leroy, S., De Vliegher, S., Perreten, V., and Fox, L.K. (2013b). A *Staphylococcus xylosus* isolate with a new mecC allotype. *Antimicrobial agents and chemotherapy* 57**,** 1524-1528.

Hata, D.J. (2010). "Molecular Methods for Identification and Characterization of *Acinetobacter* spp," in *Molecular Diagnostics*. Elsevier), 313-326.

Hiramatsu, K., Ito, T., Tsubakishita, S., Sasaki, T., Takeuchi, F., Morimoto, Y., Katayama, Y., Matsuo, M., Kuwahara-Arai, K., and Hishinuma, T. (2013). Genomic basis for methicillin resistance in *Staphylococcus aureus*. *Infection & chemotherapy* 45**,** 117-136.

Hiramatsu, K., Katayama, Y., Matsuo, M., Sasaki, T., Morimoto, Y., Sekiguchi, A., and Baba, T. (2014). Multi-drug-resistant *Staphylococcus aureus* and future chemotherapy. *Journal of Infection and Chemotherapy* 20**,** 593-601.

Howell, A., Dubrac, S., Noone, D., Varughese, K.I., and Devine, K. (2006). Interactions between the YycFG and PhoPR two-component systems in *Bacillus subtilis*: The PhoR kinase phosphorylates the non-cognate YycF response regulator upon phosphate limitation. *Molecular Microbiology* 59**,** 1199-1215.

Hulett, F.M. (2002). "The pho regulon," in *Bacillus subtilis and Its Closest Relatives*. American Society of Microbiology), 193-201.

Iglesias, A.A., and Preiss, J. (1992). Bacterial glycogen and plant starch biosynthesis. *Biochemistry and Molecular Biology Education* 20**,** 196-203.

Janda, J.M., and Abbott, S.L. (2007). 16S rRNA gene sequencing for bacterial identification in the diagnostic laboratory: pluses, perils, and pitfalls. *Journal of Clinical Microbiology* 45**,** 2761-2764.

Jeffreys, A.J., Wilson, V., Neumann, R., and Keyte, J. (1988). Amplification of human minisatellites by the polymerase chain reaction: towards DNA fingerprinting of single cells. *Nucleic Acids Research* 16**,** 10953-10971.

Jeffries, L., Cawthorne, M., Harris, M., Cook, B., and Diplock, A. (1968). Menaquinone determination in the taxonomy of *Micrococcaceae*. *Microbiology* 54**,** 365-380.

Kamps, A., Achebach, S., Fedtke, I., Unden, G., and Götz, F. (2004). Staphylococcal NreB: an O2-sensing histidine protein kinase with an O2-labile iron–sulphur cluster of the FNR type. *Molecular microbiology* 52**,** 713-723.

Kasik, R.L., and Luksas, A.J. (1971). "Preparation of an elastic cheese product". Google Patents).

Kaufmann, M.E. (1998). "Pulsed-field gel electrophoresis," in *Molecular Bacteriology*. Springer), 33-50.

Kehrenberg, C., Aarestrup, F.M., and Schwarz, S. (2007). IS21-558 insertion sequences are involved in the mobility of the multiresistance gene *cfr*. *Antimicrobial agents and chemotherapy* 51**,** 483-487.

Kehrenberg, C., Cuny, C., Strommenger, B., Schwarz, S., and Witte, W. (2009). Methicillin-resistant and-susceptible *Staphylococcus aureus* strains of clonal lineages ST398 and ST9 from swine carry the multidrug resistance gene cfr. *Antimicrobial agents and chemotherapy* 53**,** 779-781.

Kehrenberg, C., Schwarz, S., Jacobsen, L., Hansen, L.H., and Vester, B. (2005). A new mechanism for chloramphenicol, florfenicol and clindamycin resistance: methylation of 23S ribosomal RNA at A2503. *Molecular microbiology* 57**,** 1064-1073.

Kilpper, R., Buhl, U., and Schleifer, K.H. (1980). Nucleic acid homology studies between *Peptococcus saccharolyticus* and various anaerobic and facultative anaerobic Gram-positive cocci. *FEMS Microbiology Letters* 8**,** 205-210.

Kloos, W.E., Ballard, D.N., George, C.G., Webster, J.A., Hubner, R.J., Ludwig, W., Schleifer, K.H., Fiedler, F., and Schubert, K. (1998a). Delimiting the genus *Staphylococcus* through description of *Macrococcus caseolyticus* gen. nov., comb. nov. and *Macrococcus equipercicus* sp. nov., and *Macrococcus bovicus* sp. no. and *Macrococcus carouselicus* sp. nov. *International Journal of Systematic and Evolutionary Microbiology* 48 Pt 3**,** 859-877.

Kloos, W.E., George, C.G., Olgiate, J.S., Van Pelt, L., Mckinnon, M.L., Zimmer, B.L., Muller, E., Weinstein, M.P., and Mirrett, S. (1998b). *Staphylococcus hominis* subsp. *novobiosepticus* subsp. nov., a novel trehalose-and N-acetyl-D-glucosamine-negative, novobiocin-and multiple-antibiotic-resistant subspecies isolated from human blood cultures. *International Journal of Systematic and Evolutionary Microbiology* 48**,** 799-812.

Kolar, S.L., Nagarajan, V., Oszmiana, A., Rivera, F.E., Miller, H.K., Davenport, J.E., Riordan, J.T., Potempa, J., Barber, D.S., and Koziel, J. (2011). NsaRS is a cell-envelope-stress-sensing two-component system of *Staphylococcus aureus*. *Microbiology* 157**,** 2206-2219.

Krewulak, K.D., and Vogel, H.J. (2008). Structural biology of bacterial iron uptake. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1778**,** 1781-1804.

Kümmel, J., Stessl, B., Gonano, M., Walcher, G., Bereuter, O., Fricker, M., Grunert, T., Wagner, M., and Ehling-Schulz, M. (2016). *Staphylococcus aureus* Entrance into the Dairy Chain: Tracking *S. aureus* from Dairy Cow to Cheese. *Frontiers in Microbiology* 7**,** 1603.

Kwok, A.Y., and Chow, A.W. (2003). Phylogenetic study of *Staphylococcus* and *Macrococcus* species based on partial hsp60 gene sequences. *International journal of Systematic and Evolutionary Microbiology* 53**,** 87-92.

Leonardi, R., and Jackowski, S. (2007). Biosynthesis of pantothenic acid and coenzyme A. *EcoSal Plus* 2.

Liu, X., and Taber, H.W. (1998). Catabolite regulation of the *Bacillus subtilis ctaBCDEF* gene cluster. *Journal of Bacteriology* 180**,** 6154-6163.

Liu, Y., Wang, Y., Wu, C., Shen, Z., Schwarz, S., Du, X.-D., Dai, L., Zhang, W., Zhang, Q., and Shen, J. (2012). First report of the multidrug resistance gene *cfr* in *Enterococcus faecalis* of animal origin. *Antimicrobial agents and chemotherapy* 56**,** 1650-1654.

Locke, J.B., Rahawi, S., Lamarre, J., Mankin, A.S., and Shaw, K.J. (2012). Genetic environment and stability of cfr in methicillin-resistant *Staphylococcus aureus* CM05. *Antimicrobial agents and chemotherapy* 56**,** 332-340.

Long, K.S., Poehlsgaard, J., Kehrenberg, C., Schwarz, S., and Vester, B. (2006). The Cfr rRNA methyltransferase confers resistance to phenicols, lincosamides, oxazolidinones, pleuromutilins, and streptogramin A antibiotics. *Antimicrobial agents and chemotherapy* 50**,** 2500-2505.

Lory, S. (2014). "The Family *Staphylococcaceae*," in *The Prokaryotes: Firmicutes and Tenericutes,* eds. E. Rosenberg, E.F. Delong, S. Lory, E. Stackebrandt & F. Thompson. (Berlin, Heidelberg: Springer Berlin Heidelberg), 363-366.

Ludwig, W., Schleifer, K.-H., Fox, G.E., Seewaldt, E., and Stackebrandt, E. (1981). A phylogenetic analysis of staphylococci, *Peptococcus saccharolyticus* and *Micrococcus mucilaginosus*. *Microbiology* 125**,** 357-366.

Macheboeuf, P., Contreras-Martel, C., Job, V., Dideberg, O., and Dessen, A. (2006). Penicillin binding proteins: key players in bacterial cell cycle and drug resistance processes. *FEMS microbiology reviews* 30**,** 673-691.

Manna, A.C., Bayer, M.G., and Cheung, A.L. (1998). Transcriptional analysis of different promoters in the sar locus in *Staphylococcus aureus*. *Journal of Bacteriology* 180**,** 3828-3836.

Mannerová, S., Pantůček, R., Doškař, J., Švec, P., Snauwaert, C., Vancanneyt, M., Swings, J., and Sedláček, I. (2003). *Macrococcus brunensis* sp. nov., *Macrococcus hajekii* sp. nov. and *Macrococcus lamae* sp. nov., from the skin of llamas. *International Journal of Systematic and Evolutionary Microbiology* 53**,** 1647-1654.

Marmur, J., and Doty, P. (1962). Determination of the base composition of deoxyribonucleic acid from its thermal denaturation temperature. *Journal of Molecular Biology* 5**,** 109-118.

Mashruwala, A.A., and Boyd, J.M. (2017). The *Staphylococcus aureus* SrrAB regulatory system modulates hydrogen peroxide resistance factors, which imparts protection to aconitase during aerobic growth. *PloS one* 12**,** e0170283.

Mendes, R.E., Deshpande, L.M., Castanheira, M., Dipersio, J., Saubolle, M.A., and Jones, R.N. (2008). First report of cfr-mediated resistance to linezolid in human staphylococcal clinical isolates recovered in the United States. *Antimicrobial agents and chemotherapy* 52**,** 2244-2246.

Mesbah, N.M., Whitman, W.B., and Mesbah, M. (2011). "14 - Determination of the G+C Content of Prokaryotes," in *Methods in Microbiology,* eds. F. Rainey & A. Oren. Academic Press), 299-324.

Mitreva, M. (2017). "8 - The Microbiome in Infectious Diseases A2 - Cohen, Jonathan," in *Infectious Diseases (Fourth Edition),* eds. W.G. Powderly & S.M. Opal. Elsevier), 68-74.e62.

Monedero, V., Revilla-Guarinos, A., and Zuniga, M. (2017). "Physiological Role of Two-Component Signal Transduction Systems in Food-Associated Lactic Acid Bacteria," in *Advances in applied microbiology*. Elsevier), 1-51.

Moreno, V., and Kosikowski, F. (1973). Degradation of β-casein by micrococcal cell-free preparations. *Journal of Dairy Science* 56**,** 33-38.

Moser, B.A., Becnel, J.J., Undeen, A.H., Hodge, N.C., and Patterson, R.S. (1996). Fatty acid analysis of Microsporidia. *Journal of Invertebrate Pathology* 68**,** 146-151.

Mullis, K., Faloona, F., Scharf, S., Saiki, R., Horn, G., and Erlich, H. (Year). "Specific enzymatic amplification of DNA in vitro: the polymerase chain reaction", in: *Cold Spring Harbor symposia on quantitative biology*: Cold Spring Harbor Laboratory Press), 263-273.

Nahaie, M., Goodfellow, M., Minnikin, D., and Hajek, V. (1984). Polar lipid and isoprenoid quinone composition in the classification of *Staphylococcus*. *Microbiology* 130**,** 2427-2437.

Ogston, A. (1881). Report upon micro-organisms in surgical diseases. *British medical journal* 1**,** 369. b362.

Ogston, A. (1882). *Micrococcus* poisoning. *Journal of Anatomy and Physiology* 16**,** 526.

Olive, D.M., and Bean, P. (1999). Principles and applications of methods for DNA-based typing of microbial organisms. *Journal of Clinical Microbiology* 37**,** 1661-1669.

Oskouian, B., and Stewart, G.C. (1990). Repression and catabolite repression of the lactose operon of *Staphylococcus aureus*. *Journal of Bacteriology* 172**,** 3804-3812.

Parker, J. (2001). "Base Composition A2 - Brenner, Sydney," in *Encyclopedia of Genetics*, ed. J.H. Miller. (New York: Academic Press), 192.

Paul, F., Duchiron, F., and Monsan, P. (1990). "Enzyme, its method of production and its application to the preparation of methyl N-(L-aspartyl-1) L-phenylalaninate". Google Patents).

Peacock, S.J., and Paterson, G.K. (2015). Mechanisms of Methicillin Resistance in *Staphylococcus aureus*. *Annual Review of Biochemistry* 84**,** 577-601.

Piard, J., El Soda, M., Alkhalaf, W., Rousseau, M., Desmazeaud, M., Vassal, L., and Gripon, J. (1986). Acceleration of cheese ripening with liposome-entrapped proteinase. *Biotechnology letters* 8**,** 241-246.

Priest, F.G., and Goodfellow, M. (2012). *Applied microbial systematics.* Springer Science & Business Media.

Rammelkamp, C.H., and Maxon, T. (1942). Resistance of *Staphylococcus aureus* to the Action of Penicillin. *Proceedings of the Society for Experimental Biology and Medicine* 51**,** 386-389.

Randazzo, C.L., Torriani, S., Akkermans, A.D., De Vos, W.M., and Vaughan, E.E. (2002). Diversity, dynamics, and activity of bacterial communities during production of an artisanal Sicilian cheese as evaluated by 16S rRNA analysis. *Applied and Environmental Microbiology* 68**,** 1882-1892.

Reiche, B., Frank, R., Deutscher, J., Meyer, N., and Hengstenberg, W. (1988). Staphylococcal phosphoenolpyruvate-dependent phosphotransferase system: purification and characterization of the mannitol-specific enzyme IIImtl of *Staphylococcus aureus* and *Staphylococcus carnosus* and homology with the enzyme IImtl of *Escherichia coli*. *Biochemistry* 27**,** 6512-6516.

Rosenbach, A.J.F. (1884). *Mikroorganismen bei den Wundinfektionskrankheiten des Menschen.* Bergmann.

Rosey, E., Oskouian, B., and Stewart, G. (1991). Lactose metabolism by *Staphylococcus aureus*: characterization of *lacABCD,* the structural genes of the tagatose 6-phosphate pathway. *Journal of Bacteriology* 173**,** 5992-5998.

Rossello-Mora, R., Urdiain, M., and Lopez-Lopez, A. (2011). "DNA–DNA Hybridization," in *Methods in Microbiology*. Elsevier), 325-347.

Santos, I.C., Smuts, J., Choi, W.-S., Kim, Y., Kim, S.B., and Schug, K.A. (2018). Analysis of bacterial FAMEs using gas chromatography–vacuum ultraviolet spectroscopy for the identification and discrimination of bacteria. *Talanta* 182**,** 536-543.

Sasser, M. (1990). Identification of bacteria by gas chromatography of cellular fatty acids.

Sasser, M. (2006). Bacterial identification by gas chromatographic analysis of fatty acids methyl esters (GC-FAME). *Newark, NY: Microbial ID*.

Satiaputra, J., Eijkelkamp, B.A., Mcdevitt, C.A., Shearwin, K.E., Booker, G.W., and Polyak, S.W. (2018). Biotin-mediated growth and gene expression in

Staphylococcus aureus is highly responsive to environmental biotin. *Applied Microbiology and Biotechnology*, 1-11.

Satiaputra, J., Shearwin, K.E., Booker, G.W., and Polyak, S.W. (2016). Mechanisms of biotin-regulated gene expression in microbes. *Synthetic and Systems Biotechnology* 1, 17-24.

Satoh, A., Konishi, S., Tamura, H., Stickland, H.G., Whitney, H.M., Smith, A.G., Matsumura, H., and Inoue, T. (2010). Substrate-induced closing of the active site revealed by the crystal structure of pantothenate synthetase from *Staphylococcus aureus*. *Biochemistry* 49, 6400-6410.

Scheffers, D.-J., and Pinho, M.G. (2005). Bacterial cell wall synthesis: new insights from localization studies. *Microbiology and Molecular Biology Reviews* 69, 585-607.

Schleifer, K.-H. (2015). "Macrococcus," in *Bergey's Manual of Systematics of Archaea and Bacteria*. John Wiley & Sons, Ltd).

Schleifer, K. (1973). Chemical composition of staphylococcal cell walls. *Contributions to microbiology and immunology* 1, 13.

Schleifer, K., Kilpper-Balz, R., Fischer, U., Faller, A., and Endl, J. (1982a). *Identification of "Micrococcus candidus" ATCC 14852 as a Strain of Staphylococcus epidermidis and of "Micrococcus caseolyticus" ATCC 13548 and Micrococcus varians ATCC 29750 as Members of a New Species, Staphylococcus caseolyticus.*

Schleifer, K.H., and Kandler, O. (1972a). Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriological reviews* 36, 407.

Schleifer, K.H., and Kandler, O. (1972b). Peptidoglycan types of bacterial cell walls and their taxonomic implications. *Bacteriological Reviews* 36, 407-477.

Schleifer, K.H., Killper-Balz, R., Fischer, U., Faller, A., and Endl, J. (1982b). Identification of "*Micrococcus candidus*" ATCC 14852 as a strain of *Staphylococcus epidermidis* and of "*Micrococcus caseolyticus*" ATCC 13548 and *Micrococcus varians* ATCC 29750 as Members of a New Species, *Staphylococcus caseolyticus*. *International Journal of Systematic and Evolutionary Microbiology* 32, 15-20.

Schleifer, K.H., and Kloos, W.E. (1975). A simple test system for the separation of staphylococci from micrococci. *Journal of Clinical Microbiology* 1, 337-338.

Schwartz, D.C., and Cantor, C.R. (1984). Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell* 37**,** 67-75.

Schwarz, S., Werckenthin, C., and Kehrenberg, C. (2000). Identification of a Plasmid-Borne Chloramphenicol-Florfenicol Resistance Gene in *Staphylococcus sciuri*. *Antimicrobial agents and chemotherapy* 44**,** 2530-2533.

Schwendener, S., Cotting, K., and Perreten, V. (2017). Novel methicillin resistance gene *mecD* in clinical *Macrococcus caseolyticus* strains from bovine and canine sources. *Scientific Reports* 7**,** 43797.

Silvestri, L.G., and Hill, L.R. (1965). Agreement Between Deoxyribonucleic Acid Base Composition and Taxometric Classification of Gram-Positive Cocci. *Journal of Bacteriology* 90**,** 136-140.

Skerker, J.M., Prasol, M.S., Perchuk, B.S., Biondi, E.G., and Laub, M.T. (2005). Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis. *PLoS biology* 3**,** e334.

Smith, L.K., and Mankin, A.S. (2008). Transcriptional and translational control of the mlr operon, which confers resistance to seven classes of protein synthesis inhibitors. *Antimicrobial agents and chemotherapy* 52**,** 1703-1712.

Song, M.D., Wachi, M., Ishino, F., and Matsuhashi, M. (1987). Evolution of an inducible penicillin-target protein in methicillin-resistant *Staphylococcus aureus* by gene fusion. *FEBS letters* 221**,** 167-171.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. *Journal of Molecular Biology* 98**,** 503-517.

Spry, C., Kirk, K., and Saliba, K.J. (2007). Coenzyme A biosynthesis: an antimicrobial drug target. *FEMS microbiology reviews* 32**,** 56-106.

Stackebrandt, E., and Woese, C.R. (1979). A phylogenetic dissection of the family *Micrococcaceae*. *Current Microbiology* 2**,** 317-322.

Streit, W., and Entcheva, P. (2003). Biotin in microbes, the genes involved in its biosynthesis, its biochemical role and perspectives for biotechnological production. *Applied Microbiology and Biotechnology* 61**,** 21-31.

Tenover, F.C., Arbeit, R.D., Goering, R.V., Mickelsen, P.A., Murray, B.E., Persing, D.H., and Swaminathan, B. (1995). Interpreting chromosomal DNA

restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. *Journal of Clinical Microbiology* 33**,** 2233.

Thöny-Meyer, L. (1997). Biogenesis of respiratory cytochromes in bacteria. *Microbiology and Molecular Biology Reviews* 61**,** 337-376.

Toh, S.M., Xiong, L., Arias, C.A., Villegas, M.V., Lolans, K., Quinn, J., and Mankin, A.S. (2007). Acquisition of a natural resistance gene renders a clinical strain of methicillin-resistant *Staphylococcus aureus* resistant to the synthetic antibiotic linezolid. *Molecular microbiology* 64**,** 1506-1514.

Tsubakishita, S., Kuwahara-Arai, K., Baba, T., and Hiramatsu, K. (2010a). Staphylococcal Cassette Chromosome mec-Like Element in *Macrococcus caseolyticus*. *Antimicrobial Agents and Chemotherapy* 54**,** 1469-1475.

Tsubakishita, S., Kuwahara-Arai, K., Sasaki, T., and Hiramatsu, K. (2010b). Origin and molecular evolution of the determinant of methicillin resistance in staphylococci. *Antimicrobial agents and chemotherapy* 54**,** 4352-4359.

Vandamme, P., Pot, B., Gillis, M., De Vos, P., Kersters, K., and Swings, J. (1996). Polyphasic taxonomy, a consensus approach to bacterial systematics. *Microbiology Reviews* 60**,** 407-438.

Vos, P., Garrity, G., Jones, D., Krieg, N.R., Ludwig, W., Rainey, F.A., Schleifer, K.-H., and Whitman, W.B. (2011). *Bergey's manual of systematic bacteriology: Volume 3: The Firmicutes.* Springer Science & Business Media.

Wang, Y., He, T., Schwarz, S., Zhou, D., Shen, Z., Wu, C., Wang, Y., Ma, L., Zhang, Q., and Shen, J. (2012a). Detection of the staphylococcal multiresistance gene *cfr* in *Escherichia coli* of domestic-animal origin. *Journal of Antimicrobial Chemotherapy* 67**,** 1094-1098.

Wang, Y., Wang, Y., Schwarz, S., Shen, Z., Zhou, N., Lin, J., Wu, C., and Shen, J. (2012b). Detection of the staphylococcal multiresistance gene *cfr* in *Macrococcus caseolyticus* and *Jeotgalicoccus pinnipedialis*. *Journal of antimicrobial chemotherapy* 67**,** 1824-1827.

Wang, Y., Wang, Y., Wu, C.-M., Schwarz, S., Shen, Z., Zhang, W., Zhang, Q., and Shen, J.-Z. (2011). Detection of the staphylococcal multiresistance gene *cfr* in *Proteus vulgaris* of food animal origin. *Journal of antimicrobial chemotherapy* 66**,** 2521-2526.

Wayne, L., Brenner, D., Colwell, R., Grimont, P., Kandler, O., Krichevsky, M., Moore, L., Moore, W., Murray, R., and Stackebrandt, E. (1987). Report of

the ad hoc committee on reconciliation of approaches to bacterial systematics. *International Journal of Systematic and Evolutionary Microbiology* 37**,** 463-464.

Wu, S., Piscitelli, C., De Lencastre, H., and Tomasz, A. (1996). Tracking the evolutionary origin of the methicillin resistance gene: cloning and sequencing of a homologue of mecA from a methicillin susceptible strain of *Staphylococcus sciuri*. *Microbial Drug Resistance* 2**,** 435-441.

Wu, Y., Cui, C., Sun, W., Yang, B., and Zhao, M. (2009). Effects of *Staphylococcus condimenti* and *Micrococcus caseolyticus* on the volatile compounds of Cantonese sausage. *Journal of Food Process Engineering* 32**,** 844-854.

Yoovidhya, T., Combes, D., and Monsan, P. (1986). Kinetic and thermal stability studies of rulactine, a proteolytic enzyme from *Micrococcus caseolyticus*. *Biotechnology letters* 8**,** 333-338.

Zhang, H., Hackbarth, C., Chansky, K., and Chambers, H. (2001). A proteolytic transmembrane signaling pathway and resistance to β-lactams in staphylococci. *Science* 291**,** 1962-1965.

Zopf, W. (1885). *Die spaltpilze.* Eduard Trewendt.

Zwir, I., Harari, O., and Groisman, E.A. (2007). "[18] - Gene Promoter Scan Methodology for Identifying and Classifying Coregulated Promoters," in *Methods in Enzymology*, eds. M.I. Simon, B.R. Crane & A. Crane. Academic Press), 361-385.

Chapter 2

A Rapid PCR-based Method to Discriminate *Macrococcus caseolyticus* and *Macrococcus canis* from closely related *Staphylococcus* species based on the *cta*C Gene Sequence

This entire chapter has been published in the Journal of Microbiological Methods (2018).

152, 36-38.

DOI: 10.1016/j.mimet.2018.07.008

2.1 Abstract

Our method exploits the amplification of the cytochrome c oxidase subunit II (*ctaC*) gene for the screening of *Macrococcus caseolyticus* and *Macrococcus canis* in complex microbial communities and discriminating these species from strains of their sister genus *Staphylococcus*. Thirteen novel strains of these species were isolated using this approach.

Keywords: cytochrome c oxidase subunit II; screening: *Macrococcus caseolyticus*; *Macrococcus canis*; discriminating

2.2 Introduction

The *Macrococcus* genus is classified taxonomically in the family of *Staphylococcaceae* (Lory, 2014) and is currently comprised of eleven species - *Macrococcus bovicus, Macrococcus carouselicus, Macrococcus equipercicus, Macrococcus brunensis, Macrococcus hajekii, Macrococcus lamae, Macrococcus goetzii, Macrococcus epidermidis, Macrococcus bohemicus, Macrococcus caseolyticus* and *Macrococcus canis* (Brawand et al., 2017; Mašlaňová et al., 2018), with the latter two species being more closely related to species of the *Staphylococcus* genus than the other *Macrococcus* species (Mazhar et al., 2018). *M. caseolyticus* has a documented presence in dairy and meat products, and has been associated with flavour development in certain fermented foods (Wu et al., 2009; Schwendener et al., 2017). 16S rRNA sequence analysis indicates *M. caseolyticus* is most closely related to the recently defined *M. canis* species, with sequence similarities between the type strains of both species of 99.7%. The *M. canis* species, isolated from canine sources, demonstrates haemolytic activities and has been associated with canine infections (Brawand et al., 2017). Whole genome analysis has identified the presence of methicillin and other multidrug resistance determinants in both species (Baba et al., 2009; Wang et al., 2012; Cotting et al., 2017). This evidence of occurrence of *M. caseolyticus* in food and the possible transfer of these resistance plasmids through the food chain to commensal and pathogenic bacteria is a cause for concern, given the European Food Safety Authority framework for the safety evaluation of micro-organisms in the food chain (Bourdichon et al., 2012; Brodmann et al., 2017). To further our knowledge on these species and the potential risks associated with their presence in foods, we embarked on a phenotype-based screening programme to isolate novel *M. caseolyticus* and *M. canis* strains from complex environmental sources. However, this method was time consuming and complicated by overwhelming numbers of competitive *Staphylococcus* species.

In this study we address the limitations of phenotype-based screening by using a PCR-based method, utilising specific primers targeting the partial amplification of a region within the cytochrome c oxidase (COX) subunit II (*ctaC*) gene which is present in *M. caseolyticus* and *M. canis* but absent in most *Staphylococcus* species, including *S. aureus*, with the exception of species from the *Staphylococcus sciuri* group (Baba et al., 2009).

2.3 Materials and Methods

2.3.1 Bacterial Strains and Culture Conditions

All *Macrococcus* and *Staphylococcus* strains employed in this study were cultivated at 37°C for 24 h in Tryptic Soy Broth (TSB; Becton, Dickinson and Company, Berkshire, England). Reference strains used are listed in Table 1. All of the strains used were either purchased from the American Type Culture Collection (ATCC, Middlesex, Uk), Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH (DSMZ, Braunschweig, Germany) or obtained from the Dairy production centre (DPC) culture collection at the Teagasc Food Research Centre (Moorepark, Fermoy, Cork, Ireland). Information relating to all reference and novel strains isolated during this study are summarised in Table 1.

2.3.2 *M. caseolyticus* Isolation from Dairy and Non-Dairy Sources

Raw milk samples and swabs from bovine tongues were used in this study as sources of novel *M. caseolyticus* and *M. canis* strains. The samples were sourced from Curtins Research Farm (Teagasc, Moorepark, Fermoy, Cork, Ireland). Raw milk samples were serially diluted in maximum recovery diluent (MRD) (Oxoid Ltd. Hampshire, England), plated on to Mannitol Salt Agar (MSA; Sigma-Aldrich, Wicklow, Ireland) and incubated for 24 h at 37°C. Swabs from bovine tongue samples were streaked directly on to MSA agar plates and incubated for 24 h at 37°C. Resulting mannitol non-fermenters (pink colonies) from MSA were replica streaked on to MSA and Tryptic Soy Agar (TSA; Becton, Dickinson and Company, Berkshire, England) plates, which were incubated for 24 h at 37°C.

2.3.3 Design of *ctaC*-Specific Primers

The primer sequences for partial amplification of the *ctaC* gene were designed using default parameters of NCBI/Primer-BLAST tool (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) and synthesised by Sigma Aldrich (Sigma Aldrich, Wicklow, Ireland).

Table 1: *Staphylococcus* and *Macrococcus* reference strains and novel *Macrococcus caseolyticus* strains used in this study.

Bacterial Species (16 S rRNA)	**Strain designation**	**Source**	**Reference**
Macrococcus caseolyticus	[a]ATCC 13548^T	Cow's Milk	(Kloos et al., 1998)
Macrococcus caseolyticus	ATCC 13518	Cow's Milk	(Kloos et al., 1998)
Macrococcus caseolyticus	ATCC 51835	Pilot whale	(Kloos et al., 1998)
Macrococcus bovicus	[b]DSMZ15607^T	Holstein cow	(Kloos et al., 1998)
Macrococcus equipercicus	DSMZ 15609^T	Irish thoroughbred horse	(Kloos et al., 1998)
Staphylococcus vitulinus	DSMZ 15615^T	Ground lamb	(Webster et al., 1994)
Staphylococcus lentus	DSMZ 20046	Unknown	(Schleifer et al., 1983)
Staphylococcus sciuri	DSMZ 6671	Unknown	(Kloos et al., 1976)
Staphylococcus aureus	[c]DPC6868	Raw milk	ND
Staphylococcus aureus	DPC6867	Raw milk	ND
Staphylococcus epidermidis	DPC5804	Bovine mastitis isolate	ND
Staphylococcus carnosus	DPC3312	Unknown	ND
Macrococcus caseolyticus	DPC7158	Bovine Tongue	This study
Macrococcus caseolyticus	DPC7159	Bovine Tongue	This study
Macrococcus caseolyticus	DPC7160	Bovine Tongue	This study
Macrococcus caseolyticus	DPC7161	Bovine Tongue	This study
Macrococcus caseolyticus	DPC7162	Bovine Tongue	This study
Macrococcus caseolyticus	DPC7163	Bovine Tongue	This study
Macrococcus caseolyticus	DPC7164	Bovine Tongue	This study
Macrococcus caseolyticus	DPC7165	Bovine Tongue	This study
Macrococcus caseolyticus	DPC7166	Bovine Tongue	This study
Macrococcus caseolyticus	DPC7168	Bovine Tongue	This study
Macrococcus caseolyticus	DPC7169	Bovine Tongue	This study
Macrococcus caseolyticus	DPC7170	Raw Milk	This study
Macrococcus caseolyticus	DPC7171	Raw Milk	This study

[a] ATCC; American Type Culture Collection; [b]DSMZ; Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH; [c]DPC; Dairy Production Centre. ND;not determined

2.3.4 *ctaC*-Specific PCR Analysis of Reference Strains

Genomic DNA for all of the reference strains in this study was extracted using the Ultra Clean Microbial DNA Isolation Kit (Mo-Bio Laboratories, Cambridge, United Kingdom) according to the manufacturer's instructions. PCR reactions were performed in 30 µl reaction volumes using 0.3 µM primers, MyTaq™ PCR buffer at a final concentration of 1X, 2.5 units of Taq polymerase (MyTaq™ BIO-21105; Bioline, Dublin, Ireland) and 3 µl (≈50 ng/ µl) of template DNA. PCR was performed in an Eppendorf Mastercycler®Pro (Eppendorf® Mastercycler® Pro Thermal Cyclers, VWR, Dublin, Ireland). The amplification consisted of an initial denaturation step of 94°C for 5 min, with denaturation of 94°C for 30 sec, primer annealing of 55°C for 30 sec, extension of 72°C for 30 sec for 35 cycles and a final extension of 72°C for 2 min and cooling to 4°C. PCR amplicons were purified using the ISOLATE II PCR and Gel Kit (BIO-52060; Bioline, Dublin, Ireland) according to the manufacturer's instructions. The concentration and purity of the isolated amplicons were measured using the Nanodrop-Spectrophotometer (NanoDrop 1000 Spectrophotometer; Thermo Scientific, Dublin, Ireland). Sequence analysis of purified PCR products was performed by GATC Biotech AG (Cologne, Germany). BLAST was utilised to conduct comparisons to sequences available in the NCBI BLASTn database (Altschul et al., 1997).

2.3.5 *ctaC*-Specific PCR Screening of Complex Samples

For the PCR screening, template DNA was extracted from single colonies by suspending each colony in 50 µl of sterile ultra-pure water (DNase-RNase free) (Teagasc, Moorepark, Fermoy, Cork, Ireland) in a 1.5 ml micro-tube (Sarstedt, Wexford, Leinster, Ireland) and heat treating the sample for 15 min at 95°C followed by centrifugation (Eppendorf® Centrifuge 5242 R) at 1,300 × g for 3 min. 5 µl of the resulting supernatant was used as template DNA for the PCR reaction. PCR reactions were performed in 30 µl reaction volumes using 0.3 µM primers, MyTaq™ PCR buffer at final concentration of 1X, and 2.5 units of Taq polymerase (MyTaq™ BIO-21105; Bioline, Dublin, Ireland) using the PCR parameters described previously.

2.3.6 16S rRNA Analysis

Species identification of novel isolates was determined using the method described by Alander et al, (Alander et al., 1999). The 16S rRNA amplicons (≈1500 bp) were forwarded for sequence analysis to GATC Biotech AG (Cologne, Germany) and analysed by BLAST as described previously.

2.3.7 Gel Electrophoresis and Image Analysis

Analysis of PCR products was performed by electrophoresis at 60 V cm-1 for 1 h on a 200 ml 1% (w/v) agarose gel (Medical Supply Company Ltd, Dublin, Ireland), prepared using a 0.5X dilution of TRIS-acetate-EDTA buffer (Thermo Fisher Scientific, Dublin, Ireland) and stained with 1X Gel Red (VWR, Dublin, Ireland). The size of PCR amplicons were determined by using 1 kb marker (HyperLadder™ 1 kb; Bioline, Dublin, Ireland) ranging from 200 to 10037 bp and 100 bp marker (HyperLadder™ 100 bp; Bioline, Dublin, Ireland) ranging from 100 to 1013 bp.

2.3.8 Pulsed Field Gel Electrophoresis (PFGE)

DNA was prepared according to the procedure described by Simpson *et al*, with the following modifications (Simpson et al., 2002). Strains of interest were grown in 5 ml Tryptic Soy broth (TSB; Becton, Dickinson and Company, Berkshire, England) for 24 h at 37°C. 1 ml of stationary-phase cultures were spun for 3 min at 15,871 × g. Restriction digestion was performed using the *SmaI* enzyme and Cutsmart restriction buffer (B7204S; New England Biolabs, Hitchin, England). Plugs were loaded into the wells of a 200 ml 1% (w/v) pulsed-field grade agarose (Bio-Rad Laboratories, Dublin, Ireland) prepared in 0.5X dilution of TRIS-borate-EDTA buffer (Sigma Aldrich, Wicklow, Ireland) and loaded in to the PFGE chamber (CHEF-DR II system; Bio-Rad Laboratories, Dublin, Ireland). DNA fragments were resolved at 6 V/cm for 18 h with 0.5X TRIS-borate-EDTA running buffer (Sigma Aldrich, Wicklow, Ireland), maintained at 14°C (Simpson et al., 2002) and pulse ramped from 1 to 20 sec. Gels were stained in distilled water containing 0.5 µg /ml of ethidium bromide (Sigma Aldrich, Wicklow, Ireland) for 1 h and destained in distilled water for 1 h. For PFGE gel analysis, a low-range PFGE marker (N0350S, Brennan & Company, Dublin, Ireland) ranging from 2 to 194 kb was used. Gels were visualized by transillumination UV using the AlphaImager™ 3400 detection system (Alpha Innotech, BioSurplus, San Diego, US) along with AlphaEaseFC™ software (Alpha Innotech, BioSurplus, San Diego, US).

2.3.9 Phylogenetic Analysis of the Targeted *ctaC* Sequence

To construct a Phylogenetic tree, sequence alignment was carried out using MUSCLE (Edgar, 2004). MEGA version 7.0 software was used to evaluate the phylogenetic relationship with reference *M. caseolyticus* and *M. canis* strains and strains available in the NCBI-Genbank database, which includes *M. caseolyticus* JCS5402 (GenBank accession number; NC_011999.1) and IMD8019 (CP021058.1), *M. canis* KM45013^{T} (CP021059.1) and two strains from the *S. sciuri* used as outgroup; FDAARGOS_2859 (CP022046.1) and SNUDS-18

(CP020377.1). Neighbour-joining (NJ) method was used along with Kimura 2 parameter model to compute the distance formula and 100 bootstrap replicates.

2.4 Results and Discussion

2.4.1 Design of *ctaC* Primers and *in silico* Specificity

At the time of development of the *ctaC* PCR method, a solitary complete genome of *M. caseolyticus* strain JCSC5402 (GenBank accession number; NC_011999.1) was deposited in the NCBI database. The COX operon (*cta*ABCDEFG) of *M. caseolyticus* JCSC5402 is illustrated in Fig 1. Sequence similarity analysis of genes for the first three core subunits (*ctaCDE*) of COX indicated the nucleotide sequence coding of *ctaC* as being the least conserved. Therefore, the *ctaC* gene was selected as the target sequence for primer design. The primer set of *ctaC* F-5'-GAACTGTCTGCTTTACGTCC-3' and *ctaC* R-5'-GGGTACTCAAATTCCCACCAG-3' was selected, based on the predicted product size (370 bp), *in silico* specificity of the primer set, and the inability of the primer sequences to form homodimers as well as hairpins. The location of the binding sites of these primers within the *ctaC* gene from *M. caseolyticus* JCSC5402 is indicated in Fig. 1, with a predicted product size of 370 bp.

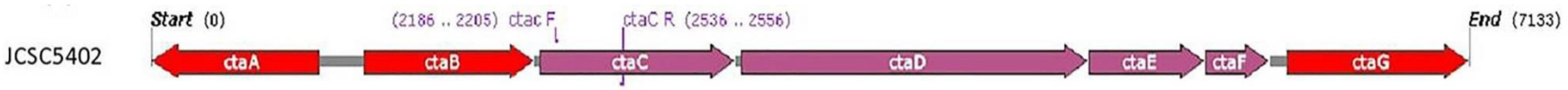


Fig.1 Schematic presentation of cytochrome c oxidase (COX) gene cluster in *M. caseolyticus* JCSC5402 strain (NC_011999.1). The arrows indicate the direction of transcription of the open reading frame. All annotated regions are colour coded: *ctaA* (biosynthesis of heme A), *ctaB* (biosynthesis of heme O) and *ctaG* (caa3 assembly factor) are shown in red and *ctaD, ctaC ctaE* and *ctaF*, the four subunits of COX are shown in purple. Location of primer binding site is indicated within the region of *ctaC* gene. A scale size in base pairs (bp) is displayed from left hand to right hand corner.

The specificity of the *ctaC* primer set was evaluated utilising each of the *Macrococcus* and *Staphylococcus* reference strains listed in Fig. 2(A) as template DNA. For the *M. caseolyticus* and *M. canis* reference strains tested, a single PCR amplicon of the expected size of 370 bp was generated (Fig. 2). BLASTn analysis of these products demonstrated 99% sequence similarity with the target *ctaC* gene from *M. caseolyticus* JCS5402, whereas the *ctaC* from DSMZ 101690^{T} shared 100% identity with *M. canis* KM45013^{T}. For *M. bovicus* and *M. equipercicus*, no amplification was observed. Similarly, no amplification was observed *S. aureus*, *S. sciuri* and *S. carnosus;* Fig. 2 (A), with the exception of *S. epidermidis* DPC5804, which produced a faint band of approx. 420 bp (data not shown). Subsequent sequencing of this product showed that it was unrelated to the target gene. Thus, the *ctaC* primer set appeared to be (a) specific for the *ctaC* gene of strains of *M. caseolyticus* and *M. canis* and (b) capable of discriminating strains of interest from other *Macrococcus* and *Staphylococcus* species when tested in pure culture.

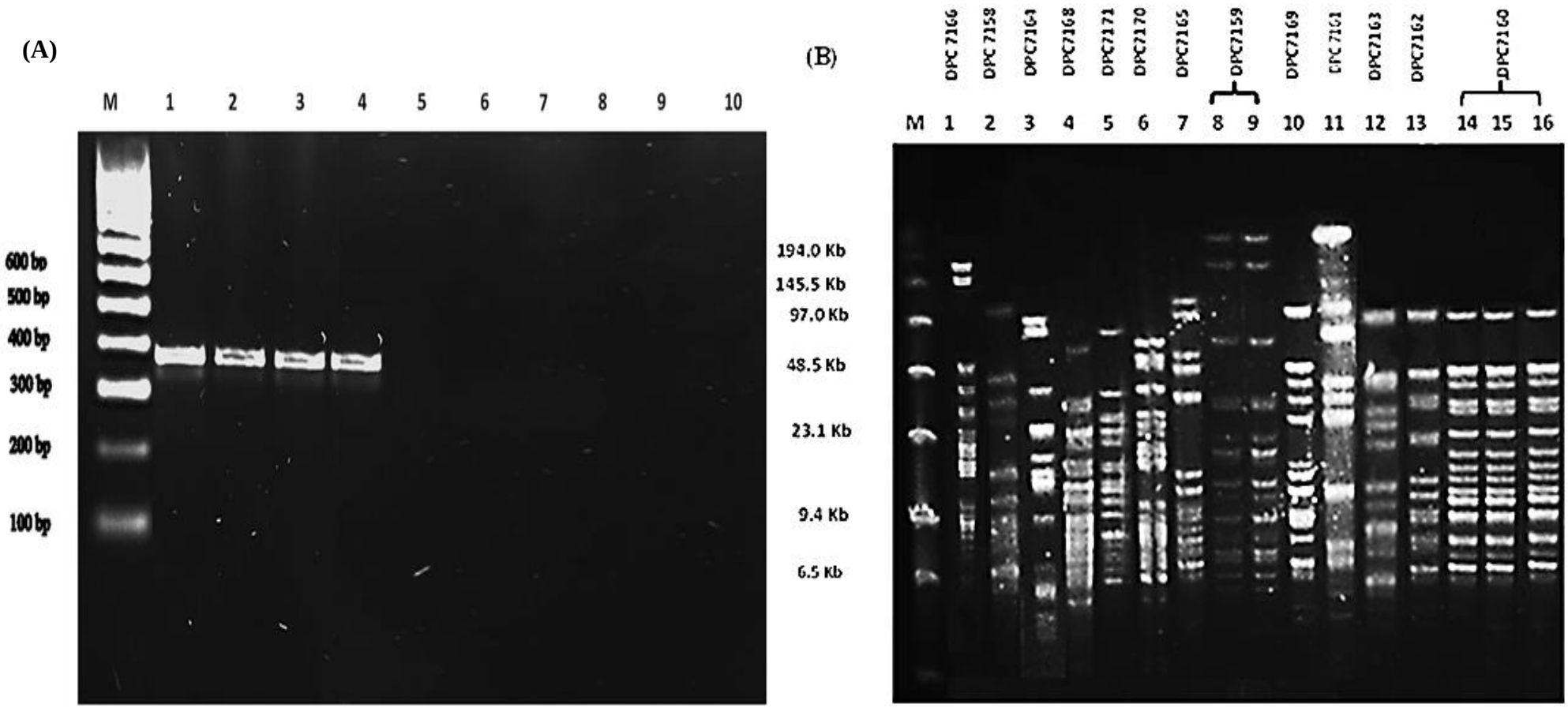


Fig.2 (A) Specificity of *ctaC* primers using pure cultures of reference strains of *Macrococcus* and *Staphylococcus.* M: 100-bp DNA ladder, Lane 1: *Macrococcus caseolyticus* ATCC 13518, Lane 2: *Macrococcus caseolyticus* ATCC 13548^{T}, Lane 3: *Macrococcus caseolyticus* ATCC 51835, Lane 4 *Macrococcus canis* DSMZ 101690^{T}, Lane 5 *Macrococcus bovicus* DSMZ15607, Lane 6: *Macrococcus equipercicus* DSMZ 15609, Lane 7: *Staphylococcus aureus* DPC 6868, Lane 8: *Staphylococcus aureus* DPC 6867, Lane 9: *Staphylococcus sciuri* DSM 6671, Lane 10: *Staphylococcus carnosus* DPC3312. ATCC; American Type Culture Collection; DSMZ; Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH; DPC; Teagasc Culture Collection.

(B) Pulse-field gel electrophoresis (PFGE) of SmaI-digested genomic DNA of 16 *ctaC* positive isolates, isolated from bovine milk and tongue. Lane M; Low Range PFGE Marker (NO350S, Brenan & Company, Dublin, Ireland) Lanes 1-16 are the representative 13 pulsotype obtained (note lane 8-9; have similar pulsotype and therefore referred to as strain DPC7159 and lane 14-16 referred to as DPC7160).

2.4.4 Nucleotide Sequence Divergence of *ctaC* Fragments

Phylogenetic analysis was conducted of the amplified region of the *ctaC* gene from each of the 13 strains along with reference strains and other strains available on the NCBI-Genbank database (Fig. 3). Results of the analysis indicated a number of the isolated strains (DPC7168, DPC7169, DPC7165, DPC7163, DPC7162, DPC7160) form a sister group with *M. canis* KM45013^{T} (CP021059.1), whereas only two strains (DPC7171 and DPC7170) clustered with reference *M. caseolyticus* strains. Interestingly, a number of the isolated strains (DPC7159, DPC7164, DPC7166, DPC7161) form their own clade (Fig. 3), suggesting that they may not be either *M. canis* or *M. caseolyticus*, as defined by the 16S analysis, but another as-yet undefined *Macrococcus* species.

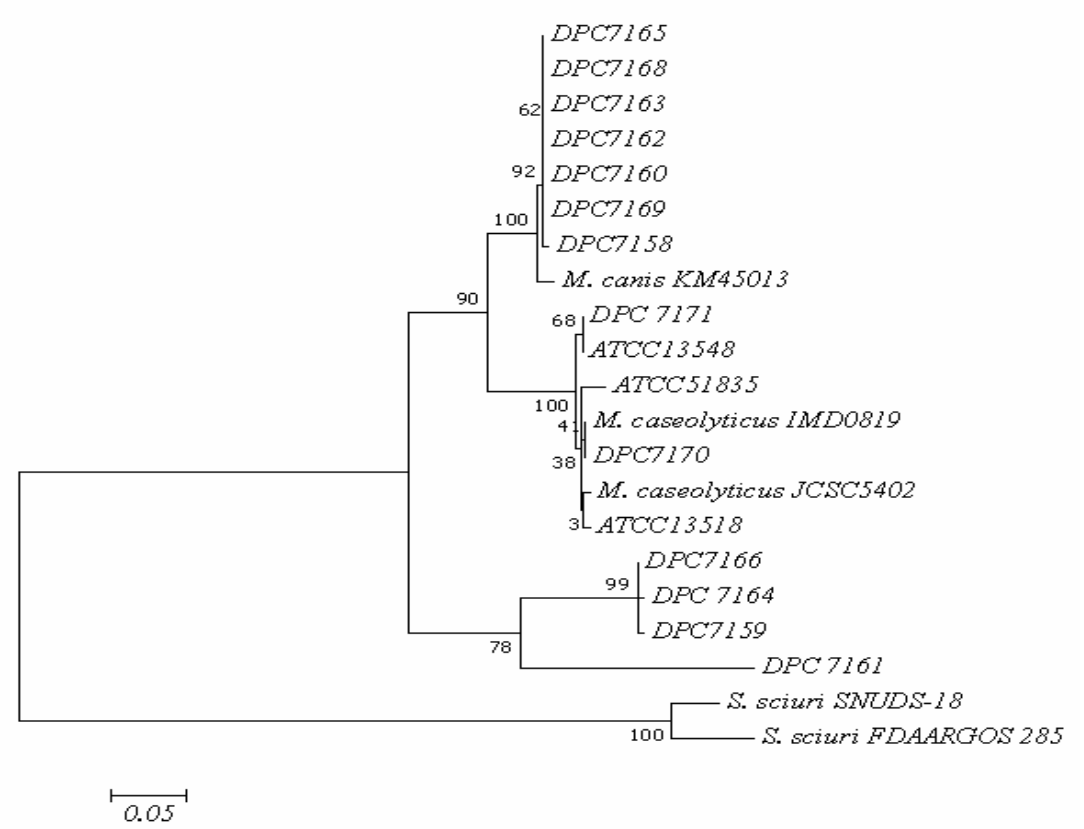


Fig.3 Phylogenetic unrooted tree representing relationship among the partial gene sequence of the *ctaC* sequence from different strains isolated in this study along with reference *M. caseolyticus* and *M. canis* strains and strains available in the NCBI-Genbank database, which includes *M. caseolyticus* JCS5402 (GenBank accession number; NC_011999.1) and IMD8019 (CP021058.1), *M. canis* KM45013^{T} (CP021059.1) and two strains from the *S. sciuri*; FDAARGOS_2859 (CP022046.1) and SNUDS-18 (CP020377.1). The tree was generated from the output of sequence alignment by MUSCLE [multiple alignment (Gap opening: -400, Gap extension:0, clustering method: UPGMB)] in MEGA version 7.0 software, phylogenetic relationship was evaluated with Neighbour-joining (NJ) method along with Kimura 2 parameter model to compute the distance formula. Bootstrap values are shown at each node as a percentage of 100 replicates. The scale bar (NJ distance) represents 5% difference in nucleotide sequences.

2.5 Conclusion

This study demonstrates that a PCR-based method designed for partial amplification of the *ctaC* gene is capable of discriminating *M. caseolyticus* and *M. canis* species from a background of *Staphylococcus*. Screening diverse sources with the developed method has resulted in the isolation of 13 distinct strains of the targeted species. Further investigation will reveal if strains of a novel *Macrococcus* species have also been isolated with this method.

2.6 Acknowledgements

This research is funded by Teagasc (ref. 6697) and the Teagasc Walsh Fellowship programme (ref. 2015055).

2.7 References

Alander, M., Satokari, R., Korpela, R., Saxelin, M., Vilpponen-Salmela, T., Mattila-Sandholm, T., and Von Wright, A. (1999). Persistence of Colonization of Human Colonic Mucosa by a Probiotic Strain, Lactobacillus rhamnosus GG, after Oral Consumption. *Applied and Environmental Microbiology* 65**,** 351-354.

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Resources* 25**,** 3389-3402.

Baba, T., Kuwahara-Arai, K., Uchiyama, I., Takeuchi, F., Ito, T., and Hiramatsu, K. (2009). Complete genome sequence of *Macrococcus caseolyticus* strain JSCS5402, reflecting the ancestral genome of the human-pathogenic staphylococci. *Journal of Bacteriology* 191**,** 1180-1190.

Bourdichon, F., Casaregola, S., Farrokh, C., Frisvad, J.C., Gerds, M.L., Hammes, W.P., Harnett, J., Huys, G., Laulund, S., and Ouwehand, A. (2012). Food fermentations: microorganisms with technological beneficial use. *International journal of food microbiology* 154**,** 87-97.

Brawand, S.G., Cotting, K., Gómez-Sanz, E., Collaud, A., Thomann, A., Brodard, I., Rodriguez-Campos, S., Strauss, C., and Perreten, V. (2017). *Macrococcus canis* sp. nov., a skin bacterium associated with infections in dogs. *International Journal of Systematic and Evolutionary Microbiology* 67**,** 621-626.

Brodmann, T., Endo, A., Gueimonde, M., Vinderola, G., Kneifel, W., De Vos, W.M., Salminen, S., and Gómez-Gallego, C. (2017). Safety of Novel Microbes for Human Consumption: Practical Examples of Assessment in the European Union. *Frontiers in Microbiology* 8**,** 1725.

Cotting, K., Strauss, C., Rodriguez-Campos, S., Rostaher, A., Fischer, N.M., Roosje, P.J., Favrot, C., and Perreten, V. (2017). *Macrococcus canis* and *M. caseolyticus* in dogs: occurrence, genetic diversity and antibiotic resistance. *Veterinary Dermatology*.

Edgar, R.C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic acids research* 32**,** 1792-1797.

Lory, S. (2014). "The Family *Staphylococcaceae*," in *The Prokaryotes: Firmicutes and Tenericutes,* eds. E. Rosenberg, E.F. Delong, S. Lory, E. Stackebrandt & F. Thompson. (Berlin, Heidelberg: Springer Berlin Heidelberg), 363-366.

Mašlaňová, I., Wertheimer, Z., Sedláček, I., Švec, P., Indráková, A., Kovařovic, V., Schumann, P., Spröer, C., Králová, S., Šedo, O., Krištofová, L., Vrbovská, V., Füzik, T., Petráš, P., Zdráhal, Z., Ružičková, V., Doškař, J., and Pantuček, R. (2018). Description and comparative genomics of *Macrococcus caseolyticus* subsp. *hominis* subsp. nov., *Macrococcus goetzii* sp. nov., *Macrococcus epidermidis* sp. nov., and *Macrococcus bohemicus* sp. nov., novel macrococci from human clinical material with virulence potential and suspected uptake of foreign DNA by natural transformation. *Frontiers in Microbiology*, 9.

Mazhar, S., Hill, C., and Mcauliffe, O. (2018). A rapid PCR-based method to discriminate *Macrococcus caseolyticus* and *Macrococcus canis* from closely-related *Staphylococcus* species based on the *ctaC* gene sequence. *Journal of Microbiological Methods* 152**,** 36-38.

Schwendener, S., Cotting, K., and Perreten, V. (2017). Novel methicillin resistance gene *mecD* in clinical *Macrococcus caseolyticus* strains from bovine and canine sources. *Scientific Reports* 7**,** 43797.

Simpson, P., Stanton, C., Fitzgerald, G., and Ross, R. (2002). Genomic diversity within the genus *Pediococcus* as revealed by randomly amplified polymorphic DNA PCR and pulsed-field gel electrophoresis. *Applied and Environmental Microbiology* 68**,** 765-771.

Wang, Y., Wang, Y., Schwarz, S., Shen, Z., Zhou, N., Lin, J., Wu, C., and Shen, J. (2012). Detection of the staphylococcal multiresistance gene cfr in *Macrococcus caseolyticus* and *Jeotgalicoccus pinnipedialis*. *Journal of Antimicrobial Chemotherapy* 67**,** 1824-1827.

Wu, Y., Cui, C., Sun, W., Yang, B., and Zhao, M. (2009). Effects of *Staphylococcus condimenti* and *Micrococcus caseolyticus* on the volatile compounds of Cantonese sausage. *Journal of Food Process Engineering* 32**,** 844-854.





Mazhar, S. 2020. A multi-omics perspective on the biology and evolution of the genus Macrococcus. PhD Thesis, University College Cork.

Please note that Chapters 3&4 (pp. 86-143) are unavailable due to a restriction requested by the author.

CORA Cork Open Research Archive http://cora.ucc.ie

Chapter 5

A Systems-wide Analysis of Proteolytic and Lipolytic Pathways Uncovers the Flavour-forming Potential of The Gram-Positive Bacterium *Macrococcus caseolyticus* subsp. *caseolyticus*

This entire chapter has been published in the Journal of Frontiers in Food Microbiology (2020).
11 (1553).
DOI: 10.3389/fmicb.2020.01533

5.1 Abstract

Macrococcus caseolyticus subsp. *caseolyticus* is a Gram-positive, commensal organism documented to be present as a component of the secondary microflora in fermented foods such as Ragusano and Fontina cheeses and Cantonese sausage. In these products, the organism appears to play a role in ripening and the development of the final organoleptic qualities. However, the role of this organism in flavor generation is not well understood. Therefore, the objective of this study was to investigate the role of *M. caseolyticus* subsp. *caseolyticus* in flavour compound formation through an examination of enzymatic, metabolomic and genomic data. A bank of *M. caseolyticus* subsp. *caseolyticus* strains derived from a variety of niches were examined. Enzyme activities analysed comprised those of the proteolytic and lipolytic cascades including cell-envelope proteinase (CEP), peptidases, esterases, lipases, aminotransferases and glutamate dehydrogenase. Strain– to strain variation was observed, often associated with niche. All strains, except those isolated from non-dairy sources, demonstrated high CEP activity. Such high CEP activity associated with dairy strains implies the importance of this characteristic in the adaptation of these strains to a dairy-specific niche. However, limited downstream peptidolytic activity, in addition to a limited ability to generate free amino acids was observed across all strains, indicating weak ability of this organism to generate amino-acid derived flavor compounds. Interestingly, the strains with high CEP activity also demonstrated high esterase activity and gas chromatography-mass spectrometry analysis of the volatile compounds produced when these strains were grown in lactose-free milk demonstrated differences in the range and types of volatiles produced. In contrast to this metabolic versatility, comparative genome analysis revealed the distribution of components of the proteolytic and lipolytic system in these strains to be conserved. Overall, this study demonstrates the potential of *M. caseolyticus* subsp. *caseolyticus* to generate diverse volatile flavor compounds. Additionally, the identification of the highly active strain -specific cell wall bound caseolytic proteases deriving extensive casein hydrolysis, serves as a promising avenue which can be potentially harnessed in the future to produce greater and more diverse flavour compounds.

Keywords: *Macrococcus caseolyticus* subsp. *caseolyticus*, fermented foods, flavour development, whole genome sequence, enzymatic assays, metabolomics.

5.2 Introduction

Members of the *Macrococcus* genus are disseminated in nature as animal commensals and are probable ancestors of staphylococcal species (Hiramatsu et al., 2014). The genus is currently composed of eleven species- *M. bovicus, M. carouselicus, M. equipercicus, M. brunensis, M. hajekii, M. lamae, M. canis, M. epidermidis, M. goetzii, M. bohemicus* and *M. caseolyticus*, which is further divided into two subspecies: *M. caseolyticus* subsp. *hominis* and *M. caseolyticus* subsp. *caseolyticus* (Mašlaňová et al., 2018). In the early 1900's, strains of *Macrococcus caseolyticus*, at the time known as *Micrococcus caseolyticus,* were isolated from raw milk samples. Early screening of these strains indicated their ability to rapidly and completely peptonize the milk, therefore the name *caseolyticus* (casein-dissolving) was suggested (Evans, 1916). Historically, studies have associated the occurrence of the *M. caseolyticus* species in food systems, as part of the secondary flora of cheese and its role in ripening and flavour development has been documented (Bhowmik and Marth, 1988; Bhowmik and Marth, 1990). The proteolytic activity of this organism on β-casein was investigated, and shown to lead to the production of short peptides. This suggested that *M. caseolyticus* was capable of producing the substrates for flavour compound production (Moreno and Kosikowski, 1973). The technological applications of enzymes extracted from this organism and their use in the production of novel cheese products have been described in a number of studies and patents (Hargrove and Mcdonough, 1964; Desmazeaud and Hermier, 1968; Alkalaf et al., 1987). The more conventional role of *M. caseolyticus* in the flavour development of Cantonese sausage has also been investigated and the majority of the volatiles generated were identified as metabolites of free fatty acid catabolism (Wu et al., 2009).

The development of flavour in dairy products is a particularly complex process which involves three main processes: glycolysis (sugar metabolism), proteolysis (degradation of proteins) and lipolysis (degradation of lipids)(McSweeney, 2017). Whilst, the metabolisms of sugars mainly lactose, are a source of many flavor compounds, the pathways that are explored in this study are proteolysis and lipolysis. Proteolysis has been regarded as one of the most important processes in the development of flavour and the enzymatic reactions involved are well defined in the lactic acid bacteria (LAB), the group of organisms most widely associated with flavour formation in dairy products (Smit et al., 2005). The proteolytic cascade commences

with the breakdown of casein into small peptides by the action of surface bound proteinases, often referred to as cell enveloped proteinases (CEP). The peptides are then transported into the cell and further degraded by the coordinated action of peptidases with different, but often partially overlapping, specificities for amino acids (Gobbetti et al., 2007). The free amino acids (FAA) generated can directly contribute to flavour, but their further metabolism is identified as a key process in flavour formation (McSweeney and Sousa, 2000). There are several pathways that lead to flavour compound generation originating from the catabolism of FAA, initiated by the activity of various enzymes such as aminotransferases (AT), lyases, decarboxylases, deminases and dehydratases (Jensen and Ardö, 2010). However, the majority of the most important flavour compounds have been reported to originate in the transamination pathway (Jensen and Ardö, 2010). Aminotransferases catalyse the transfer of an amino group to α-keto acid, which are important precursors for the generation of various volatile flavour compounds (Jensen and Ardö, 2010). The transamination reactions are dependent on the presence of an amino group receptor, usually α-ketoglutarate, which is produced by glutamate dehydrogenase (GDH). This key enzyme is described as a rate limiting factor for transamination and therefore essential for the production of precursors which lead to the generation of flavour compounds by members of the LAB (Tanous et al., 2002; Kieronczyk et al., 2004). The lipolytic pathway is also an important process in flavour development and involves the hydrolysis of lipids present in milk to free fatty acids (FFA) and glycerols, mono-or diglycerides by the action of esterases or lipases. The liberation of FFA or short- and intermediate chain fatty acids contribute directly to flavour or serve as precursors for the biosynthesis of numerous flavour-contributing volatile compounds (McSweeney and Sousa, 2000).

The purpose of this study was to investigate the metabolic pathways involved in flavour compound formation in *M. caseolyticus* subsp. *caseolyticus*. Anecdotal evidence suggested that certain strains of this subspecies had a positive impact on the flavour profile of some cheese types. To investigate this further, we analysed a bank of six *M. caseolyticus* subsp. *caseolyticus* strains (isolated from sources including whale skin, bovine milk and semi-hard cheese) through enzymatic assays, whole genome sequencing and comparative genome analysis, and metabolomics data generated by gas chromatography-mass spectrometry (GC-MS). Additionally, we

have investigated free amino acid and nitrogen utilization capabilities of these strains to further evaluate their capability in producing amino-acid derived flavour compounds. To understand their overall flavour-forming potential, genomic and phenotypic information was coupled with metabolomics data generated by GC-MS and other methodologies.

5.3 Materials and Methods

5.3.1 Bacterial Strains and Culture Conditions

A total of six *M. caseolyticus* subsp. *caseolyticus* strains were employed in this study and are presented in Table 1. All six strains were cultivated at 37°C for 24 h in Tryptic Soy Broth (TSB; Becton, Dickinson and Company, Berkshire, England). In addition, other control strains were employed for comparative analysis such as *Lactococcus lactis* subsp. *cremoris* Wg2 was used as a positive control for CEP activity and was cultivated in LM17 (Merck, Darmstadt, Germany). *Lactobacillus paracasei* DPC4206 was used as a positive control for general aminopeptidases (PepN, PepC), proline specific dipeptidase (PepX) aromatic aminotransferase (ArAT) and glutamate dehydrogenase (GDH) activities. In addition, *L. paracasei* DPC4536 was also used as a positive control for GDH activity and were cultivated in MRS media (Oxoid, Basingstoke, UK). *Yarrowia lipolytica* DPC6266 was used as a positive control for lipase activity and was propagated in broth medium containing yeast extract (YE) (Merck, Germany), all control strains were incubated aerobically at 30°C.

5.3.2 Comparative Genome Analysis and Orthologous Groups Identification

The whole genome sequences (WGS) of five out of the six, *M. caseolyticus* subsp. *caseolyticus* strains are available in the public database and the accession numbers are presented in Table 1. Details of genome sequencing and assembly were previously reported (Mašlaňová et al., 2018; Mazhar et al., 2019). Functional genome distribution (FGD) was conducted on the five *M. caseolyticus* subsp. *caseolyticus* representative strains to identify conserved and non-conserved open reading frame (ORFs) encoding enzymes involved in the proteolytic and lipolytic system using CompACTor v 0.18 at an e-value threshold of $1e^{-10}$ with FGDfinder v0.022 tool (Altermann, 2012). Additionally, core-genome and singleton analysis was carried out with OrthoVenn (a web platform for orthologous gene clustering) (Wang et al., 2015). Inferred proteins for each of the genomes by Prokka version 1.11 were used as input (Seemann, 2014).

5.3.3 Qualitative Analysis of Proteolytic and Lipolytic Activity

To evaluate the proteolytic activity of *M. caseolyticus* subsp. *caseolyticus* strains, reconstituted skim milk (RSM) agar was prepared from skim milk powder (Kerry ingredients, Cheshire, UK) at 10% (w/v) and agar (Agar; Sigma-Aldrich, Wicklow, Ireland) at 1.5% (w/v). The inoculated plates were then incubated for 24 h at 37°C. To determine lipolytic activity, tributyrin agar (Sigma-Aldrich) was prepared according to the manufacturer's instructions, with the modification prescribed by Bertuzzi, A. (2017). *Y. lipolytica* DPC6266 was used as a positive control. The plates were incubated for 48 h at 37°C. A positive result for protease and lipase activity was scored on the basis of the presence of halo of clearing around the growth of the organism. The test assay was performed in triplicate.

5.3.4 Quantitative analysis of the enzymes involved in proteolysis

5.3.4.1 Determination of CEP Activity

CEP activity was determined using a modification of the method previously described by Stefanovic et al. (2017), which is based on the EnzCheck® kit Green Fluorescence E-6638 (Molecular Probes, Eugene, OR, USA). *M. caseolyticus* subsp. *caseolyticus* strains were grown in 50 ml of TSB for 24 h at 37°C. *L. lactis* subsp. *cremoris* Wg2 used as a positive control and was propagated in 50 ml of LM17 for 24 h at 37°C. Cells were centrifuged (4000 g, 10 min, 4°C), and washed three times with 50 mmol l^{-1} Tris-HCl buffer pH 7.8 with 2 mmol $l^{-1}CaCl_2$ added. Components of the kit were prepared according to manufacturer's instructions; 100 µl of cell suspension and 100 µl of prepared BODIPY® FL casein solution were mixed in 96-well microplate (Sarstedt, Wexford, Ireland) and incubated for 24 h at 37°C. Fluorescence (Ex/ Em 505/513 nm) was measured on a Synergy 2 reader (Bio-Tek Multi Detection Plate Reader, Winooski, VT, USA), using optimal filters: 485/ 20 nm for extinction and 528/20 nm for emission. A proteinase K solution (2 µg ml^{-1}) was used as a positive control. Enzyme activities for each strain were expressed as direct fluorescence readings. All strains were evaluated in triplicate. A set of trypsin standards from 0.2 ng ml^{-1} to 70 µg ml^{-1} were prepared and their activity was measured in a similar fashion.

5.3.4.2 Preparation of Cell Free Extract

To obtain cell free extract (CFE), *M. caseolyticus* subsp. *caseolyticus* strains were incubated for 24 h in 50 ml TSB were centrifuged (4000 g, 10 min, 4°C) and washed twice with 50 mmol l^{-1} sodium phosphate buffer pH 7.5 and suspended in the same buffer to a final volume of 5 ml. Cells were disrupted by sonication (Soniprep 150; MSE LTD, London, UK) in five cycles of 15 s sonication on maximum amplitude (20 amplitude microns) and 45 s of cooling on ice. Sonicated samples were centrifuged (12 000 g, 10 min, 4°C) to remove cell debris. Cell counts (CFU per ml) were evaluated for each strain before and after sonication.

5.3.4.3 Determination of Aminopeptidases, Aminotransferases and Glutamate-Dehydrogenase Activities

Aminopeptidases activity was measured according to the method defined by Jensen and Ardo (2010), with the modifications defined by Stefanovic et al. (2017). Chromogenic substrates (L-Lysine p-nitroanilide (pNA) (Sigma-Aldrich), H-Gly-Pro-pNA, H-Arg-pNA, H-Glu-pNA, and H-Ala-Phe-Pro-pNA (Bachem, Bubendorf, Switzerland) for PepN, PepX , PepC, PepA and PepV, respectively, were prepared as 1 mmol l^{-1} solutions in 50 mmol l^{-1} sodium phosphate buffer pH 7.5. The assay mixture contained 50 µl of substrate solution and 50 µl of CFE. Absorbance was measured at 405_{nm} (Synergy HT; Bio-Tek Multi Detection Plate Reader) after 60 min of incubation at 37°C. The amount of p-nitroaniline released was determined by including a standard curve obtained for standard samples of p-nitroaniline ranging between 0 to 80 nmol. Aminopeptidase activities were expressed as nmol of p-nitroanilne released per min and mg of protein. *L. paracasei* DPC4206 was used as a control, as its PepN, PepX and PepC activity was reported by Stefanovic et al. (2017). Blanks contained 50 mmol l^{-1} sodium phosphate buffer instead of CFE. Development of yellow color in the samples, originating from p-nitroaniline, and no color development in the blank after incubation were considered as a sign of enzyme activity of CFE. Protein content was determined by using Qubit™ Protein Assay Kit (Thermo Fisher Scientific, Dublin, Ireland).

Aromatic aminotransferase (ArAT) activity was performed by following the conversion of phenylalanine to phenylpyruvate, as described by Stefanovic et al. (2017). *L. paracasei* DPC4206 was used as a control as its ArAT activity has been

reported previously (Stefanovic et al., 2017). Glutamate dehydrogenase (GDH) assay was performed based on the principle described by Kieronczyk et al. (2003), with the modifications defined by Stefanovic et al. (2017). *L. paracasei* DPC4206 and DPC4536 were used as a control as its GDH activity has been described previously (Stefanovic et al., 2017). Protein content was determined as described above and the results were expressed as the number of units of activity per mg of protein. The total quantity of the enzyme that resulted in an increase of absorbance of 0.01 per min corresponded to one unit (U) of activity.

5.3.5 Milk Protein Hydrolysis and Free Amino Acid analysis

To prepare samples for the determination of milk protein hydrolysis and free amino acid (FAA) levels, after propagation of *M. caseolyticus* subsp. *caseolyticus* strains for 24 h at 37°C in TSB, cultures were inoculated at 1% in to 20 mls of commercial ultra-high-temperature (UHT) lactose free milk (LFM) (Friendly Farms lactose free milk, Aldi, Ireland). The inoculated LFM (triplicate for each strain) was incubated for 24 h at 37°C. Cell counts were evaluated at time 0 h and 24 h. After 24 h incubation, samples were stored at -20°C and were defrosted at room temperature before the analysis of milk protein hydrolysis and FAA using Reversed-Phase High-Performance Liquid Chromatography (RP-HPLC) and Ion Exchange column, respectively.

RP-HPLC was used to qualitatively assess the extent of hydrolysis of the major milk proteins by the action of cell wall bound proteinases. The analysis was performed according to the method described previously (Mounsey and O'Kennedy, 2009). Separation was performed on an Agilent Poroshell 300SB-C18 (75 × 2.1 mm i.d.) column (Agilent Technologies, CA, USA). The HPLC system consisted of an Agilent 1200 Separation Module with DAD Detector and Agilent Chemstation Software. All samples were evaluated in triplicate.

Samples were first subjected to deproteinization for the analysis of free amino acid (FAA) as described previously (McDermott et al., 2016), on the soluble N extracts using a Jeol JLC-500/V amino acid analyser (Jeol, Garden city, Herts, UK) fitted with a Jeol Na^+ high performance cation exchange column. The chromatographic analysis was conducted at pH 2.2. All samples were evaluated in triplicate and results are expressed as µg ml^{-1} of LFM.

5.3.6 Nitrogen Substrate Utilisation with Biolog Phenotypic Microarray PM3 Plate

The ability of *M. caseolyticus* subsp. *caseolyticus* strains to utilize a range of nitrogen sources was analysed with high throughput phenotypic microarrays PM3 plate (Biolog, Hayward, CA, USA) according to the published procedures (Bochner et al., 2001). All reagents and materials for the phenotypic studies were purchased from Biolog (Biolog, USA). The cells from passage four were scrapped from the surface of TSA plates and suspended in PM3 inoculation fluid containing Dye Mix H; 100µl of a 1:200 dilution of cell suspension at 81% transmittance was added to each well of the PM3 plate. IF-0a GN/GP base inoculating fluid was prepared and plates were inoculated and incubated in the OmniLog incubator for 72 h. Data were collected every 15 min and analysed using the Biolog Kinetic and Parametric software (Biolog, Hayward, CA, USA). Phenotype diversities were evaluated based on the area differences under the kinetic curves of color formation. The experiment was conducted twice. The data from PM3 was quantitatively analysed with two modules kinetic and parametric of the Biolog phenotype Microarray software. Area under the curve values (AUC) were extrapolated from the parametric module for each well after subtracting A1 well with kinetic module.

5.3.7 Quantitative Analysis of Enzymes Involved in Lipolysis

To determine esterase activity, the conversion of p-nitrophenol butyrate to p-nitrophenol, and butyric acid and p-nitrophenol octanoate to p-nitrophenol, and octanoic acid, was measured using a transparent 96 well microplate (Sarstedt) as described by Bertuzzi (2017). The assay mixture was composed of a buffer (100 mmol l-1 sodium phosphate, 150 mmol l-1 sodium chloride, 0.5 % v/v triton X-100, at pH 7) and a substrate (50 mmol l^{-1} p-nitrophenol butyrate; 50 mmol l^{-1} p-nitrophenol octanoate in acetonitrile). In each well, 50 µl of buffer, 50 µl of CFE and 10 µl of substrate were mixed and absorbance was measured after 1h of incubation at 37°C at 400_{nm} (Synergy HT; Bio-Tek Multi Detection Plate Reader). The amount of p-nitrophenol released was determined from a standard curve obtained for a set of standards ranging from 0 to 500 nmol of p-nitrophenol. Protein content was determined as described above. The activity was expressed as µmol of p-nitrophenol released per mg of protein.

5.3.8 Semi-Quantitative Assay for Hydrolytic Activities (API ZYM)

A range of hydrolytic activities were determined calorimetrically on 19 naphtyl substrates using the API-ZYM kit system (BioMérieux, Hampshire, UK). The assay was performed according to the manufacturer's instructions in triplicate.

5.3.9 Volatile Compounds Analysis by Gas Chromatography-Mass Spectrometry (GC-MS)

LFM inoculated at 1% with *M. caseolyticus* subsp. *caseolyticus* strains as previously described, were used for this analysis and 3 g of each sample was added to a 20 ml amber screw capped La-Pha-Pack headspace vials with silicone/polytetrafluoroethylene septa (Apex Scientific, Kildare, Ireland). Samples were equilibrated to 40°C for 10 mins with pulsed agitation of 5 sec at 500 rpm using an agitator on a Gerstel MPS autosampler (Anatune, Cambridge, UK). For this headspace solid phase micro-extraction (HS_SPME) a single 50/30 um Carboxen^{TM} /divinylbenzene/ polydimethylsiloxane (DVB/CAR/PDMS) fiber was used (Agilent Technologies). The fiber was exposed to the headspace above the samples for 20 min at depth of 1 cm at 40°C with agitation. The fiber was retracted and injected into the GC inlet and desorbed for 2 min at 250°C into a SPL injector with a SPME liner. The fiber was conditioned between runs using a bake out station at 270°C for 3 min using nitrogen to ensure no carry-over between samples. Injections were made on a Shimadzu 2010 Plus GC (Mason Technology Ltd, Dublin, Ireland) with an DB-624 UI (60m x 0.32mm x 1.8µm) column (Agilent Technologies) using a split/splitless injector with a 1:10 split. A merlin microseal was used as the septum (Sigma-Aldrich). The temperature of the column oven was set at 40°C, held for 5 min, increased at 5°C/min to 230°C then increased at 15°C/min to 260°C, held for 5 min yielding at total GC run time of 65 min. Helium was used as a carrier gas held at a constant flow of 1.2ml/min. The detector was a Shimadzu TQ8030 mass spectrometer detector (Mason Technology, Dublin, Ireland), ran in single quad mode. The ion source temperature was 220°C and the interface temperature was set at 260°C. The MS mode was electronic ionization (70ev) with the mass range m/z scanned between 35 and 250 amu.

All samples were analysed in the same GC run. A set of external standards (dimethyl sulphide, benzaldehyde, cyclohexanone, butyl acetate, acetone and ethanol (Sigma-Aldrich) at concentrations of 10 ppm were run at the start of the sample set to ensure that both the HS-SPME extraction and MS detection were within specification. Blanks (empty vials) were injected regularly to monitor possible carry-over.

Chromatograms obtained by GC analysis were converted to .cdf format and processed by TargetView® (Markes International, Llantrisant, UK). Identification of compounds was based on the results of a comparison with the NIST 2011 Mass Spectral Library (Scientific Instrument Services, NJ, USA) and an in-house library produced from external standards (where available) and confirmed by calculating linear retention indices (Van den Dool and Kratz, 1963).

5.3.10 Satistical analysis

All enzymatic assays results were statistically analysed using one-way analysis of variance (ANOVA) (Biological replicates n =3) with Minitab (Minitab 17, Minitab Inc, Coventry, UK) followed by least significant difference (LSD) *post hoc test.* ANOVA were also used for testing the significance of differences in PM3, FAA analysis and GC-MS data. Selected substrates from PM3 and total GC-MS volatile profiles were then visualized as heat maps with hierarchical clustering using single linkage in R with the pheatmap package (R Core Team 2015, R foundation for Satistical Computing Austria, https://R-project.org) and principal component analysis (PCA) was used for the analysis of GC-MS data with R statistical software package (https://R-project.org).

5.4 Results

5.4.1 High CEP activity in *M. caseolyticus* subsp. *caseolyticus* strains is associated with a dairy origin

Six *M. caseolyticus* subsp. *caseolyticus* strains which were widely distributed both geographically and by the source from which they were isolated (including cheese sample, bovine milk and whale skin) were examined for their proteolytic activities (Table 1). Preliminary screening of the proteolytic activity was determined by examining casein degradation on skim milk agar plates, which is observed as the development of a transparent zone of clearing around the bacterial growth. Of the six strains examined, zones of clearing were observed for strains DPC6291, DPC7170, DPC7171 and ATCC13548^{T}, whereas strains ATCC13518 and ATCC51835 were negative for this phenotype after 24 h incubation (Fig. 1A). This suggests that while dairy-derived strains were capable of casein degradation on skim milk agar plates, non-dairy associated strains were not.

Table 1: Details of the six *M. caseolyticus* subsp. *caseolyticus* strains analysed in this study**.**

Bacterial strain	Country	Isolated from	Isolation Year	Accession number	Reference
ATCC13548^{T}	USA	Raw milk	1916	PZJF00000000	(Evans, 1916)
ATCC51835	NC,USA	Whale skin	1995	SDQL00000000	(Kloos et al., 1998)
ATCC13518	unknown	unknown	1980	ND	(Roberts, 1985)
DPC6291	Cork,Ireland	cheese	2017	SDQM00000000	(Mazhar et al., 2018; Mazhar et al., 2019)
DPC7170	Cork,Ireland	Cow's milk	2017	SDQK00000000	(Mazhar et al., 2018; Mazhar et al., 2019)
DPC7171	Cork,Ireland	Cow's milk	2017	SDQJ00000000	(Mazhar et al., 2018; Mazhar et al., 2019)

To quantify the level of CEP activity demonstrated by these strains, a kit based on the proteolysis of BODIPY® FL-labeled casein derivatives was employed. This assay is based on the principle that the measured increase in fluorescence is proportional to the proteinase activity. While all six *M. caseolyticus* subsp. *caseolyticus* strains demonstrated CEP activity when measured by this assay, the levels varied significantly between dairy-derived and non-dairy strains (Fig. 1B). The CEP activity was expressed as measured fluorescence and ranged from 130.7 arbitrary fluorescence units for strain ATTC51835, isolated from whale skin, to 682 arbitrary fluorescence units for strain DPC6291, isolated from cheese (Fig. 1B), which corresponds to fluorescence measured when standard solutions of trypsin in the range of 0.4– 10.0 µg ml^{-1} were used (data not shown). *L. lactis* subsp. *cremoris* Wg2 was used as a CEP-positive control strain, having been confirmed as such in a previous study (Nikolić et al., 2009; Stefanovic et al., 2017). CEP activity for Wg2 was measured at 220 arbitrary fluorescence units. The dairy-derived *M. caseolyticus* strains DPC6291, DPC7170, DPC7171 and ATCC13548 all had statistically higher levels of activity than Wg2 (Fig. 1B; Table S1).

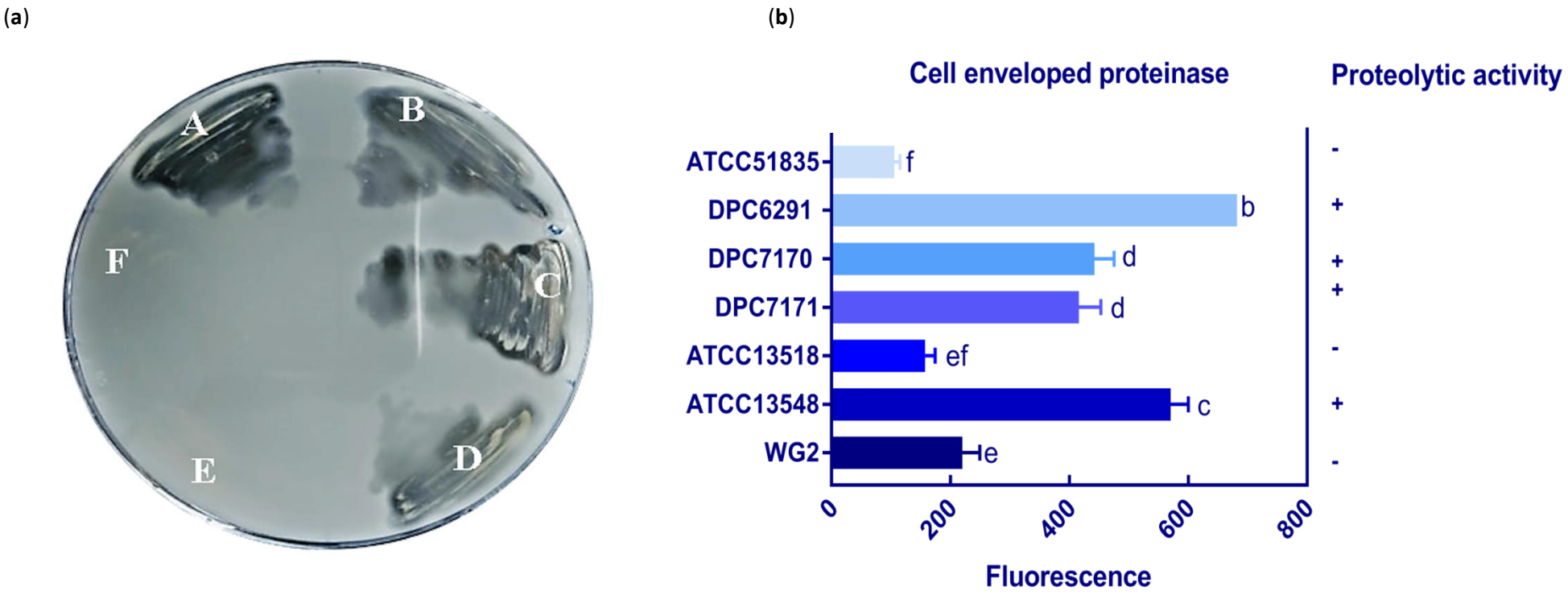


Fig. 1 (a) Proteolytic action of six *M. caseolyticus* subsp. *caseolyticus* strains on RSM agar medium indicated strains; (A) DPC6291, (B) DPC7170, (C) DPC7171 and (D) ATCC13548 are proteolytic as degradation of the substrate (casein), incorporated in the agar plate by enzyme protease is observed as the development of a transparent zone in these plates in comparison to (E) ATCC13518 and (F) ATCC51835 which were negative after 24 h incubation. **(b)** Cell envelope proteinase (CEP) activities of *M. caseolyticus* subs *caseolyticus* as determined by EnzCheck® kit following incubation at 37 °C for 24 h. Bars sharing the same letter show no significant difference according to least significant difference (LSD) test ($p<0.05$). Strains were analysed in triplicate. Error bars present standard deviation. The graph presents activities of seven representative strains, including the ATCC 51835 with the lowest activity observed, and *Lactococcus lactis* subsp. *cremoris* Wg2, which was used as a positive control

To confirm the specificity of the observed cell enveloped proteinase activities, reversed-phase high-performance liquid chromatography (RP-HPLC) profiles were generated for each strain following growth in UHT lactose-free milk (LFM). Again, *L. lactis* subsp. *cremoris* Wg2 was used as a CEP-positive control. *M. caseolyticus* subsp. *caseolyticus* DPC6291, isolated from cheese, demonstrated extensive hydrolysis of the majority of the milk and whey proteins (Fig. 2A). This was also observed for the other CEP-active strains described above (DPC7170, DPC7171, ATCC13548 (data not shown). In contrast, very little hydrolysis of the milk and whey proteins was observed for strain ATCC51835 (Fig. 2B), isolated from whale skin and showing comparatively low CEP activity in the previous assay. This was also the case for ATCC13518, whose source is unknown, and for *L. lactis* subsp. *lactis* WG2. Proteolytic digestion of κ-casein and αS_2-casein was highest for DPC6291 and varied amongst DPC7170, DPC7171 and ATCC13548, whereas no hydrolysis of these fractions was observed for ATCC51835 and ATCC13518. All strains demonstrated the ability to hydrolyze the αS_1-casein and β-casein fractions; however, these fractions were predominantly hydrolyzed with the dairy strains DPC6291, DPC7170, DPC 7171 and ATCC13548. Limited activity was observed with ATCC51835 and ATCC13518 against αS_1-casein and β-casein and peaks for β-casein hydrolysis overlapped and were similar to the control strain Wg2. In fact, ATCC51835 and ATCC13518 strains demonstrated comparatively weak proteolytic activity towards majority of the casein fractions. In addition, the whey proteins preceding the casein fractions were, for the most part, intact across all strains; however, DPC6291 was the only strain to demonstrate moderate ability to hyrdolase β-lactoglobulin. Overall, the extent of hydrolysis of the casein fractions, as observed in this assay, correlated with the activities observed in the CEP assay. The strains showing a high CEP activity, as revealed by the BODIPY® FL-labeled casein assay, illustrated strong affinity towards different fractions of casein, and were derived from dairy-associated sources. In contrast, those with low CEP activity exhibited a low extent of hydrolysis of casein (Fig. 2B) and were from non-dairy sources as in the case of ATCC51835, derived from whale skin, or in the case of ATCC13581, the source of which is unknown.

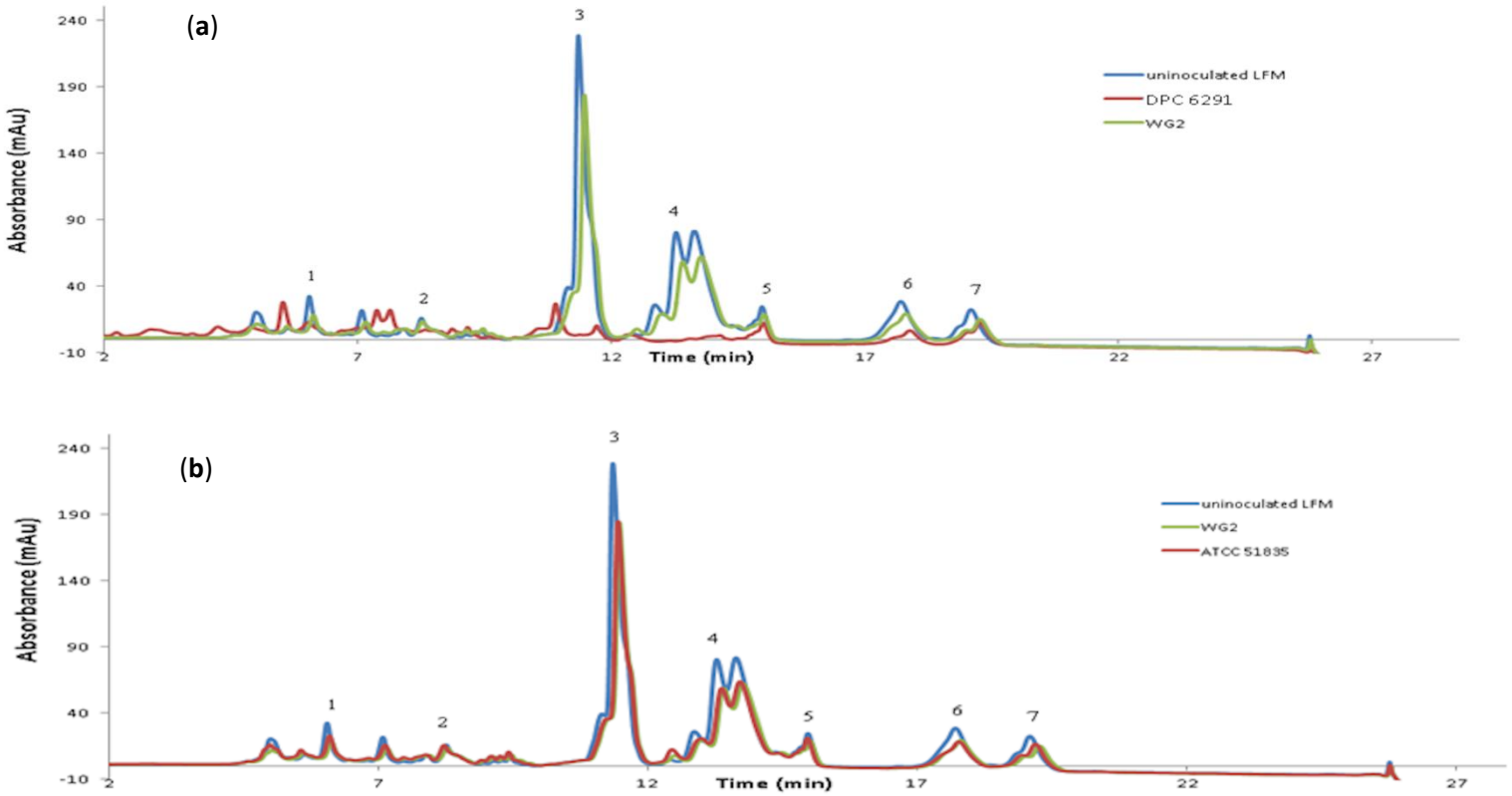


Fig. 2 RP-HPLC profiles for the following six strains (a) DPC6291, (b) ATCC51835 analysed along with controls uninoculated LFM and *Lactococcus lactis* subsp. *cremoris* Wg2. Peaks representing fraction of casein and whey proteins (1) K-casein, (2) αs2-casein (3) αs1- casein (4) β- casein (5) α-lactalbumin (6) β-lactoglobulin a (7) β-lactoglobulin b.

5.4.2 Limited Downstream Proteolytic Enzyme Activity in *M. caseolyticus* subsps. *caseolyticus* strains in Comparison to LAB Flavour Formers

The activities for PepN, PepX, PepC, PepA and PepV peptidases were examined in the *M. caseolyticus* subsp. *caseolyticus* strain bank. *L. paracasei* DPC4206 was used as a positive control for activities of PepN, PepX and PepC as these had been previously reported (Stefanovic et al., 2017). Surprisingly, all *M. caseolyticus* strains showed very limited activities towards the various substrates tested (Fig. 3). The PepN activities measured ranged from 0.3 to 6.7 nmol para-nitroaniline (per min per mg protein) for DPC6291 and ATCC13548, respectively. PepX activities ranged from 3.07 for DPC7170 to 7.43 nmol para-nitroaniline (per min per mg protein) for ATCC13548, while, PepC activities ranged from 1.26 for DPC7170 to 5.14 nmol para-nitroaniline (per min per mg protein) for DPC6291. PepV activities ranged from 3.09 for DPC7171 to 8.53 nmol para-nitroaniline (per min per mg protein) for DPC7170 and PepA activities ranged from 0.59 to 5.43 nmol para-nitroaniline (per min per mg protein) for ATCC 51835 and ATCC13548, respectively. PepN, PepC, PepX and PepV activities expressed by the positive control strain DPC4206 were in the range of 58.7, 48.23, 46.4 and 44.02 nmol para-nitroaniline (per min per mg protein), respectively. In all cases (except PepA), the activities of the peptidases exhibited by the *M. caseolyticus* subsp. *caseolyticus* strains were significantly lower in comparison to the *L. paracasei* DPC4206, a LAB strain with proven peptidolytic ability (Fig. 3, Table S1).

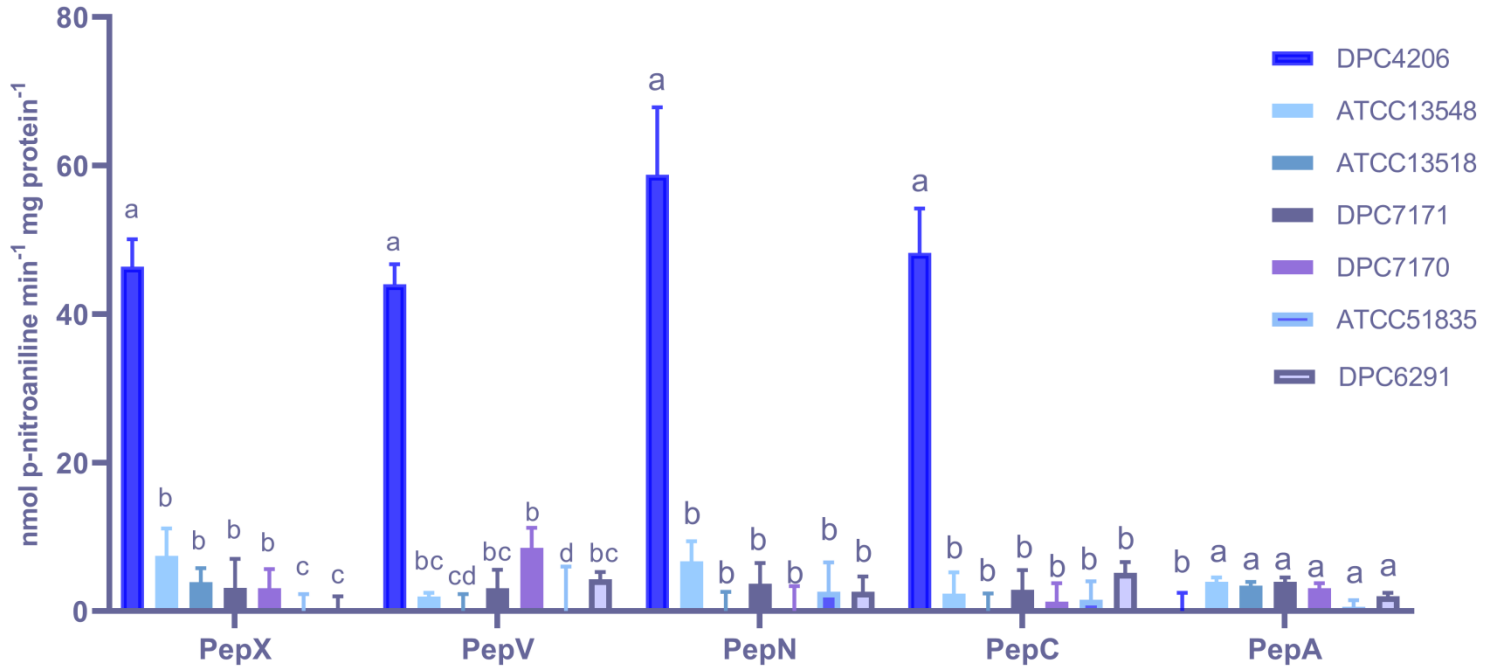


Fig. 3: Aminopeptidase (PepX, PepV , PepN, PepC and PepA) activities of six *M. caseolyticus* subsp. *caseolyticus* (ATCC13548, ATCC13518, DPC7171, DPC7170, DPC6291 and ATCC51835) along with Lactobacillus paracasei DPC4206 determined by measuring cleavage of corresponding chromogenic substrates (L-Lys-pNA, Arg-pNA and Gly-Pro-pNA, H-Glu-pNA and H-Ala-Phe-Pro-pNA) for PepN, PepC and PepX, PepA and PepV respectively. Results are expressed as nmol of released p-nitroaniline min-1 mg protein-1. Bars sharing the same letter show no significant difference according to least significant difference (LSD) test ($p<0.05$). Strains were analysed in triplicate. Error bars present standard deviation.

To examine the ability of the strains to generate free amino acids (FAA), the levels of FAA in LFM milk following fermentation with each of the six *M. caseolyticus* subsp. *caseolyticus* strains were analysed with HPLC. The analysis indicated no significant ($P < 0.05$) differences between the test strains and the uninoculated control-LFM, except for the amino acids histidine, proline and tryptophan (Fig. 4, table S1). The most significant difference was the release of tryptophan by the protease active strains, highest in DPC6291 of 103.22 µg ml^{-1}/lactose free milk. These results suggest that the *M. caseolyticus* subsp. *caseolyticus* strains tested, irrespective of source, display poor peptidolytic activity and as a result, cannot generate significant levels of FAA when provided with a rich protein source such as milk.

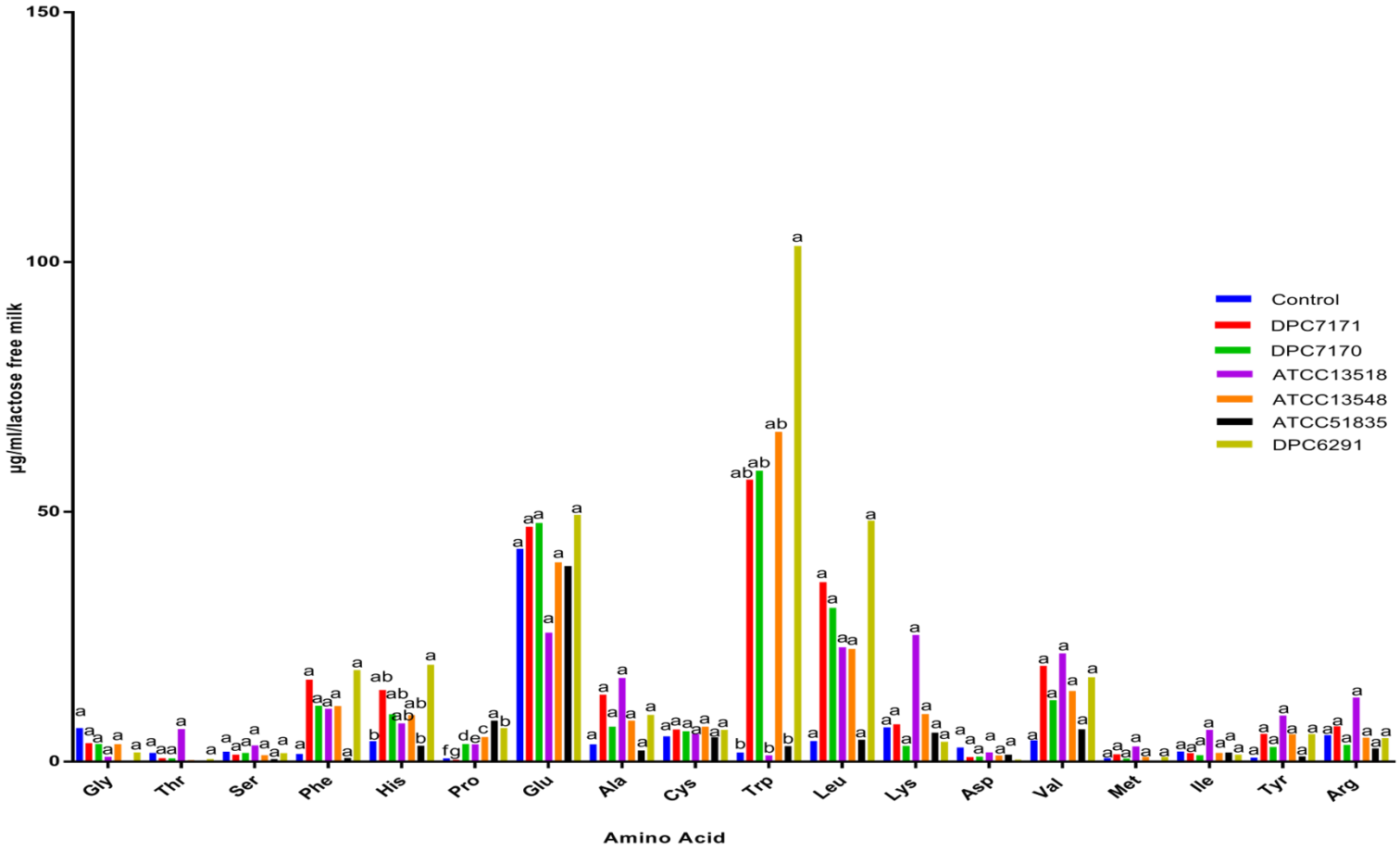


Fig.4: Free amino acid analysis of the six *M. caseolyticus* subsp. *caseolyticus*. All samples were evaluated in triplicate. Results are expressed as µg ml-1 of lactose free milk. Bars sharing the same letter show no significant difference according to least significant difference (LSD) test ($p<0.05$).

The activities of the other downstream enzymes with an important role in the proteolytic cascade leading to flavour compound formation were also examined. When ArAT activity was measured, the *M. caseolyticus* subsp. *caseolyticus* strains showed significantly lower levels of activity when compared to *L. paracasei* DPC4206 which was used as positive control strain as its ArAT activity has been previously reported (Stefanovic et al., 2017). The ArAT activity expressed by DPC4206 in this study was 4.07 μmoles of phenylpyruvate released per mg protein (slightly higher than the previous study). The activity of ArAT measured for *M. caseolyticus* subsp. *caseolyticus* ranged from 1.5 μmoles of phenylpyruvate released per mg protein for ATCC13518 to 2.07 μmoles of phenylpyruvate released per mg protein for DPC6291. Statistical analysis revealed no significant differences between the *M. caseolyticus* subsp. *caseolyticus* strains for ArAT activity (Fig. 5a). The GDH activity of the strains was also analysed and compared to *L. paracasei* strains DPC4206 and DPC4536, strains whose GDH activities had been determined in previous study (Stefanovic et al., 2017). The GDH activity for *M. caseolyticus* subsp. *caseolyticus* strains ranged from 4.2 for strains ATCC51835 and ATCC7171 to 7.1 U mg^{-1} of protein for ATCC13548. The GDH activity for the control strains DPC4206 and DPC4536 ranged from 13.3 to 17.3 U mg^{-1} of protein (correlating with previous report) and differences in GDH activities were shown to be significant (Fig. 5b). Overall, the ArAT and GDH activities measured across the *M. caseolyticus* subsp. *caseolyticus* strains were significantly lower than the LAB strains tested in this study.

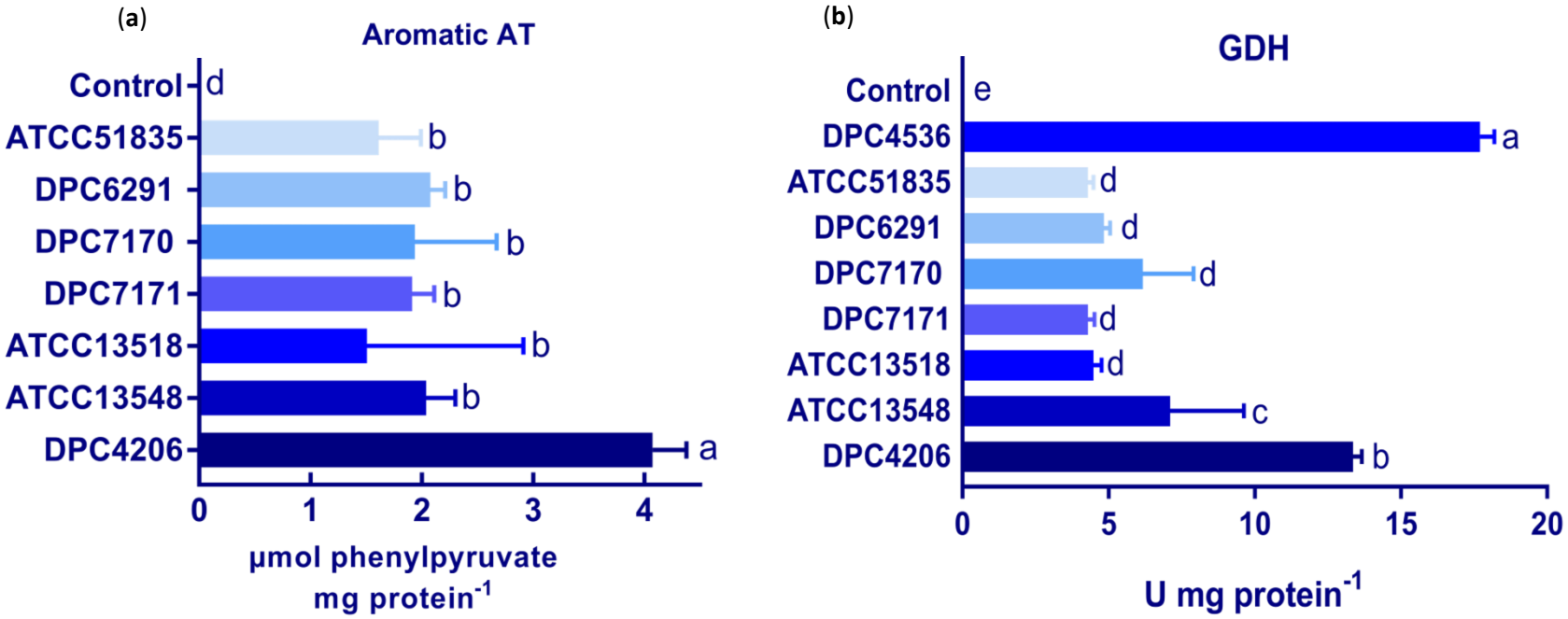


Fig. 5(a) Aromatic aminotransferase (ArAT) activities of *M. caseolyticus* subsp. *caseolyticus* strains along with *Lactobacillus paracasei* DPC4206 determined by measuring the absorbance of phenylpyruvate, the final product of transamination between phenylalanine and α-ketoglutarate. Results are expressed as µmol of released phenylpyruvate/(min*mg of protein). (b) Glutamate dehydrogenase (GDH) activities of strains of *M. caseolyticus* subsp. *caseolyticus* along with *Lactobacillus paracasei* DPC4206 and DPC 4536 by following change in absorbance during a reaction catalysed by GDH enzyme in which glutamic acid is converted to α-ketoglutarate in the presence of NAD+ . Results are presented as units of enzyme activity per mg of protein, where the unit represents the amount of enzyme giving an increase of absorbance of 0.01 per 1 min. Bars sharing the same letter show no significant difference according to least significant difference (LSD) test ($p<0.05$). Strains were analysed in triplicate. Error bars present standard deviation.

The ability of the *M. caseolyticus* subsp. *caseolyticus* strain bank to metabolise a diverse range of nitrogen sources was examined with high throughput PM3 plates from Biolog (Biolog, CA, USA). Overall, the six strains examined showed little activity, as they were only able to metabolise 13.6% of the nitrogen sources tested, the majority of which were fatty acids (e-amino-N-caproic acid, D,L-a-amino-caprylic acid, D-amino-N-valeric acid) and di-peptides (Ala-His, Gly-Glu, Met-Ala). This further confirms the limited ability of the members of this subspecies to catabolize free amino acids. Differences in area under the curve (AUC) for all substrates were analysed using ANOVA and activities demonstrated no significant differences between strains. A detailed list of the AUC values of each substrate in each well of PM3 along with ANOVA analysis can be found in Table S1. The heat map in Fig.6 illustrates the substrates most effectively metabolised by these strains in the PM3 plate.

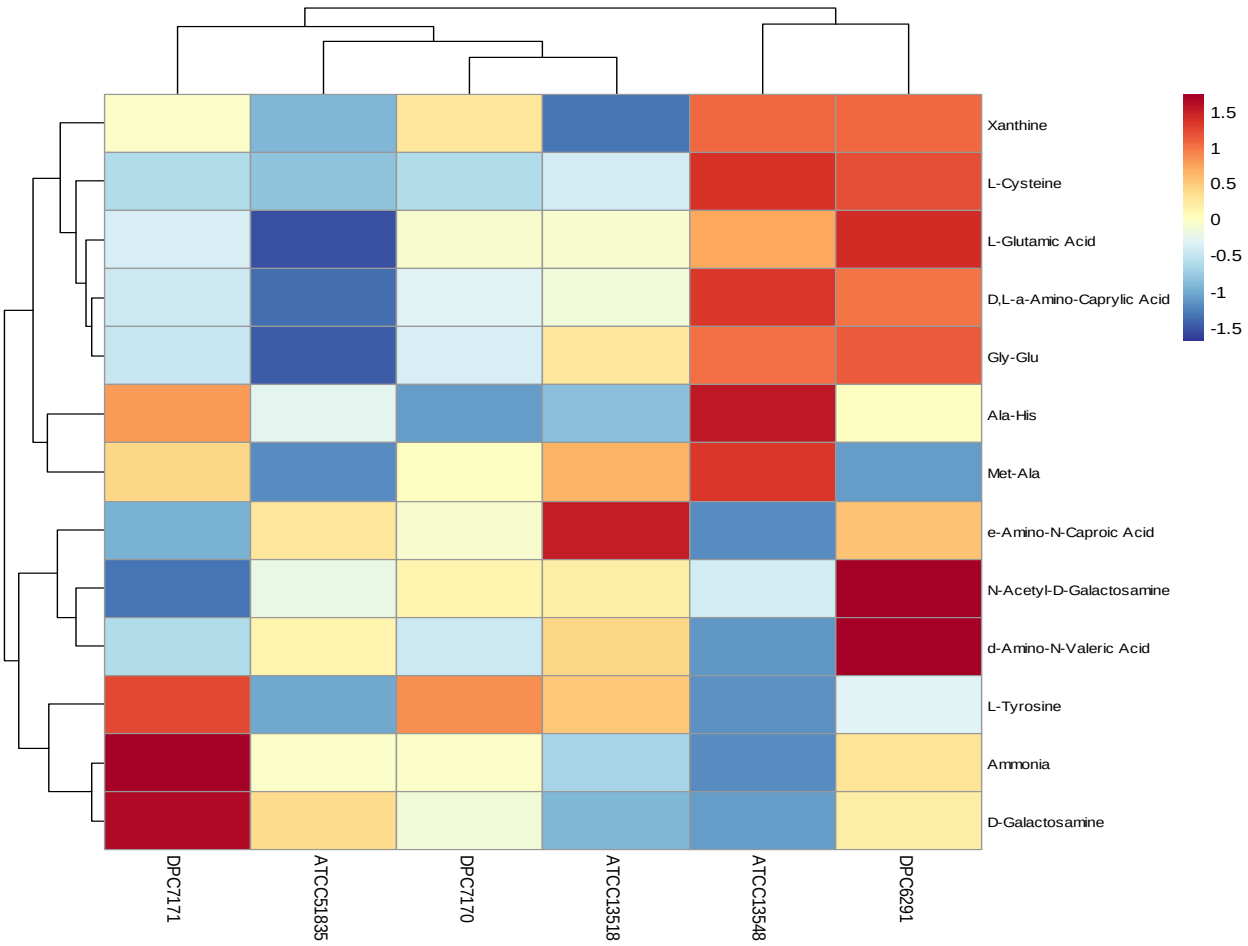


Fig.6: Heat map of substrates that were most effectively metabolised by *M. caseolyticus* subs *caseolyticus* strains in PM3 plate.

5.4.3 High Esterase Activity Observed in strains with High CEP Activity

The breakdown of lipids also plays an important role in flavour development. Lipase activity for the *M. caseolyticus* subsp. *caseolyticus* strains were qualitatively analysed on tributyrin agar as outlined by Bertuzzi, A. (2017). All strains demonstrated weak hydrolytic activity in comparison to the control strain *Y. lipolytica* DPC6266 on tributyrin agar (data not shown). Meanwhile, esterase activity was measured using both p-nitrophenol butyrate and p-nitrophenol octanoate as substrates. The highest activity towards p-nitrophenol butyrate was obtained with DPC6291 of 0.79 µmol of p-nitrophenol released per mg of protein and lowest with ATCC51835 of 0.33 µmol of p-nitrophenol released per mg of protein. Likewise, the highest activity towards p-nitrophenol octanoate was observed in DPC 6291 of 0.72 µmol of p-nitrophenol released per mg of protein and lowest with ATCC 51835 of 0.33 µmol of p-nitrophenol released per mg of protein (Fig. 7). The esterase activities were statistically significant and comparatively higher amongst the high CEP active strains (DPC6291, DPC7170, DPC7171 and ATCC13548). The esterase activity of dairy related *M. caseolyticus* subsp. *caseolyticus* strains was also significant in comparison to the *L. paracasei* DPC4206.

Additionally, hydrolytic activities towards a number of lipase, esterase, proteinase and peptidase substrates were evaluated for all the *M. caseolyticus* subsp. *caseolyticus* strains along with *L. paracasei* DPC4206 and *Y. lipolytica* DPC6266 determined using the semi-quantitative API-ZYM kit system. All *M. caseolyticus* subsp. *caseolyticus* strains were positive for esterase activity (C4:0), and weakly positive for esterase (C8:0), whereas *Y. lipolytica* DPC 6266 was positive towards both and DPC4206 was negative for both esterase (C4:0) and esterase (C8:0). All *M. caseolyticus* subsp. *caseolyticus* strains were negative for lipase, peptidase, trypsin and α-chymotrypsin like proteinase and glycosidase, in comparison to *L. paracasei* DPC4206 which demonstrated hydrolytic activities of the peptidase enzymes leucine arylamidase, valine arylamidase, and cystine arylamidase. Glycosidase and phosphatase activities were also observed in DPC4206. *Y. lipolytica* DPC6266 was also positive for leucine arylamidase, acid phosphatase, phosphohydrolyase and β-glucosidase, and, negative for lipase (C14) using this assay (Table 2). The results from this analysis further confirm the lipolytic activity towards butyrate (C4), whereas no peptidolytic activity was observed across the members of the *M. caseolyticus* subsp. *caseolyticus*.

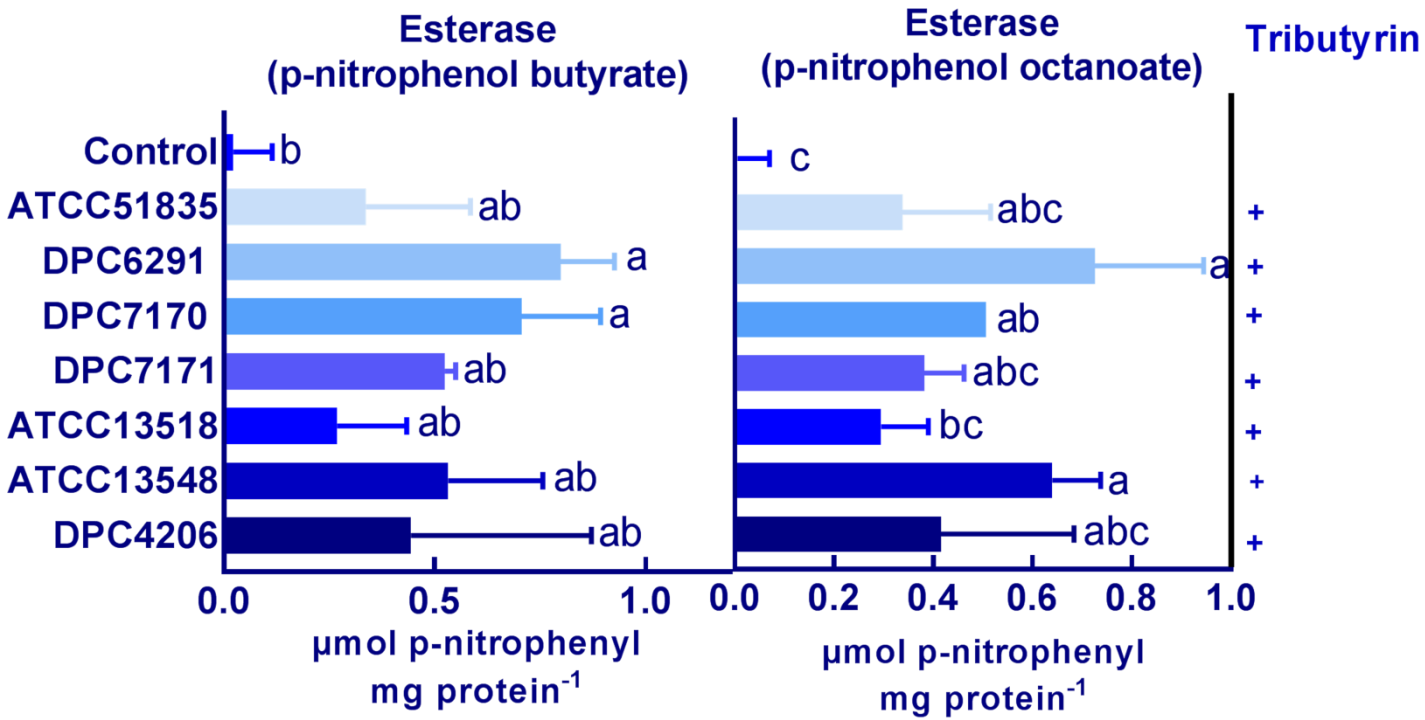


Fig. 7: Esterase activity on p-nitrophenol butyrate and p-nitrophenol octanoate for the six *M. caseolyticus* subsp. *caseolyticus* strains along with DPC 4206 are presented. Activity is expressed in µmol of p-nitrophenol released per mg of protein^{-1}. Bars sharing the same letter show no significant difference according to least significant difference (LSD) test ($p<0.05$). Strains were analysed in triplicate. Error bars present standard deviation. Tributyrin hydrolysis :positive; +.

Table: 2 Semi-Quantitative assay of enzyme activities of *M. caseolyticus* subs *caseolyticus* strains and control strains.

		Strains							
		DPC6291	ATCC 51835	ATCC13548	ATCC13518	DPC7170	DPC7171	DPC 4206	DPC6266
Esterase activity	Esterase (C4:0)	+	+	+	+	+	+	-	+
	Esterase (C8:0)	(+)	(+)	(+)	(+)	(+)	(+)	+	+
Lipase activity	Lipase (C14:0)	-	-	-	-	-	-	-	-
Peptidase activity	Leucine arylamidase	-	-	-	-	-	-	+	+
	Valine arylamidase	-	-	-	-	-	-	+	-
	Cystine arylamidase	-	-	-	-	-	-	+	-
Proteinase activity	Trypsin	-	-	-	-	-	-	-	-
	α-chymotrypsin	-	-	-	-	-	-	-	-
Phosphatase activity	Acid phosphatase	-	-	-	-	-	-	+	+
	Alkaline phosphatase	-	-	-	-	-	-		
	Phosphohydrolyase	+	+	+	+	+	+	+	+
Glycosidase activity	α-Galactosidase	-	-	-	-	-	-	-	-
	β-Galactosidase	-	-	-	-	-	-	+	-
	β-Glucuronidase	-	-	-	-	-	-	-	-
	α-Glucosidase	-	-	-	-	-	-	+	-
	β-Glucosidase	-	-	-	-	-	-	+	+
	β-Glucosaminidase	-	-	-	-	-	-	+	-
	α-Mannosidase	-	-	-	-	-	-	+	-
	α-Fucosidase	-	-	-	-	-	-	+	-
Color shading shows green(Positive), yellow (Weak positive) and red (Negative).									

5.4.4 Comparative Genomics Reveals the Conservation of the Proteolytic and Lipolytic System Components in *M. caseolyticus* subsp. *caseolyticus* strains

A systematic genome-wide analysis of components of the proteolytic and lipolytic systems from five draft genomes of strains from the *M. caseolyticus* subsp. *caseolyticus* strain bank was conducted. Functional genome distribution (FGD) was performed on all of five strains to identify ORFs specific to each of the strains and also ORFs that were conserved between these strains (Altermann, 2012). Mining of conserved genes within the five genomes using FGD at an e-value of $1e^{-10}$ with zero mismatches identified 1,314 conserved ORFs. The numbers of non-conserved genes varied from 113 to 150 for each of the strains. Subsequently, the presence of ORFs encoding genes for enzymes involved in proteolysis and lipolysis was examined across all five genomes. Interestingly, irrespective of the niche from which they were derived, the distribution of the genes encoding the components of the proteolytic and lipolytic processes in *M. caseolyticus* subsp. *caseolyticus* appeared to be conserved across strains (A detailed list of genes with GI codes can be found in Table S1; data not shown). The number of genes encoding the proteinases, peptide transporters, peptidases, aminotransferases, dehydrogenases, lyases, lipases and esterases are shown in Table 3. The distribution of proteolytic components such as the cell wall bound proteinase (*prt*P) and the presence of oligopeptide (OPP) transport gene cluster (opp*ABCDF*) were widely distributed across all five genomes. The general broad specificity peptidases, PepN and PepC that are repoted to be widely distributed across LAB were not identified (Liu et al., 2010). Conversely, the presence of exopeptidases with different substrate specificities such as PepA that has a narrow activity towards only acidic amino acids (Glu and Asp substrates), PepP (a proline-specific peptidase), PepM (a methionine-specific peptidase), PepV (broad specificity dipeptidase), PepT (capable of hydrolyzing only tripeptides), carboxypeptidases and endopeptidases such as PepF were identified across the five genomes. Further, genes encoding the downstream enzymes of the proteolytic pathways involved in the transamination processes, which includes ATs (*ilv*E, *roc*D, *his*C, *glm*S) and GDH (*gdh*A) were found to be conserved across the five studied genomes. The distributions of lipolytic enzymes such as monoacylglycerol lipase (*mgl*P and *mgI*L) which are responsible for the hydrolysis of monoacylglycerol into free fatty acid and glycerol, and

carboxylesterase which catalyse the hydrolysis of various types of esters were also widely distributed across the five genomes.

Additionally, OrthoVenn 2 was used identify the distribution of shared orthologous clusters and singletons across the five genomes (Wang et al., 2015). According to this web-based software, the genomes shared 2,238 clusters constituting pan-genome, whereas, the core-genome represented in all strains was estimated in 1,780 clusters (Fig. 8), whose functions were mostly assigned to metabolic processes including the proteolytic and lipolytic catabolic pathways (a detailed list of core orthologous clusters with Gene ontology ID can be found in Table S2- data not shown). The accessory genome, composed of singleton gene clusters unique to each strain, were composed of genes whose functions were mostly unknown, or associated with virulence factors, resistance determinants and mobile genetic elements (MGEs) such as transposons, phages and plasmid proteins. In summary, both the FGD and OrthoVenn 2 comparative genomic analysis of the five strains did not reveal genetic content differences with regards the components of the proteolytic and lipolytic cascade, except for the presence of gene *yyd*H encoding a zinc metalloprotease belonging to MEORPS family M50b in DPC6291 and citrate lyase subunit beta in DPC7170.

Table 3: Distribution of enzymes responsible for flavour development in selected *Macrococcus caseolyticus* subsp. *caseolyticus* genomes. The number of identified genes is indicated. Color shading shows absence (yellow) and presence single gene (light green) or multiple genes (dark green).

Functional Category	Suggested gene name	Description/Function	Strains				
			DPC6291	DPC7170	DPC7171	ATCC13548	ATCC51835
Proteases	PrtP	Cell wall bound Proteinase	1	1	1	1	1
	clpA	Clp protease ATP-binding subunit	2	2	2	2	2
	clpP	Clp Proteolytic subunit	1	1	1	1	1
	clpP2	Clp protease proteolytic subunit 2	1	1	1	1	1
	clpX	Clp Proteolytic subunit	1	1	1	1	1
	yugP	Zinc metallopeptidase	1	1	1	1	1
	yydH	Zinc metalloprotease	1	0	0	0	0
	ytmA	Membrane-bound serine protease	1	1	1	1	1
Peptide transporters and Amino acid permeases	OppA	Oligopeptide binding protein	1	1	1	1	1
	OppB	Permease protein	1	1	1	1	1
	OppC	Permease protein	1	1	1	1	1
	OppD	ATP- binding protein	1	1	1	1	1
	OppF	ATP- binding protein	1	1	1	1	1
	brnQ	Branched-chain amino acid transport system II carrier protein	1	1	1	1	1
	lctP	L-lactate permease	1	1	1	1	1
	gntP	Gluconate permease	1	1	1	1	1
	metI	D-methionine transport system permease	1	1	1	1	1
	GlnQ	Glutamine transport ATP-binding protein	1	1	1	1	1
	GlnP	Glutamine permease	1	1	1	1	1
	opuCB	ABC-type proline/glycine betaine transport system, permease	1	1	1	1	1
	opuBB	glycine betaine transport system	1	1	1	1	1

	ansP	L-asparagine permease	1	1	1	1	1
	tcyC	L-cystine import ATP-binding protein	1	1	1	1	1
	tcyB	L-cystine transport system permease	1	1	1	1	1
	metP	Methionine permease	1	1	1	1	1
	metN	Methionine import ATP-binding protein	1	1	1	1	1
	yfmD	Fe(3+)-citrate import system permease	1	1	1	1	1
	yfmE	Fe(3+)-citrate import system permease	1	1	1	1	1
	lysP	Lysine-specific permease	1	1	1	1	1
	argO	Arginine exporter protein	1	1	1	1	0
Peptidases	Map	Methoionine aminopeptidase	1	1	1	1	1
Unique Aminopeptidases	PcP	Pyrrolidone carboxyl peptidase	1	1	1	1	1
	Ldc	LD-Carboxypeptidases	1	1	1	1	1
	PepA	Glutamyl aminopeptidase	1	1	1	1	1
	ampS	Aminopeptidase S	1	1	1	1	1
Endopeptidases	PepF	Oligoendopeptidase F	2	2	2	2	2
Dipeptidase	PepV	Dipeptidase PepV	1	1	1	1	1
Tripeptidases	PepT	Peptidase T	1	1	1	1	1
Proline peptidase	PepP	Aminopeptidase P	2	2	2	2	2
Aminotransferases	ilvE	Branched-chain-amino-acid aminotransferase	1	1	1	1	1
	rocD	Ornithine aminotransferase	1	1	1	1	1
	hisC	Histidinol-phosphate aminotransferase	1	1	1	1	1
	glmS	Glutamine--fructose-6-phosphate aminotransferase	1	1	1	1	1
Dehydrogenases	gdhA	Glutamate dehydrogenase/leucine dehydrogenase	1	1	1	1	1
	bdh	NADH-dependent butanol dehydrogenase	1	1	1	1	1
	aldA	Aldehyde dehydrogenase	1	1	1	1	1
	aladh	Alanine dehydrogenase	1	1	1	1	1
	ldh	L-lactate dehydrogenase	2	2	2	2	2

Lyases	mccB	Cystathionine gamma-lyase	1	1	1	1	1
	argH	Argininosuccinate lyase	1	1	1	1	1
	hal	histidine ammonia-lyase	1	1	1	1	1
	citE	Citrate lyase subunit beta	0	1	0	0	0
Lipases and Esterases	paaI	PaaI family thioesterase	1	1	1	1	1
	mglP	Monoacylglycerol lipase	1	1	1	1	1
	mglL	Monoacylglycerol lipase	1	1	1	1	1
	glpQ	Glycerophosphodiester phosphodiesterase	2	2	2	2	2
	est	Carboxylesterase	1	1	1	1	1

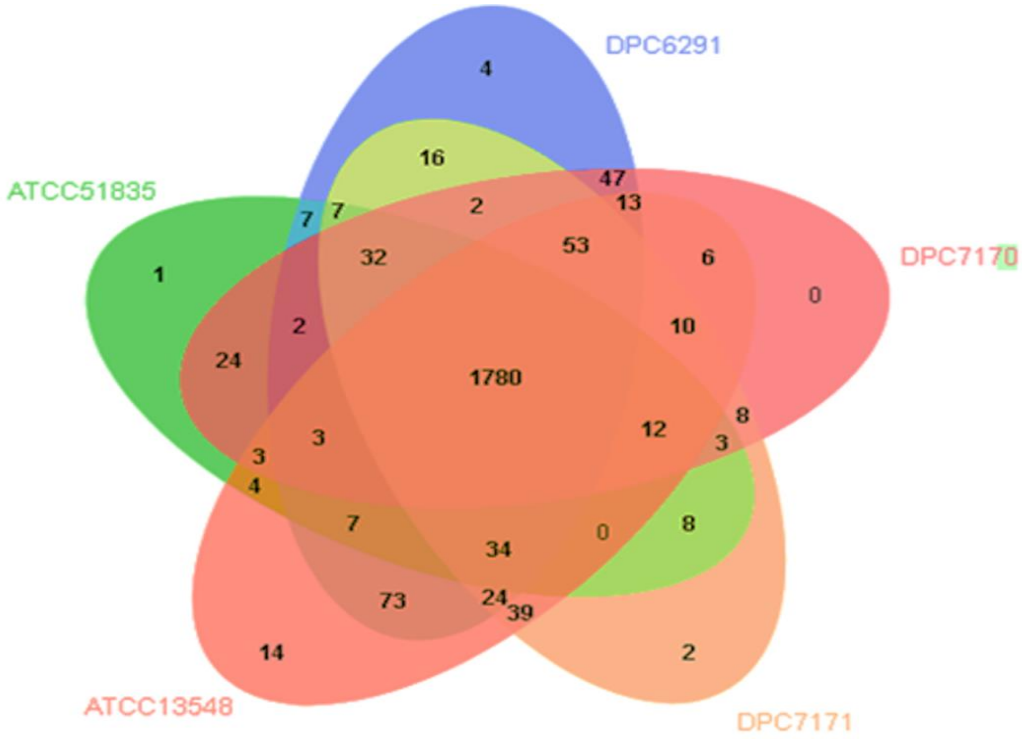


Fig. 8: Venn diagram showing shared orthologous protein clusters amongst the five *Macrococcus caseolyticus* subsp. *caseolyticus* strains. A total of 2238 clusters were identified, of which 462 were orthologous and 1776 single copy gene cluster identified with the default parameters, 1e−5 e-value cutoff for all protein similarity comparisons and 1.5 inflation value for the generation of orthologous clusters. The numbers in the diagram indicate overlapped conserved gene clusters or un-overlapped specific gene clusters in every single strain.

5.4.5 Metabolomic Analysis with GC-MS Separates *M. caseolyticus* subsp. *caseolyticus* in to Two Distinct Groups Based on High/Low CEP and Esterase Activity

HS-SPME GC-MS was used to analyze the production of volatile flavour compounds generated as a result of the metabolic activities of the selected *M. caseolyticus* subsp. *caseolyticus* strains in LFM milk after 24 h incubation. A total of 74 volatile compounds were detected in the samples including the control, of which the majority were esters (35.1%) followed by ketones (16.2%), alcohols (6.7%), aldehydes (5.4%), benzenes (5%) and some sulphurs (4.1%). Twenty-three of these volatiles were only present in the test samples. Specifically, these were twelve esters, two ketones, one benzene, two aldehydes, one alcohol, one sulphur, two acids and two phenols (Table 4).

The PCA bioplot based on the volatiles detected describes 29.3% and 23% total variation between the first and second component, respectively. There is a clear separation of the strains from the control except for strain ATCC13518 (Fig. 9). DPC6291 which has a very high CEP and esterase activity was completely separated from the rest of the strains; its position was associated with relative high levels of esters, methyl ketones, straight chain aldehyde, fatty acids and nitrogen, sulphur and phenol compounds (pyrazines, sulphurs and phenol; derived from FAA catabolism). Some of these compounds at a lower abundance were also present in the control. Overall, DPC6291 was significantly ($P< 0.05$) associated with the production of methyl butanoate, propyl butanoate, nonanal, methanethiol, acetic acid, butanoic acid (derived from lipolysis) and p-cresol (Table 4). The relatively high abundance of these fatty acid esters, straight chain aldehydes and nitrogen compounds act as a dominating factor in the discrimination of DPC6291 from other strains.

DPC7170, DPC7171 and ATCC13548 strains were positioned together and were linked to numerous esters and some methyl ketones (Fig. 9). The three strains were significantly ($P < 0.05$) associated with the production of hexyl butanoate, isopentyl hexanoate, butyl hexanoate, butyl butanoate, isoamyl isobutanoate, 2-methylbutyl butanoate, 4-pentenyl butyrate and amyl isobutyrate. The abundance of ester compounds originating from microbial esterification of different acids (derived via

lipolysis, oxidation of amino acids or glycolysis) act as a main factor in separating these strains (Table 4).

Also, the strains with comparatively low CEP and esterase actitivtiy, ATCC51835 and ATCC13518 were positioned separately from each other (Fig. 9). ATCC51835 was separated from the other strains and was associated with abundance ($P< 0.05$) of esters and a ketone such as ethyl decanoate, methyl hexanoate and 2-undecanone (Fatty acid oxidation). ATCC13518 was positioned with the control uninionculated LFM and was mainly associated with abundance ($P< 0.05$) of benzeneacetaldehyde (precursor phenylalanine), phenylethyl alcohol (precursor phenylalanine) 3-methyl butanal (FAA metabolism) and 2,3 heptanedione (FFA metabolism). Altogether, these volatiles are associated with wide range of classes (esters, benzenes, aldehydes and ketones) which act as a dominating factor and separates these low CEP and esterase active strains ATCC51835 and ATCC13518 from the comparatively highly active strains (DPC6291, DPC7170, DPC7171 and ATCC13548).

Table 4: Volatiles detected with HS-SPME GCMS absent in control and associated with test strains, reported with relative identification CAS number and the linear retention index (LRI). All results presented are obtained in this study.

Compounds	CAS	LRI	DPC6291	DPC7170	DPC7171	ATCC13548	ATCC13518	ATCC51835
			Trial 1					
Methyl butanoate	623427	748	+*	+	+	+	+	+
Propyl butanoate	644495	922	+*	+	+	+	+	+
Methyl hexanoate	106707	949	ND	ND	ND	ND	+*	+*
Butyl butanoate	109217	1019	ND	+*	+*	+*	ND	ND
Isoamyl isobutanoate	2050013	1080	ND	+*	+*	+*	ND	ND
2-Methylbutyl butanoate	51115641	1086	ND	+*	+*	+*	ND	ND
4-Pentenyl butyrate	30563316	1093	ND	+*	+*	+*	ND	ND
Amyl isobutyrate	2445729	1117	ND	+*	+*	+*	ND	ND
Butyl hexanoate	626824	1215	ND	+*	+*	+*	ND	ND
Hexyl butanoate	2639636	1216	ND	+*	ND	ND	ND	ND
Isopentyl hexanoate	2198610	1276	ND	+*	+*	ND	ND	ND
Ethyl decanoate	110383	1420	ND	ND	ND	ND	ND	+*
2,3-Heptanedione	96048	876	ND	ND	ND	ND	+*	ND
2-Undecanone	112129	1331	ND	ND	+	+	+	+*
Benzeneacetaldehyde	122781	1111	ND	ND	+	+	+*	ND
3-Methyl butanal	590863	692	+	+	+	+	+*	+
Nonanal	124196	1145	+*	+	ND	ND	ND	ND
Isopropyl Alcohol	67630	548	ND	+*	ND	ND	ND	ND
Methanethiol	74931	463	+*	ND	ND	ND	ND	ND
Acetic acid	64197	704	+*	+	+	+	+	+
Butanoic acid	107926	871	+*	+	+	+*	+	+
p-Cresol	106445	1182	+*	ND	ND	ND	ND	ND
Phenylethyl Alcohol	60128	1194	ND	ND	ND	ND	+*	ND

Legend: +*, significantly higher from other strains (P< 0.05); +, volatile detected but not significant; ND, Not detected

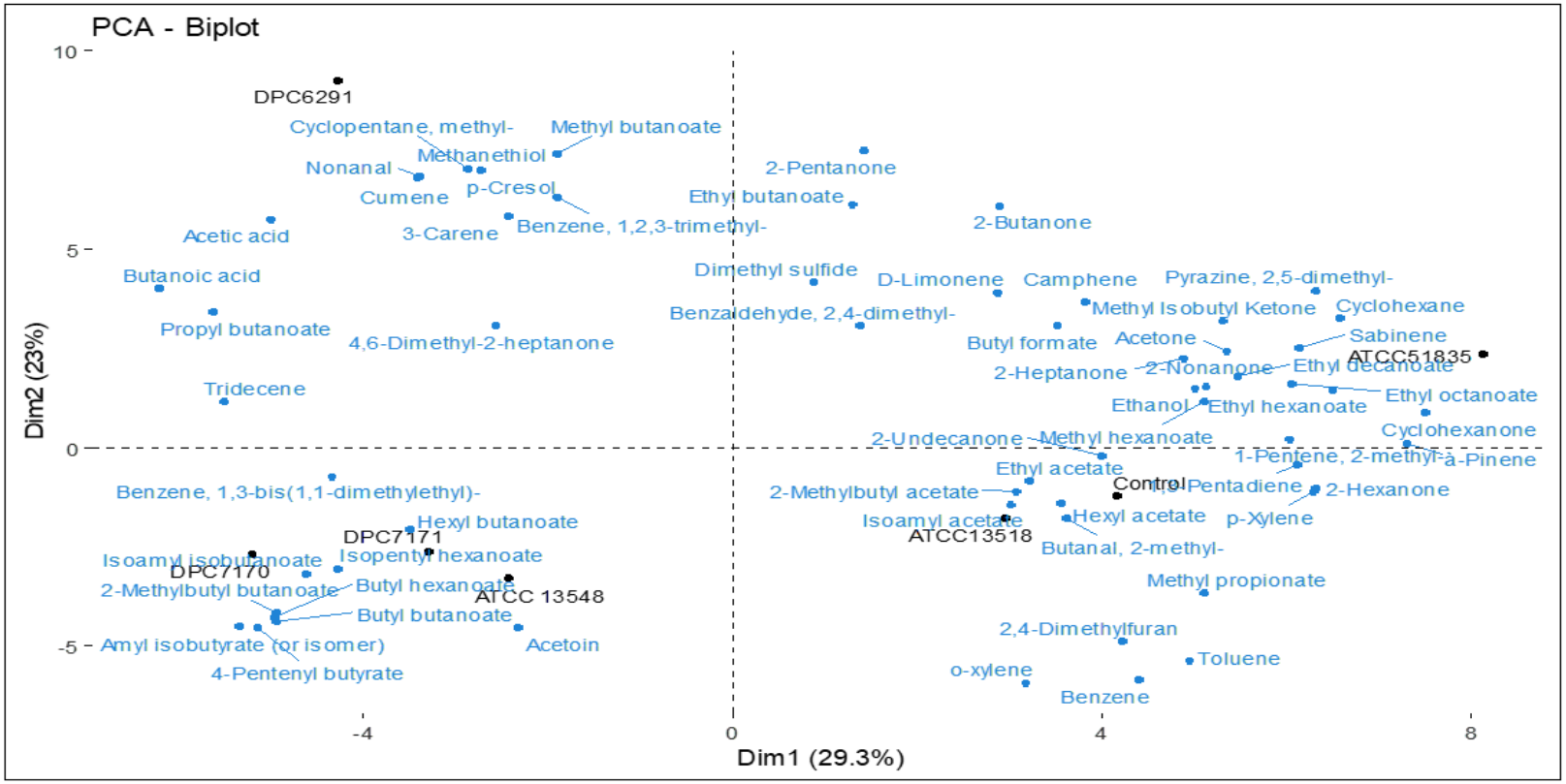


Fig.9 Principle component analysis (PCA) Bioplot illustrates the compounds responsible for the separation between test strains after 24 h incubation in LFM detected via HS-SPME GC-MS. Uninionculated LFM is used as a control

A hierarchical clustered heat map illustrates the relative abundance of variable compounds associated with different strains correlates with the PCA's illustrating the clear separation of high CEP and esterase active strains forming a separate clade (ATCC6291, DPC7170, DPC7171 and ATCC13548) from those with relatively low activity (ATCC13518 and ATCC51835) clustering closely with the control (Fig. 10).

The cell enumerations of *M. caseolyticus* subsp. *caseolyticus* strains in LFM pre and post-incubation are presented in Table 5. Compounds selected as flavour contributing which were absent in the control were according to previously published reviews of compounds considered mainly as flavour contributors in cheeses are presented in Table 6 (Curioni and Bosset, 2002).

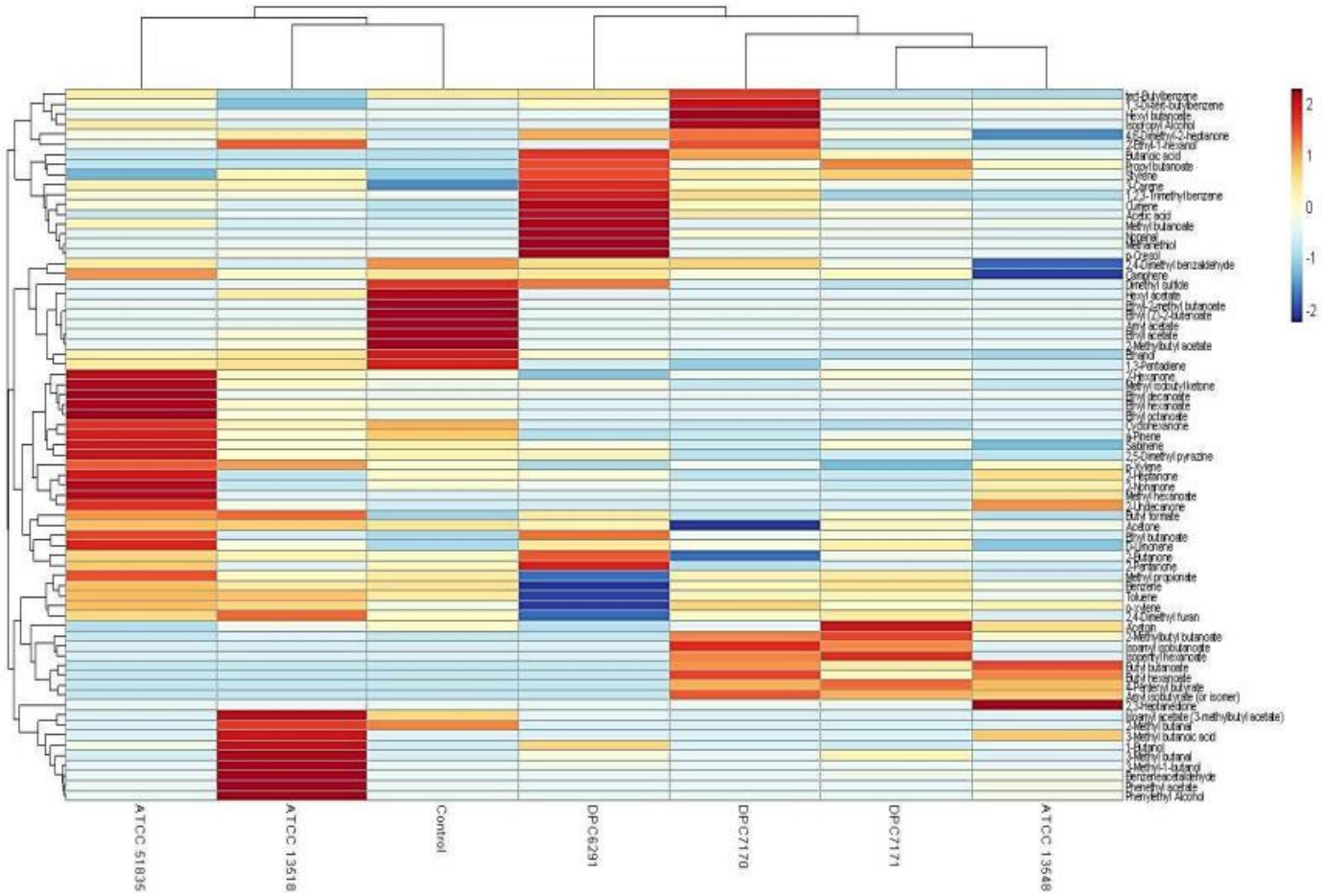


Fig.10: Hierarchical clustered heat map illustrates Trial 1 result of relative abundance of variable compounds separating high CEP and esterase active (ATCC13548, DPC7171, DPC7170 and DPC6291) strains from those with comparatively low CEP and esterase activity (ATCC 135185 and ATCC 51835) *M. caseolyticus* subsps. *caseolyticus* after 24 h incubation in LFM detected via HS-SPME GC-MS.

Table 5: Cell enumeration presented in log CFU ml^{-1}, in Lactose free milk before (t = 0 h) and after incubation at 37°C (t = 24 h). pH values; in LFM after incubation at 37°C (t = 24 h).

Strain	t=0h	t=24h	pH=24
DPC 6291	3.80±0.27	6.43±0.416	5.27±0.026
ATCC 51835	3.58±0.02	6.5±0.3	6.376±0.14
ATCC 13548	3.66±0.12	6.28±0.131	5.7±0.09
ATCC 13518	3.87±0.08	6.5±0.264	6.43±0.085
DPC 7170	3.22±0.65	6.12±0.155	5.5±0.1
DPC 7171	3.68±0.155	7.2±0.1	5.72±0.064
Control	ND	ND	6.8±0.05

Results are shown as mean values ± SD of triplicate experiments. ND=not detected.

Table 6: Flavour contributing Volatiles significantly assosicated with *M. caseolyticus* subsps *caseolyticus* strains (P< 0.05) from two trials.

Compound	Flavour description	DPC6291	DPC7170	DPC7171	ATCC 51835	ATCC13548	ATCC 13518	RI
acid								
Acetic acid	Vinegar, peppers, green, fruity floral, sour	+	+*	+	ND	+*	+	689
Butanoic acid	Sweaty, butter, cheese, strong, acid, rancid, dirty sock	+*	+	+	+	+*	+	866
aldehyde								
3-Methyl-butanal	Malty, dark chocolate, almond, cocoa, coffee	+	+	+	+	+	+*	693
Nonanal	Green, citrus, fatty, floral	+*	+	+	ND	ND	ND	1148
ketone								
2-Undecanone	Floral,fruity, green, musty, tallow	ND	ND	+	+*	+	+*	1330
ester								
Ethyl decanoate	"Queso de flor" ("FLOR" CHEESE)	ND	ND	ND	+*	ND	ND	1419
Methyl butanoate	Sweet, fruity	+	+	+*	ND	+	+	748
Butyl butanoate	Pineapple, banana, sweet	ND	+*	+*	ND	+*	ND	1019
Butyl hexanoate	Fruity, pineapple, waxy, green, juicy, apple	ND	+*	+*	ND	+*	ND	1215
Phenolic compounds								
p-Cresol	Cowy-barny	+*	ND	ND	ND	ND	+	1182
sulfur compound								
Methanethiol	Rotting cabbage, cheese, vegetative, sulphur	+*	ND	ND	ND	ND	ND	460

5.5 Discussion

Flavour is result of a combination of both taste and aroma, and the volatiles responsible for the typical flavour and aroma of fermented products are produced mainly by the metabolism of proteins, fats and carbohydrates. Of these, proteolysis is identified as particularly important for flavour development in fermented dairy products and the components of this pathway have been well defined in LAB (Liu et al., 2010). These dairy-associated microorganisms have been extensively used in food fermentations as flavour generators. However, currently, there is a drive to examine the metabolic diversity of strains that might not normally be associated with dairy products, as such strains may serve as a tools for the production of novel and distinct flavour profiles (McAuliffe et al., 2019). Therefore, in this study, we have explored the metabolic potential of *M. caseolyticus* subsp. *caseolyticus* to contribute to flavor formation. This species is found to be associated with certain fermented food types and has been documented to have a history of safe use and potential technological benefit according to the International Dairy Federation (IDF) Inventory of Microbial Food Cultures (Bourdichon et al., 2012), although information regarding its specific role in flavour generation is limited.

Prior to the commencement of this study, we had isolated a single strain of *M. caseolyticus* subsp. *caseolyticus* (DPC6291) from semi-hard cheese. Preliminary examination of the proteolytic capability of this strain was examined on RSM agar, and indicated rapid and extensive casein degradation after an overnight incubation. Similarly, *M. caseolyticus* subsp. *caseolyticus* ATCC13548 strain had been investigated previously by Bhowmik and Marth in 1988, which demonstrated complete and rapid degradation of β-casein and other fractions of casein, suggested the ability of this strain to generate flavour and aroma-forming compounds (Bhowmik and Marth, 1988). In an effort to examine the flavor forming potential of this subspecies, we established a strain bank of *M. caseolyticus* subsp. *caseolyticus* from a variety of sources as described previously (Mazhar et al., 2018), and performed a systems-wide analysis of the pathways potentially contributing to flavor formation in fermented foods.

The cell- enveloped proteinases, encoded by *prt*P and its homologs, plays an important role in flavour development, as its activity towards casein hydrolysis results in the

provision of substrates for the subsequent steps of the proteolytic cascade. Our genome-wide comparative analysis on the five *M. caseolyticus* subsp. *caseolyticus* out of the six strains from our strain bank revealed the presence of a single copy of *prt*P across all of the five strains examined (DPC6291, DPC7170, DPC7171, ATCC13548 and ATCC51835). However, the CEP activity of the six *M. caseolyticus* subsp. *caseolyticus* strains analysed in this study indicated strain-to strain variability with dairy-derived strains displaying extensive capability in hydrolyzing casein (DPC6291, DPC7170, DPC7171 and ATCC13548). This variability was also evident when RP-HPLC profiles were examined, which clearly demonstrated the extensive hydrolysis of casein fractions with high CEP-active, dairy-associated strains (DPC6291, ATCC13548, DPC7170 and DPC7171) and comparatively lower activity towards casein fractions were observed with a less CEP-active, non-dairy derived strain ATCC51835, and also with strain ATCC13518, the source of which is unknown. In addition, the high protease action of the dairy-derived strains was apparent on RSM plates and when these strains were inoculated in LFM milk leading to coagulation of the milk after an overnight incubation at pH values (between the range 5.15 – 5.72) above those defined for acid-induced coagulation (pH ~4.6)(Phadungath, 2005). This coagulation was absent in less CEP-active strains, which further highlights the extensive milk protein hydrolysis capabilities of the high CEP-active strains. Overall, the presence of *prt*P homologs across the dairy and the non-dairy associated genomes did not correlate with the phenotype. Further examination of the amino acid sequence encoded by the *prt*P gene indicated comparatively high sequence similarities amongst the dairy derived strains (> 99%) and relatively less sequence similarities between the dairy and the non-dairy derived strain ATCC51835 (~ 98%). These mutations in the *prt*P gene could be the cause of the variability obtained in the phenotypic expression of the proteinase. The genome analysis has also identified the presence of CLP ATPases proteases across the five genomes which are reported to be active towards caseins (Gottesman et al., 1990; Bertuzzi, 2017). Therefore, the significant protease activity observed in dairy derived strains could be either as a result of the action of caseolytic protease CEP (*prt*P) or a combined action with the CLP ATPases proteases.

The presence of peptide transporters oligo-peptide (OPP) across all five genomes suggests the ability for peptide uptake and their subsequent metabolism in these strains. However, other transport systems which are well described in LAB such as

Di/tripeptide (DtpT/DtpP) were not identified, implying a limitation in the uptake of nitrogen sources in the form of dipeptides or tripeptides and the subsequent utilization of such peptides by these organisms (Liu et al., 2010). The next step after oligo-peptide uptake in proteolysis is the internal hydrolysis of peptides to FAA by the action of peptidases. In this study, all six *M. caseolyticus* subsp. *caseolyticus* strains demonstrated limited activity towards general broad specificity and proline specific peptidases which correlates with the absence of genes encoding for these enzymes in the genome. All six strains were further analysed for their activity towards Glu and Ala-Phe-Pro-pNA substrates for PepA and PepV peptidases, which were identified in the genomes. However, all strains demonstrated limited activity towards these substrates indicating either these enzymes may not be efficient or the substrates used are not appropriate to evaluate their activity.

A previous investigation conducted by our group with a bank of *Lactobacillus casei* strains, established that ArAT was a suitable test case for the determination of general AT activity (Stefanovic et al., 2017). Therefore, in this study, we investigated the ArAT activities across the six strains as a representation of the general AT activity. The presence of number of AT genes in the genome correlated with the phenotypic data, which revealed that all tested strains demonstrated ArAT activity. However, the activity was comparatively lower than *L. paracasei* DPC4206 (positive control strain). GDH activity was detected in all strains with ATCC13548 as statistically significant; however the level of activity was lower in comparison to *L. paracasei* DPC4206 and DPC4536 and other strains of LAB reported in previous studies (Stefanovic et al., 2017), but corresponds to the activity reported for *Staphylococcus saprophyticus* strain DPC5671 and *M. caseolyticus* DPC6291 investigated by Bertuzzi, (2017).

Moreover, FAA and Phenotypic microarray PM3 analysis conducted further confirmed limited peptidyl hydrolases capability of the six strains. The evaluation of FAA release in fermentates of the six test strains indicated a significant release of only histidine, proline and tryptophan out of 20 amino acids analysed, in comparison to the control. From the PM3 analysis, it was revealed that strains were capable of metabolising only 13.6% of the total nitrogen sources tested, of which the majority were fatty acids (e-Amino-N-Caproic acid, D,L-a-Amino-Caprylic acid, D-Amino-N-Valeric acid) and di-peptides (Ala-His, , Met-Ala, Gly-Glu). The hydrolysis di-

peptides which may correlate with the presence of the PepV and PepA peptidases present in the genomes. The broad specificity dipeptidase activity of PepV may act on the Ala-His and Met-Ala substrates while the glutamyl aminopeptidase PepA may act on the Gly-Glu substrate (Christensen et al., 1999). In addition, in all six strains, L-cysteine was the most significantly metabolised nitrogen source, correlating with the presence of the *suf*S gene, encoding cysteine desulfurase, an enzyme that transforms L-cysteine to L-alanine and S-sulfanylcysteine, in all five genomes. Therefore, L-cysteine could represent one of the few preferred nitrogen sources required for the growth of this organism (Table: S1).

Altogether, our conclusion from our enzyme and genomic analysis is that the proteolytic system of *M. caseolyticus* subsp. *caseolyticus* differs considerably from those of the well described LAB species. Their ability to use the peptide substrates generated from casein is constrained by their limited peptidolytic activity. The absence of general peptidases and the weak activities of those present along with lower dehydrogenase (GDH) and aminotransferase (ArAT) activities required for the catabolism of amino acids suggests an inability of this organism to catabolize a wide array of amino acids and therefore, to produce significant quantities of amino acid-derived flavor compounds. The implications of this in a mixed strain culture system, such as in the manufacture of hard and semi-hard cheese, is that these enzymes are most likely provided by other strains in the mix, thus complementing the limited proteolytic activity of *M. caseolyticus* subsp. *caseolyticus*.

Comparative-genome analysis also indicated the presence of genes encoding lipase enzymes such as monoacylglycerol lipase (*mgl*P and *mgl*L) and carboxylesterase. The phenotypic data correlated with the genomic data as the lipase and esterolytic activity was observed in all strains. However, the esterase activity was significant in strain DPC6291 and DPC7170, and was higher than the *L. paracasei* DPC4206 strain. The esterase activity of DPC6291 on p-nitrophenol butyrate has been previously reported to be the most significant amongst the highly lipolytic strain *Y. lipolytica* DPC6266 by Bertuzzi, (2017)., and our results correlate with this study. Another study also identified the volatile compounds in cantonese sausage inoculated with a strain of *M. caseolyticus* subsp. *caseolyticus* to originate mainly from degradation and oxidation of lipids (Wu et al., 2009). The enzymatic analysis, together with genomic analysis

and the utilization of FFA in the PM3 nitrogen plates (capracylic, valeric and caproic acid), indicate the capability of this organism to metabolise FFA.

Finally, the volatile profiles of the six strains inoculated in LFM milk presented some associations with the enzymatic activities analysed. A majority of the compounds absent in the uninionculated LFM milk control were straight and branched chain esters, compound's considered to be metabolites of FFAs. Specifically, DPC6291, ATCC13548, DPC7170 and DPC7171 demonstrated significant esterase activity, whereas, all strains demonstrated lipolytic activity on tributyrin agar. Interestingly, the high CEP active strain DPC6291 was associated with the production of the phenol compound, p-cresol and the production of the sulphur compound methanethiol. The generation of p-cresol may be associated with the ability of DPC6291 to metabolise tryptophan as identified in FAA analysis, whereas the production of methanethiol could be associated with enzymes present in the genome such as PepM (methionine-specific activity) which functions to remove N-terminal methionine residues from proteins and cystathionine γ –lyase which catalyses the production of methanethiol from methionine or they could have been produced from the precursors or contaminants present in the milk. These compound have been listed as main odorant in number of cheeses such as British farmhouse cheddar, smear ripened and mould ripened cheeses (Molimard and Spinnler, 1996; McSweeney and Sousa, 2000; Curioni and Bosset, 2002). DPC6291 was also associated with the production of the aldehyde nonanal (derived from β-oxidation of FFAs), which is associated with green, citrus and fatty aroma identified in soft cheeses (Sablé and Cottenceau, 1999; Collins et al., 2003). ATCC13548, DPC7170 and DPC7171 were significantly associated with the production of esters. The key odorants produced by these strains included hexyl butanoate, isopentyl hexanoate, butyl hexanoate, butyl butanoate which are linked with sweet, fruity and floral notes (Curioni and Bosset, 2002). The majority of the compounds identified in the relatively low protease and esterase active strains, ATCC 51835 and ATCC 13518 were also present in the control but in significantly lower amounts. Therefore, in the hierarchical clustering, these strains cluster more closely to the control, ATCC13518 being the most similar. There was one compound identified as a potent odorant absent in control and significantly ($P< 0.05$) produced only by ATCC13518, 3-methyl-butanal (leucine transamination) associated with green, malty aroma. This compound could have originated as a result of methylenation of butanal

present in the control. 3-methyl-butanal is identified as a potent odorant in Camembert aged cheddar and in a number of other cheese varieties (Griffith and Hammond, 1989; Curioni and Bosset, 2002). In comparison to other strains, a limited number of volatiles were significantly associated with ATCC51835 (n=3). One of these is a ketone, 2-undecanone, identified as key aroma compounds in Camembert cheese, and reported to originate from fatty acid oxidation (Curioni and Bosset, 2002; Bertuzzi et al., 2018). Overall, the volatile analyses of *M. caseolyticus* subsp. *caseolyticus* strains confirms their metabolic diversity as they demonstrated different capacities for the production of flavour compounds correlating with the enzymatic analysis as majority of these compounds are mainly associated with FFA metabolism. Hierarchical clustering of the high CEP and esterase active strain demonstrates that these dairy derived strains are more metabolically active in LFM milk, sharing several similar volatile compounds as these strains cluster together and sperate from the control (uninionculated LFM milk), whereas the non-dairy derived strains with relatively low CEP and esterase activity sharing similar profile with the control and were also comparatively less metabolically active in LFM milk, therefore separated and clustered together.

5.6 Conclusion

This study demonstrates the genomic, phenotypic ability and the metabolic diversity of six *M. caseolyticus* subsp. *caseolyticus* strains for the production of flavour compounds. The genomic comparison revealed the components of the proteolytic and lipolytic system to be conserved. On the other hand, the observed variability in their activities especially in case of CEP and esterases activities indicates that these may be a consequence of different regulation and not due to the different number of key enzyme encoding homologs. The high CEP activity of *M. caseolyticus* subsp. *caseolyticus* strains resulting in extensive casein hydrolysis indicates the ability of these strains to generate high levels of substrates to feed the subsequent steps in the proteolytic cascade. However, limited peptidase activity was observed correlating with the absence of general peptidases in the genomes of these strains. Overall, *M. caseolyticus* subsp. *caseolyticus* demonstrated the ability to generate diverse volatiles with some potent odorants, which makes them potentially useful for further investigation as adjuncts. Work is ongoing to examine the potential of the *M. caseolyticus* subsp. *caseolyticus* strains in combination with high peptidolytic LAB strains, this synergistic effect on the metabolism of proteins may potentially enhance generation of flavour compounds originating from the proteolytic pathway. However, a note of caution should be advised as our genome analysis has also revealed the presence of multiple antibiotic resistance and virulence genes in members of this subspecies (unpublished data). While these specific genes may be considered niche adaptation factors, they could potentially contribute to these organisms being unsuitable for food applications and this information should be considered on a case-by-case in any future use in food systems.

5.7 Acknowledgements

Thanks to David Mannion for assistance with R-software for statistical analysis and to Bernard Corrigan for conducting RP-HPLC analysis. This research is funded by the Teagasc Walsh Fellowship programme (ref. 2015055).

5.8 References

Alkalaf, W., Vassal, L., Desmazeaud, M., and Gripon, J. (1987). Use of Rulactine as ripening agent in semi-hard cheese. *Lait (France)*.

Altermann, E. (2012). Tracing lifestyle adaptation in prokaryotic genomes. *Frontiers in Microbiology* 3**,** 48-48.

Bertuzzi, A. (2017). Modification of cheese flavour through the use of surface microbiota.

Bertuzzi, A.S., Walsh, A.M., Sheehan, J., Cotter, P.D., Crispie, F., Mcsweeney, P.L., Kilcawley, K.N., and Rea, M.C. (2018). Omics-based insights into flavor development and microbial succession within surface-ripened cheese. *MSystems* 3**,** e00211-00217.

Bhowmik, T., and Marth, E.H. (1988). Protease and peptidase activity of *Micrococcus* species. *Journal of Dairy Science* 71**,** 2358-2365.

Bhowmik, T., and Marth, E.H. (1990). Rote of *Micrococcus* and *Pediococcus* species in cheese ripening: A Review1. *Journal of Dairy Science* 73**,** 859-866.

Bochner, B.R., Gadzinski, P., and Panomitros, E. (2001). Phenotype microarrays for high-throughput phenotypic testing and assay of gene function. *Genome Research* 11**,** 1246-1255.

Bourdichon, F., Casaregola, S., Farrokh, C., Frisvad, J.C., Gerds, M.L., Hammes, W.P., Harnett, J., Huys, G., Laulund, S., and Ouwehand, A. (2012). Food fermentations: microorganisms with technological beneficial use. *International Journal of Food Microbiology* 154**,** 87-97.

Christensen, J.E., Dudley, E.G., Pederson, J.A., and Steele, J.L. (1999). Peptidases and amino acid catabolism in lactic acid bacteria. *Antonie van Leeuwenhoek* 76**,** 217-246.

Collins, Y.F., Mcsweeney, P.L., and Wilkinson, M.G. (2003). Lipolysis and free fatty acid catabolism in cheese: a review of current knowledge. *International Dairy Journal* 13**,** 841-866.

Curioni, P.M.G., and Bosset, J.O. (2002). Key odorants in various cheese types as determined by gas chromatography-olfactometry. *International Dairy Journal* 12**,** 959-984.

Desmazeaud, M., and Hermier, J. (Year). "Isolement, purification et propriétés d'une protéase exocellulaire de *Micrococcus caseolyticus*", in: *Annales de Biologie Animale Biochimie Biophysique*: EDP Sciences), 565-577.

Evans, A.C. (1916). The bacteria of milk freshly drawn from normal udders. *The Journal of Infectious Diseases***,** 437-476.

Gobbetti, M., De Angelis, M., Di Cagno, R., and Rizzello, C. (2007). "The relative contributions of starter cultures and non-starter bacteria to the flavour of cheese," in *Improving the Flavour of Cheese*. Elsevier), 121-156.

Gottesman, S., Clark, W.P., and Maurizi, M.R. (1990). The ATP-dependent Clp protease of *Escherichia coli*. Sequence of *clpA* and identification of a Clp-specific substrate. *Journal of Biological Chemistry* 265**,** 7886-7893.

Griffith, R., and Hammond, E.G. (1989). Generation of Swiss cheese flavor components by the reaction of amino acids with carbonyl compounds. *Journal of Dairy Science* 72**,** 604-613.

Hargrove, R.E., and Mcdonough, F.E. (1964). "Process of making low-fat ripened skim milk cheese, US3156568A". Google Patents).

Hiramatsu, K., Katayama, Y., Matsuo, M., Sasaki, T., Morimoto, Y., Sekiguchi, A., and Baba, T. (2014). Multi-drug-resistant *Staphylococcus aureus* and future chemotherapy. *Journal of Infection and Chemotherapy* 20**,** 593-601.

Jensen, M.P., and Ardö, Y. (2010). Variation in aminopeptidase and aminotransferase activities of six cheese related *Lactobacillus helveticus* strains. *International Dairy Journal* 20**,** 149-155.

Kieronczyk, A., Skeie, S., Langsrud, T., Le Bars, D., and Yvon, M. (2004). The nature of aroma compounds produced in a cheese model by glutamate dehydrogenase positive *Lactobacillus* INF15D depends on its relative aminotransferase activities towards the different amino acids. *International Dairy Journal* 14**,** 227-235.

Liu, M., Bayjanov, J.R., Renckens, B., Nauta, A., and Siezen, R.J. (2010). The proteolytic system of lactic acid bacteria revisited: a genomic comparison. *BMC Genomics* 11**,** 36.

Mašlaňová, I., Wertheimer, Z., Sedláček, I., Švec, P., Indráková, A., Kovařovic, V., Schumann, P., Spröer, C., Králová, S., Šedo, O., Krištofová, L., Vrbovská, V., Füzik, T., Petráš, P., Zdráhal, Z., Ružičková, V., Doškař, J., and Pantuček, R. (2018). Description and comparative genomics of *Macrococcus*

caseolyticus subsp. *hominis* subsp. nov., *Macrococcus goetzii* sp. nov., *Macrococcus epidermidis* sp. nov., and *Macrococcus bohemicus* sp. nov., novel macrococci from human clinical material with virulence potential and suspected uptake of foreign DNA by natural transformation. *Frontiers in Microbiology*, 9.

Mazhar, S., Altermann, E., Hill, C., and Mcauliffe, O. (2019). Draft genome sequences of *Macrococcus caseolyticus, Macrococcus canis, Macrococcus bohemicus,* and *Macrococcus goetzii*. *Microbial Resource Announcement* 8**,** e00343-00319.

Mazhar, S., Hill, C., and Mcauliffe, O. (2018). A rapid PCR-based method to discriminate *Macrococcus caseolyticus* and *Macrococcus canis* from closely-related *Staphylococcus* species based on the *ctaC* gene sequence. *Journal of Microbiological Methods* 152**,** 36-38.

Mcauliffe, O., Kilcawley, K., and Stefanovic, E. (2019). Symposium Review: Genomic Investigations of Flavor Formation by Dairy Microbiota. *Journal of Dairy Science* 102**,** 909-922.

Mcdermott, A., Visentin, G., De Marchi, M., Berry, D.P., Fenelon, M.A., O'connor, P.M., Kenny, O.A., and Mcparland, S. (2016). Prediction of individual milk proteins including free amino acids in bovine milk using mid-infrared spectroscopy and their correlations with milk processing characteristics. *Journal of Dairy Science* 99**,** 3171-3182.

Mcsweeney, P.L. (2017). "Biochemistry of cheese ripening: introduction and overview," in *Cheese: Chemistry, Physics and Microbiology, Third edition*. Elsevier), 379-387.

Mcsweeney, P.L., and Sousa, M.J. (2000). Biochemical pathways for the production of flavour compounds in cheeses during ripening: A review. *Le Lait* 80**,** 293-324.

Molimard, P., and Spinnler, H. (1996). Compounds involved in the flavor of surface mold-ripened cheeses: Origins and properties. *Journal of Dairy Science* 79**,** 169-184.

Moreno, V., and Kosikowski, F. (1973). Degradation of β-casein by micrococcal cell-free preparations. *Journal of Dairy Science* 56**,** 33-38.

Mounsey, J.S., and O'kennedy, B.T. (2009). Stability of β-lactoglobulin/micellar casein mixtures on heating in simulated milk ultrafiltrate at pH 6.0. *International Journal of Dairy Technology* 62**,** 493-499.

Nikolić, M., Tolinački, M., Fira, D., Golić, N., and Topisirović, L. (2009). Variation in specificity of the PrtP extracellular proteinases in *Lactococcus lactis* and *Lactobacillus paracasei* subsp. *paracasei*. *Folia Microbiologica* 54**,** 188-194.

Phadungath, C. (2005). The mechanism and properties of acid-coagulated milk gels. *Songklanakarin Journal of Science and Technology* 27**,** 433-448.

Sablé, S., and Cottenceau, G. (1999). Current knowledge of soft cheeses flavor and related compounds. *Journal of Agricultural and Food Chemistry* 47**,** 4825-4836.

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30**,** 2068-2069.

Smit, G., Smit, B.A., and Engels, W.J. (2005). Flavour formation by lactic acid bacteria and biochemical flavour profiling of cheese products. *FEMS Microbiology Reviews* 29**,** 591-610.

Stefanovic, E., Kilcawley, K.N., Rea, M.C., Fitzgerald, G.F., and Mcauliffe, O. (2017). Genetic, enzymatic and metabolite profiling of the *Lactobacillus casei* group reveals strain biodiversity and potential applications for flavour diversification. *Journal of Applied Microbiology* 122**,** 1245-1261.

Tanous, C., Kieronczyk, A., Helinck, S., Chambellon, E., and Yvon, M. (2002). "Glutamate dehydrogenase activity: a major criterion for the selection of flavour-producing lactic acid bacteria strains," in *Lactic Acid Bacteria: Genetics, Metabolism and Applications*. Springer), 271-278.

Van Den Dool, H., and Kratz, P.D. (1963). "A generalization of the retention index system including linear temperature programmed gas-liquid partition chromatography".).

Wang, Y., Coleman-Derr, D., Chen, G., and Gu, Y.Q. (2015). OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Research* 43**,** W78-84.

Wu, Y., Cui, C., Sun, W., Yang, B., and Zhao, M. (2009). Effects of *Staphylococcus condimenti* and *Micrococcus caseolyticus* on the volatile compounds of Cantonese sausage. *Journal of Food Process Engineering* 32**,** 844-854.

Table S1: Results of the least significant test (LSD) ($p<0.05$) performed after Analysis of Variance (ANOVA) for all enzymatic analyses of strains of *M. caseolyticus* in enzymatic assays (cell envelope proteinase (CEP), aminopeptidases PepN, PepC, PepX, PepA and di-peptidase PepV, aromatic aminotransferase (ArAT), glutamate dehydrogenase (GDH)), Free amino acid analysis and PM3 analysis.

CEP (Cell enveloped activities)

DPC/ATCC strains and controls	Fluorescence	Stat. different groups
Positive control Prot K (1 µg/mL)	1141.1	A
6291	682.11	B
13548	570.4	C
7170	442.5	D
7171	416.2	D
Wg2	220.36	E
13518	157.69	EF
51835	105.99	F

PepX (Aminopeptidase X) activities:

DPC strain or sample	nmol p-NA/(min*mg protein)	Stat. different groups
6291	-29.53	C
13548	7.43	B
7170	3.07	B
7171	3.14	B
4206	46.4	A
13518	3.9	B
51835	-30.13	C

PepN (Aminopeptidase N) activities:

DPC strain or sample	nmol p-NA/(min*mg protein)	Stat. different groups
6291	0.32	B
13548	6.71	B
7170	-1.96	B
7171	3.7	B
4206	58.75	A
13518	-22.4	B
51835	2.59	B

PepC (Aminopeptidase C) activities:

DPC strain or sample	nmol p-NA/(min*mg protein)	Stat. different groups
6291	5.147	B
13548	2.357	B
7170	1.26	B
7171	2.867	B
4206	48.23	A
13518	-0.15	B
51835	1.54	B

PepV (Dipeptidase pepV) activities

DPC strain or sample	nmol p-NA/(min*mg protein)	Stat. different groups
6291	4.296	BC
13548	1.965	BC
7170	8.53	B
7171	3.09	BC
4206	44.02	A
13518	-11.31	CD
51835	-18.93	D

PepA (Aminopeptidase A) activities:

DPC strain or sample	nmol p-NA/(min*mg protein)	Stat. different groups
6291	2	A
13548	5.43	A
7170	3.08	A
7171	3.96	A
4206	-21.66	B
13518	3.441	A
51835	0.59	A

ArAT (aromatic aminotransferase) activites:

DPC strain or sample	µmol Ph-pyruvate/mg protein	Stat. different groups
6291	2.07	B
13548	2.04	B
7170	1.93	B
7171	1.91	B
4206	4.073	A
13518	1.50	B
51835	1.61	B

GDH (glutamate dehydrogenase) activities

DPC strain or sample	U/mg protein	Stat. different groups
6291	4.8373	D
13548	7.17	C
7170	6.17	D
7171	4.299	D
4206	13.3604	B
13518	4.484	D
51835	4.29	D
4536	17.3936	A

Esterase (p-nitrophenol octanoate) activities

DPC strain or sample	p-nitrophenol/mg protein	Stat. different groups
6291	0.726	A
13548	0.6397	AB
7170	0.507889	AB
7171	0.3816	ABC
4206	0.416	A B C
13518	0.2945	BC
51835	0.338	ABC

Esterase (p-nitrophenol butyrate) activities

DPC strain or sample	p-nitrophenol/mg protein	Stat. different groups
6291	0.7988	A
13548	0.532	AB
7170	0.705	A
7171	0.5231	AB
4206	0.443	AB
13518	0.2686	AB
51835	0.337	AB

Free amino acid analysis (Ion Exchange column)

Amino acid:	**Cysteic acid**	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	9.55475	A
DPC7171	14.33875	A
DPC7170	17.53675	A
ATCC13518	36.97	A
ATCC13548	17.405	A
ATCC51835	12.3875	A
DPC6295	22.456	A
Amino acid:	**Taurine**	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	14.357	A
DPC7171	8.9795	A
DPC7170	16.122	A
ATCC13518	15.3305	A
ATCC13548	16.21725	A
ATCC51835	15.5885	A
DPC6295	11.49325	A
Amino acid:	**Aspartic acid**	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	2.82225	A
DPC7171	0.89275	A

DPC7170	1.01175	A
ATCC13518	1.82125	A
ATCC13548	1.2115	A
ATCC51835	1.411	A
DPC6295	0.434	A
Amino acid:	**Glycine**	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	6.68675	A
DPC7171	3.69	A
DPC7170	3.557	A
ATCC13518	0.96125	A
ATCC13548	3.5015	A
ATCC51835	0.0115	A
DPC6295	1.82525	A
Amino acid:	**Threonine**	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	1.7745	A
DPC7171	0.68675	A
DPC7170	0.6625	A
ATCC13518	6.492	A
ATCC13548	0.289	A
ATCC51835	0.196	A
DPC6295	0.5425	A
Amino acid:	**Serine**	

DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	1.7745	A
DPC7171	0.68675	A
DPC7170	0.6625	A
ATCC13518	6.492	A
ATCC13548	0.289	A
ATCC51835	0.196	A
DPC6295	0.5425	A
Amino acid:	**Phenylalanine**	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	1.53775	A
DPC7171	16.40175	A
DPC7170	11.15575	A
ATCC13518	10.5765	A
ATCC13548	11.13875	A
ATCC51835	0.71375	A
DPC6295	18.3515	A
Amino acid:	**Histidine**	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	4.108	B
DPC7171	14.3575	AB
DPC7170	9.46825	AB
ATCC13518	7.65425	AB

ATCC13548	9.318	AB
ATCC51835	3.16575	B
DPC6295	19.3655	A
Amino acid:	**Proline**	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	0.657	F
DPC7171	0.388	G
DPC7170	3.536	D
ATCC13518	3.444	E
ATCC13548	4.9415	C
ATCC51835	8.215	A
DPC6295	6.7195	B
Amino acid:	**Glutamic acid**	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	42.6195	A
DPC7171	46.999	A
DPC7170	47.81625	A
ATCC13518	25.83	A
ATCC13548	39.9245	A
ATCC51835	39.14375	A
DPC6295	49.39325	A

Amino acid:	**Alanine**	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	3.4705	A
DPC7171	13.39625	A
DPC7170	6.97925	A
ATCC13518	16.74775	A
ATCC13548	8.22175	A
ATCC51835	2.231	A
DPC6295	9.349	A
Amino acid:	**Cysteine**	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	5.07725	A
DPC7171	6.4275	A
DPC7170	6.07325	A
ATCC13518	5.53175	A
ATCC13548	6.98675	A
ATCC51835	4.89825	A
DPC6295	6.3425	A
Amino acid:	**Tryptophan**	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	1.80675	B
DPC7171	56.475	AB

DPC7170	58.232	AB
ATCC13518	1.22825	B
ATCC13548	66.0055	AB
ATCC51835	3.08775	B
DPC6295	103.2255	A
Amino acid:	**Leucine**	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	4.1415	A
DPC7171	35.95575	A
DPC7170	30.77825	A
ATCC13518	22.93525	A
ATCC13548	22.57125	A
ATCC51835	4.36325	A
DPC6295	48.256	A
Amino acid:	**Lysine**	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	6.848	A
DPC7171	7.494	A
DPC7170	3.152	A
ATCC13518	25.38	A
ATCC13548	9.523	A
ATCC51835	5.804	A

DPC6295	3.985	A
Amino acid:	GABA	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	0	A
DPC7171	0.163	A
DPC7170	0.05325	A
ATCC13518	66.40025	A
ATCC13548	0.03075	A
ATCC51835	0.03175	A
DPC6295	0.15175	A
Amino acid:	**Valine**	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	4.2715	A
DPC7171	19.17275	A
DPC7170	12.31175	A
ATCC13518	21.6525	A
ATCC13548	14.1495	A
ATCC51835	6.48675	A
DPC6295	16.9245	A
Amino acid:	**Methionine**	
DPC and ATCC strains	µg/ml	Stat. different groups

Control (Lactose free milk)	0.7115	A
DPC7171	1.464	A
DPC7170	0.62325	A
ATCC13518	3.02525	A
ATCC13548	0.924	A
ATCC51835	0.1665	A
DPC6295	0.93625	A
	AAcid	
Amino acid:	**Isoleucine**	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	2.01775	A
DPC7171	1.72175	A
DPC7170	1.28225	A
ATCC13518	6.3645	A
ATCC13548	1.77775	A
ATCC51835	1.79475	A
DPC6295	1.42	A
Amino acid:	**Tyrosine**	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	0.80575	A
DPC7171	5.55375	A
DPC7170	2.96625	A

ATCC13518	9.21125	A
ATCC13548	5.4825	A
ATCC51835	1.04375	A
DPC6295	5.50575	A
Amino acid:	**Arginine**	
DPC and ATCC strains	µg/ml	Stat. different groups
Control (Lactose free milk)	5.331	A
DPC7171	7.077	A
DPC7170	3.352	A
ATCC13518	12.85	A
ATCC13548	4.8	A
ATCC51835	2.668	A
DPC6295	4.739	A

Differences in nitrogen substrate utilisation using Area under the curve values tested with Analysis of Variance (ANOVA) for each well of the PM3 plate:

Well:	**A02**		**A03**		**A04**		**A05**	
Substrates:	**Ammonia**		**Nitrite**		**Nitrate**		**Urea**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	2979	A	1638.5	A	-975	A	-420.5	A
7171	5930	A	2438.5	A	-1172.5	A	-1090.5	A
13518	1860	A	445	A	-354.5	A	-82	A
13548	1002	A	-186.5	A	-2180	A	-1811.5	A
51835	2937.5	A	2591	A	-1028.5	A	640	A
6291	3549.5	A	1955	A	270	A	336.5	A
Well:	**A06**		**A07**		**A08**		**A09**	
Substrates:	**Biuret**		**L-Alanine**		**L-Arginine**		**L-Asparagine**	
DPC and ATCC strains		Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	-982.5	A	-4308	A	-1638.5	A	-2428	A
7171	-4337.5	A	-725	A	-1709.5	A	607	A
13518	-313.5	A	-5869	A	954	A	-25	A
13548	-1828	A	-1276	A	-835.5	A	1688.5	A
51835	-1557	A	-5086.5	A	2453	A	739	A
6291	-810.5	A	3248	A	1494.5	A	3207	A
Well:	**A10**		**A11**		**A12**		**B01**	

Substrates:	**L-Aspartic Acid**		**L-Cysteine**		**L-Glutamic Acid**		**L-Glutamine**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	-542.5	A	6850	A	2578	A	-1309	A
7171	-2454	A	6804.5	A	1877.5	A	-1088.5	A
13518	3435.5	A	7739.5	A	2601	A	-857.5	A
13548	2007.5	A	14768	A	4379.5	A	-404.5	A
51835	1901.5	A	5995	A	-719	A	745	A
6291	5126.5	A	14146	A	5866.5	A	5216	A
Well:	**B02**		**B03**		**B04**		**B05**	
Substrates:	**Glycine**		**L-Histidine**		**L-Isoleucine**		**L-Leucine**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	-6597	A	-6740	A	-4289	A	398.5	A
7171	-595.5	A	-943	A	-2921.5	A	2047.5	A
13518	-3975.5	A	-8919.5	A	-6256.5	A	-2367	A
13548	1732.5	A	-5237	A	-2917.5	A	-1021.5	A
51835	-4466	A	-1845	A	-9259	A	-4147	A
6291	-2227	A	-3717.5	A	-2559.5	A	-2458.5	A
Well:	**B06**		**B07**		**B08**		**B09**	
Substrates:	**L-Lysine**		**L-Methionine**		**L-Phenylalanine**		**L-Proline**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	-398.5	A	1065	A	-3180	A	-84.5	A

7171	718	A	4438.5	A	1692.5	A	1828	A
13518	-2199.5	A	-612.5	A	-2717	A	2951	A
13548	-4410.5	A	-1123	A	-3896	A	-491	A
51835	2235.5	A	-1059	A	-742.5	A	1547.5	A
6291	183	A	694	A	-5182	A	-51	A
Well:	**B10**		**B11**		**B12**		**C01**	
Substrates:	**L-Serine**		**L-Threonine**		**L-Tryptophan**		**L-Tyrosine**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	-2695.5	A	2761.5	A	-5456.5	A	8173.5	A
7171	782	A	3375	A	-3978	A	8945	A
13518	-2302	A	8743	A	-4763.5	A	7371	A
13548	-1656	A	14902	A	-328.5	A	3783.5	A
51835	-1718.5	A	-635	A	-7816.5	A	4028	A
6291	-3683.5	A	530.5	A	-4648.5	A	5650	A
Well:	**C02**		**C03**		**C04**		**C05**	
Substrates:	**L-Valine**		**D-Alanine**		**D-Asparagine**		**D-Aspartic Acid**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	-486.5	A	-4956	A	-3634	A	-3236	AB
7171	3319.5	A	-2336	A	-1230.5	A	-5689	B
13518	-1239.5	A	-5385.5	A	-3964.5	A	-330	AB
13548	-1668	A	-4333.5	A	-4536.5	A	-2826	AB
51835	-237.5	A	-2000	A	-824.5	A	323	AB

6291	-204.5	A	214	A	2474	A	8881.5	A
Well:	**C06**		**C07**		**C08**		**C09**	
Substrates:	**D-Glutamic Acid**		**D-Lysine**		**D-Serine**		**D-Valine**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	-24	A	-3235.5	A	-15189	AB	1161.5	A
7171	966	A	-3297	A	-21004.5	B	4046	A
13518	720	A	-1034.5	A	-13425.5	AB	1872.5	A
13548	-1650	A	-2983.5	A	-6489	AB	-1285.5	A
51835	1836	A	790.5	A	-6431.5	AB	1983	A
6291	9136.5	A	1025.5	A	-22.5	A	2342	A
Well:	**C10**		**C11**		**C12**		**D01**	
Substrates:	**L-Citrulline**		**L-Homoserine**		**L-Ornithine**		**N-Acetyl-D,L-Glutamic Acid**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	173	A	-938.5	A	-931.5	A	-3430	A
7171	3102	A	1851	A	-7334.5	A	-3332.5	A
13518	-406	A	6516	A	1002.5	A	-5732	A
13548	-2131.5	A	11177	A	662.5	A	-1804	A
51835	1056.5	A	-1165	A	1066.5	A	-7078.5	A
6291	1121.5	A	2847	A	1603	A	1088.5	A
Well:	**D02**		**D03**		**D04**		**D05**	

Substrates:	**N-Phthaloyl-L-Glutamic Acid**		**L-Pyroglutamic Acid**		**Hydroxylamine**		**Methylamine**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	-17053.5	A	-216	A	-5121.5	A	-1199.5	A
7171	-19181.5	A	9782	A	-6929.5	A	-2599	A
13518	-18934	A	-1527	A	-10190.5	A	-2010.5	A
13548	-7745	A	-2917.5	A	-4252.5	A	-3799.5	A
51835	-21842.5	A	1641.5	A	-6041	A	1136.5	A
6291	-9783.5	A	984.5	A	-1423	A	1634	A
Well:	**D06**		**D07**		**D08**		**D09**	
Substrates:	**N-Amylamine**		**N-Butylamine**		**Ethylamine**		**Ethanolamine**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	-713	A	78	A	-199.5	A	815	A
7171	-587.5	A	-696.5	A	-1154.5	A	4597.5	A
13518	-1825	A	176.5	A	331	A	-211	A
13548	-3100.5	A	-2953.5	A	-2582.5	A	-1540.5	A
51835	1709.5	A	2104	A	1873.5	A	1752.5	A
6291	632	A	1510.5	A	1036.5	A	1602	A
Well:	**D10**		**D11**		**D12**		**E01**	
Substrates:	**Ethylenediamine**		**Putrescine**		**Agmatine**		**Histamine**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups

7170	352	A	4170.5	A	1370	A	-1214	A
7171	2902.5	A	5813	A	1874	A	1655	A
13518	-987.5	A	3498	A	5160	A	-3229.5	A
13548	-121	A	2479	A	2296	A	-3018	A
51835	496.5	A	3667	A	13.5	A	2032	A
6291	7000	A	4529.5	A	-972.5	A	1740.5	A
Well:	**E02**		**E03**		**E04**		**E05**	
Substrates:	**b-Phenylethylamine**		**Tyramine**		**Acetamide**		**Formamide**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	-2502	A	-629	A	502.5	A	1019	A
7171	-2217	A	-1645.5	A	-646.5	A	-260	A
13518	-6171	A	-3979.5	A	-265	A	835.5	A
13548	-5622.5	A	-4942.5	A	-2654	A	-2159	A
51835	-346	A	857	A	2044.5	A	2513	A
6291	-1137	A	-262	A	2670.5	A	2699.5	A
Well:	**E06**		**E07**		**E08**		**E09**	
Substrates:	Glucuronamide		D,L-Lactamide		D-Glucosamine		D-Galactosamine	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	121.5	A	-286	A	-4338	A	1435.5	A
7171	214.5	A	-1022.5	A	-4797	A	3760.5	A
13518	28.5	A	-1285	A	-4352	A	377	A

13548	-3268.5	A	-2871.5	A	-4381	A	153	A
51835	2536.5	A	1875.5	A	-1623	A	2110.5	A
6291	1433	A	-210.5	A	-3747.5	A	1888	A
Well:	**E10**		**E11**		**E12**		**F01**	
Substrates:	**D-Mannosamine**		**N-Acetyl-D-Glucosamine**		**N-Acetyl-D-Galactosamine**		**N-Acetyl-D-Mannosamine**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	-1252	A	1402	A	1925.5	A	1099	A
7171	-2365	A	1448	A	600.5	A	-463	A
13518	-121.5	A	1623.5	A	1996	A	394.5	A
13548	1412	A	-434	A	1420.5	A	-3390.5	A
51835	153.5	A	2464	A	1632.5	A	1562.5	A
6291	1372.5	A	1629.5	A	3375.5	A	3061	A
Well:	**F02**		**F03**		**F04**		**F05**	
Substrates:	**Adenine**		**Adenosine**		**Cytidine**		**Cytosine**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	-10289	A	-10343	A	4776	A	-2441	A
7171	-7081.5	A	-9617.5	A	5120	A	-3246	A
13518	-12698.5	A	-15394	A	-1363.5	A	-3243.5	A
13548	-7816.5	A	-8280.5	A	-6054.5	A	-7886	A
51835	-11378	A	-17152	A	364	A	-76.5	A
6291	-8142	A	-7499	A	5771	A	-2584.5	A

Well:	**F06**		**F07**		**F08**		**F09**	
Substrates:	**Guanine**		**Guanosine**		**Thymine**		**Thymidine**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	-15573	A	-3916.5	A	-1534	A	-3395.5	A
7171	-14875	A	-5190.5	A	-1912	A	-9019	A
13518	-14139.5	A	-9348.5	A	-2042.5	A	-6609	A
13548	-6354.5	A	-6975.5	A	-6371.5	A	-4386	A
51835	-21901.5	A	-2825.5	A	-1288	A	-2453	A
6291	-7138	A	1251	A	-1051.5	A	3498.5	A
Well:	**F10**		**F11**		**F12**		**G01**	
Substrates:	**Uracil**		**Uridine**		**Inosine**		**Xanthine**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	2188	A	2134	A	771	A	4688.5	A
7171	810	A	2453	A	120	A	4216.5	A
13518	2021.5	A	-2012	A	-3715	A	2164	A
13548	166.5	A	-433.5	A	-368	A	5927.5	A
51835	1225	A	382.5	A	-1271	A	2797.5	A
6291	3052	A	3039.5	A	3321	A	5945	A
Well:	**G02**		**G03**		**G04**		**G05**	
Substrates:	**Xanthosine**		**Uric Acid**		**Alloxan**		**Allantoin**	

DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	-3584.5	A	-6491	A	-2463.5	A	91.5	A
7171	-4902.5	A	-8565.5	A	-4356	A	-1452	A
13518	-6671	A	-7033.5	A	-4813	A	-101	A
13548	-1456.5	A	-5423.5	A	5778	A	367.5	A
51835	-2725.5	A	-4348.5	A	-6159.5	A	161	A
6291	-3026.5	A	413.5	A	7883.5	A	2332.5	A
Well:	**G06**		**G07**		**G08**		**G09**	
Substrates:	**Parabanic Acid**		**D,L-a-Amino-N-Butyric Acid**		**g-Amino-N-Butyric Acid**		**e-Amino-N-Caproic Acid**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	-3510.5	A	-4000.5	A	-612	A	1750	A
7171	-7387	A	575.5	A	1250	A	872.5	A
13518	-4159	A	-4096.5	A	-1280	A	3259.5	A
13548	-3083	A	-3038	A	-2802	A	631.5	A
51835	-4000.5	A	-2240	A	450.5	A	2074	A
6291	-686.5	A	1105.5	A	833	A	2336.5	A
Well:	**G10**		**G11**		**G12**		**H01**	
Substrates:	**D,L-a-Amino-Caprylic Acid**		**d-Amino-N-Valeric Acid**		**a-Amino-N-Valeric Acid**		**Ala-Asp**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups

7170	17846.5	A	1677.5	A	-730	A	-2245	A
7171	16841	A	1561	A	-2987	A	-3574	A
13518	19034.5	A	2229.5	A	228	A	-2292.5	A
13548	29853	A	1223	A	2704	A	5666	A
51835	10269	A	2048.5	A	-1403	A	-2519	A
6291	27382	A	3093	A	3707.5	A	10663.5	A
Well:	H02		H03		H04		H05	
Substrates:	**Ala-Gln**		**Ala-Glu**		**Ala-Gly**		**Ala-His**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	-2789	A	-704.5	A	9923.5	A	744	A
7171	-38.5	A	-3775.5	A	9674.5	A	2810.5	A
13518	-4879	A	1680	A	5198	A	1025.5	A
13548	-1177.5	A	2412	A	9817.5	A	3612.5	A
51835	-1357.5	A	1461.5	A	-1679.5	A	1640.5	A
6291	989.5	A	9172	A	11185	A	1957.5	A
Well:	**H06**		**H07**		**H08**		**H09**	
Substrates:	Ala-Leu		Ala-Thr		Gly-Asn		Gly-Gln	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	-4795	A	8869.5	A	-12281.5	A	-376.5	A
7171	-1364	A	7135	A	-12220	A	3519	A
13518	-8629	A	9565	A	-12500.5	A	-4757	A
13548	-2624	A	17290.5	A	-5050	A	1639.5	A

51835	-4725	A	-1798.5	A	-6187	A	-4428	A
6291	-1770.5	A	13119	A	-4098	A	11358	A

Well:	**H10**		**H11**		**H12**	
Substrates:	**Gly-Glu**		**Gly-Met**		**Met-Ala**	
DPC and ATCC strains	AUC values	Stat. different groups	AUC values	Stat. different groups	AUC values	Stat. different groups
7170	7650	A	5704.5	A	4480.5	A
7171	7099.5	A	8550.5	A	5745.5	A
13518	11028	A	8485.5	A	6599	A
13548	14952	A	14067.5	A	8770.5	A
51835	1992	A	-527	A	608	A
6291	15401	A	1556.5	A	934.5	A





Mazhar, S. 2020. A multi-omics perspective on the biology and evolution of the genus Macrococcus. PhD Thesis, University College Cork.

Please note that Chapter 6(pp. 222-283) is unavailable due to a restriction requested by the author.

CORA Cork Open Research Archive http://cora.ucc.ie

Chapter 7
General Discussion

7.1 General Discussion

The genus *Macrococcus* is currently comprised of eleven validly published species, *Macrococcus bovicus, Macrococcus carouselicus, Macrococcus equipercicus, Macrococcus brunensis, Macrococcus hajekii, Macrococcus lamae, Macrococcus goetzii, Macrococcus bohemicus, Macrococcus epidermidis, Macrococcus canis* and *Macrococcus caseolyticus*. The species of *M. caseolyticus* is further divided into two subspecies: *M. caseolyticus* subsp. *hominis* and *M. caseolyticus* subsp. *caseolyticus* (Mašlaňová et al., 2018). Historically *M. caseolyticus* has undergone extensive taxonomic revisions and it was not until Kloos *et al* in 1998, proposed the separation of *M. caseolyticus* from the genus *Staphylococcus* to an independent, yet closely related genus *Macrococcus* (Kloos et al., 1998). The genus *Macrococcus* and *Staphylococcus* are monophyletic, sharing the same ancestor, and are estimated to have diverged from each other over 200 million years ago (Hiramatsu et al., 2014). In contrast to the staphylococcal species, earlier reports have defined the species of micrococci to be avirulent to their animal host (Baba et al., 2009). However, recent reports have revealed the pathogenic potential of a strain belonging to the *M. caseolyticus* species affiliated with high mortality rates in mice and broiler chickens (Li et al., 2018). Additionally, members of *M. caseolyticus* and *M. canis* have also been associated with canine infections and multiple antibiotic resistances (Cotting et al., 2017a; Schwendener and Perreten, 2018). The acquisition of novel methicillin resistance homolog *mec*D raises concerns with the potential risk of its transfer to the pathogenic strains of *Staphylococcus aureus.* The gene *mec*D confers resistance to all classes of β-lactams including cephalosporin's "last resort" β-lactam antibiotics that are used to treat the serious methicillin resistance *Staphylococcus aureus* (MRSA) infections (Schwendener and Perreten, 2018). While new trends suggest members of *M. caseolyticus* and *M. canis* are potential reservoirs of methicillin resistance genes which conflicts with the anecdotal evidence suggesting the potential application of *M. caseolyticus* species in the generation of flavour of fermented products. This led us to establish a strain bank collection of *M. caseolyticus* and *M. canis* from diverse.

Chapter 2 describes the limitations of the preliminary screening trials conducted utilising conventional approach based on a variety of phenotypic traits that could differentiate and isolate *M. caseolyticus* and *M. canis* from other closely related *Staphylococcus* species. At the time, there were no species-specific methods

developed to isolate and discriminate the targeted *Macrococcus* species. Therefore, the aim here was to develop a rapid PCR based screening method to distinguish a bank of novel *M. caseolyticus* and *M. canis* strains from a background of staphylococcal species. This method utilised primers targeting the partial amplification of a conserved region within the terminal enzyme of the electron transport chain, cytochrome C oxidase, that is absent in majority of *Staphylococcus* species. Three screening trials were conducted on raw milk and bovine tongue, and a total of five hundred potential isolates were examined with the designed cytochrome c oxidase subunit II (*ctaC*) PCR, from which only sixteen isolates were positively amplified. At the time, the 16 S rRNA gene sequencing identified all strains as members of the *M. caseolyticus* species. Further, pulse field gel electrophoresis (PFGE) then revealed thirteen novel *M. caseolyticus* strains isolated in this study with the proposed PCR-based method.

The results indicate the primers amplified species-specific sequences, allowing 100% successful discrimination of the targeted *Macrococcus* species from the strains of other *Macrococcus* and *Staphylococcus* species. However, a phylogenetic study based on comparative sequence analysis of the amplified *ctaC* gene indicated that the thirteen novel strains separated into four distinct groups, whereas, only strains in group 1 and 2 shared significant nucleotide sequence similarity with the type strain of *M. caseolyticus* and *M. canis*. This method together with 16 S rRNA gene sequencing could not accurately determine the taxonomic position of the strains representing group 3 and 4 and therefore, these strains were further investigated using whole genome sequencing (WGS). Overall, this method achieves rapid discrimination and isolation of the targeted species from a background of complex microbial communities. However, the strains isolated in this study require further investigation using more specific methods that could potentially allow for greater resolution with respect to discrimination between the targeted species and novel species identification.

The aim in Chapter 3 was to resolve the taxonomic placement of the DPC7161^{T} strain isolated in Chapter 2 with the support of a polyphasic approach based on the combination of genomic, phenotypic and chemotaxonomic characteristics. This concensus approach offers a more complete characterization of microorganisms than the previous traditional microbiological methods based solely on morphological, physiological and biochemical characteristics. Therefore, in Chapter 3 approaches based on WGS similarity or distance referred to as overall genome relatedness indices

(OGRI), were used to decipher the phylogenetic relationships between DPC7161^T and all of the eleven type strains of the currently know *Macrococcus* species. This included average nucleotide identity (ANI), digital DNA-DNA hybridisation (dDDH) and core genome based phylogenomic analysis. The ANI and dDDH values between DPC7161^T and the type strains of all validly published *Macrococcus* species were significantly lower than the cut off value proposed for species delineation, 95-96%, and 70%, respectively. Additionally, the phylogenetic analysis based on core genome and 16 S rRNA gene sequences confirms the independent position of DPC7161^T within the *Macrococcus* genus. Furthermore, extensive phenotypic characterisation using API systems (API 50CH, API ID 32 Staph, and API ZYM) and metabolic fingerprints captured with GEN III MicroPlate^{TM} lead to identifying key phenotypic characters that can be used to distinguish the proposed novel *Macrococcus* species from the closely related species of *M. caseolyticus* by the ability to ferment and produce acid from D-mannitol and D-sorbitol. However, as this study is based on a single strain representative of the proposed species, the phenotypic characters may not be an accurate representation of all of its members. Taking together the genomic and phenotypic data, distinct cellular fatty acid profile and peptidoglycan structure, this research proposes that DPC7161^T should be recognised as the type strains of the novel proposed species *Macrococcus linguae* sps. nov.

In Chapter 4 same approach was employed using OGRI indexing along with core genome-based phylogeny to investigate the taxonomic position of the remaining ten strains isolated in Chapter 2. The results from ANI and dDDH revealed these strains to be members of the *M. canis* and *M. goetzii* species. However, the dDDH values obtained between the ten strains and their respective species type strains *M. canis* KM54013^T and *M. goetzii* CCM4927^T were below the threshold proposed for delineating subspecies (79-80%) (Meier-Kolthoff et al., 2014). This suggested that *M. canis* and the *M. goetzii* species are potentially composed of two distinct subspecies. Further analysis with WGS-based phylogeny and comprehensive phenotypic investigations confirmed the division of the *M. canis* and *M. goetzii* species into two proposed subspecies; *M. canis* subsp. *canis* subsp. nov., *M. canis* subsp. *bovinus* subsp. nov., *M. goetzii* subsp. *goetzii* subsp. nov, and *M. goetzii* subsp. *corkensis* subsp. nov.

The previous taxonomic description of the *M. canis* species by Brawand *et al*., (2017) was based on strains isolated from canine skin samples including a few strains isolated from infectious sites. These strains may now represent the subspecies *M. canis* subsp. *canis* subsp. nov. Interestingly, most of the strains in this study displayed complete haemolysis of sheep blood and all were positive for the CAMP reaction, which was associated with the strains ability to produce delta haemolysin (Brawand et al., 2017). These were the first reported strains of the genus to have β-haemolytic phenotype, which were later associated with the genes encoding for the bicomponent gamma-hemolysin (*hlg*B and *hlg*C) present in the genome of the type strain KM54013^{T} (Brawand et al., 2018). Conversely, our seven strains representing *M. canis* subsp. *bovinus* subsp. nov., were sourced from bovine tongue samples and were not associated with infections. The strains in our study lack the ability to produce β-haemolysis or CAMP factor, with further WGS-based comparative analysis of our strains indicated absence of the genes encoding for the biocomponent of gamma-hemolysin (*hlg*B and *hlg*C) (see Chapter 6). This supports the rationale to separate the *M. canis* species into two distinct subspecies based on their haemolytic capability, which may serve future clinical importance.

The previous description of the *M. goetzii* species described by Mašlaňová *et al*., 2018, was based on a single strain CCM4927^{T}, sourced from human clinical material. The study employed a polyphasic taxonomic approach to correctly identify the taxonomic position of this type strain within the genus *Macrococcus*. Conversely, our study has aimed to identify the taxonomic placement of the three strains representing *M. goetzii* subsp. *corkensis* subsp. nov., sourced from bovine tongue through employing comparative genomics with a core genome-based phylogenetic approach on all the currently known species and subspecies of the *Macrococcus* genus. The use of genomic data to construct phylogenomic trees for bacterial species and subspecies classifications have been suggested in numerous studies (Chun et al., 2018; Chung et al., 2018). However, this approach has not been employed in any of the previous taxonomic related studies of macrococci. Our phylogenomic analysis has clearly identified the separation of the proposed members of the subspecies as well as providing a better resolution of the evolutionary relationships amongst all the currently known members of the *Macrococcus* genus.

Overall, our results have demonstrated the taxonomic separation of the *M. canis* and *M. goetzii* into two distinct groups based on the combination of both genomic and phenotypic data, including: (a) the dDDH values between the subspecies being higher than the species-level cut-off values, (b) the dDDH between subspecies were lower than the subspecies cut-off values, (c) the subspecies forming a separate branch in the core genome-based phylogenomic tree, and (d) phenotypic traits that can be used to distinguish the proposed novel subspecies from those described before.

In Chapter 5 we explored the flavour forming potential of *M. caseolyticus* subsp. *caseolyticus.* Historically, there were reports suggesting the proteolytic capability of the type strain ATCC13548^T towards β-casein, displaying extensive hydrolysis and suggesting that this strain is capable of producing substrates for flavour compound generation (Bhowmik and Marth, 1988). Another study reports the flavour forming potential of a single strain of this species isolated from Cantonese sausage. The majority of the volatiles generated in this study originated mainly from the degradation and oxidation of lipids (Wu et al., 2009). The development of flavour is a complex process which involves both proteolytic (degradation of proteins) and lipolytic (degradation of lipids) pathways (McSweeney, 2017) and to-date, there has been no study that has performed a thorough analysis of the key enzymes involved in these two pathways in *M. caseolyticus* subsp. *caseolyticus*. Therefore, we investigate the enzymatic activities of the proteolytic and lipolytic cascade, including cell-envelope proteinase (CEP), peptidases, esterases, lipases, aminotransferases (ATs) and glutamate-dehydrogenase (GDH) of six strains, isolated from dairy and non-dairy associated niches through employing chromogenic and agar-based assays. The dairy associated strains displayed high CEP and esterase activities in comparison to the non-dairy derived strains. Thus, this variability between strains is likely a result of niche adaptation by the dairy-associated *M. caseolyticus* subsp. *caseolyticus* strains. Interestingly, genome-comparative analysis of the genes encoding proteolytic and lipolytic enzymes were conserved across the dairy and non-dairy derived strains, including the CEP encoded by the *prt*P gene. Further, comparative multiple sequence alignment of the *prt*P gene indicated the presence of several amino acids that differed from the consensus sequence in the non- dairy derived strain. This could explain the differences obtained in the phenotypic expression of the protease between strains. However, limited downstream peptidolytic activity and the generation of free amino

acids (FAA) were obtained across both, dairy and non-dairy associated strains. In this study, the limited release of FAA correlated with the lack of peptidase activities and the absence of general peptidases in their genome (PepN and PepC), indicating the organism has limited potential to generate FAA-derived flavours.

In contrast to the proteolytic capability, esterolytic activities of *M. caseolyticus* subsp. *caseolyticus* were higher than the *Lactobacilus paracasei* DPC4206 strains (investigated in this study) and previously reported lipolytic strain of *Yarrowia lipolytica* DPC6266 (Bertuzzi, 2017). This phenotypic data correlated with the volatile profiles generated by these strains inoculated in lactose free milk, as majority of the compounds produced were esters. The findings presented in this study have confirmed the metabolic ability of *M. caseolyticus* subsp. *caseolyticus* to generate flavour compounds and has identified highly active caseolytic proteases in dairy-associated strains of *M. caseolyticus* subsp. *caseolyticus* driving the extensive casein hydrolysis. This high protease action merits further investigation of these strains in a combined system which should include high peptidolytic strains to examine the synergistic effects on the metabolism of the proteins.

Further comparative genome analysis of *M. caseolyticus* subsp. *caseolyticus* in Chapter 6 has revealed the presence of multiple antimicrobial resistance and virulence factors in some members of this species. This evidence seems at odds with the documented safe and technological beneficial use of this organism in food published by the International Dairy Federation (IDF) (Bourdichon et al., 2012). The emerging data of the pathogenic potential of *M. caseolyticus* associated with several virulence factors and resistance genes associated with mobile genetic elements (MGE's) suggests that strains from this organism are unsuitable for food application (Cotting et al., 2017b; Li et al., 2018). This statement is in accordance with the European Food Safety Authority (EFSA) guidance on acquired resistance in microbes and its unsuitability to food applications (Additives and Feed, 2012). Therefore, these organisms should not be intentionally added to food as its addition to food may then lead to transmitting these genes through the food chain to other commensal organisms and that will further the spread to pathogens. However, our research may be of industrial benefit as the high protease activity associated with dairy-derived strains of *M. caseolyticus* subsp. *caseolyticus* can be exploited in the future for flavour diversification and potentially be of a use in accelerating cheese ripening.

In Chapter 6 we investigate the genomic diversity within the genus *Macrococcus,* through conducting comparative pan-gnome analysis on thirty-two strains representing all of its members. The pan-core genome analysis revealed a very low conserved genomic core (25%) and a high percentage of accessory genome (75%). In addition, resistome analysis identified a total of eight antimicrobial resistance (AMR) genes associated with MGEs present on genomic islands (GIE) that have been acquired *via* horizontal gene transfer from close and phylogenetically distant relatives. To-date *mec*D gene conferring resistance to all classes of β-lactams, including cephalosporins, was solely identified in *M. caseolyticus* species. Our results have identified the presence of *mec*D in another member of *Macrococcus*; *M. bohemicus* DPC7215 integrated on a chromosomal island sharing 100% amino acid identity with the *mec*D island described previously in *M. caseolyticus* subsp. *caseolyticus* IMD8019. Additionally, the presence of the *Staphylococcus aureus* plasmid pS0385-2, carrying streptomycin resistance was also identified in *M. bohemicus* DPC7215. Together, this suggests that horizontal gene transfer (HGT) is a key driver in the spread of AMR resistance across the members of the *Macrococcus* genus. The phenotypic expression of antimicrobial resistance in subsets of strains correlated with the resistome analysis. The results from these analyses lead to the identification of a potential novel kanamycin resistance gene in *M. linguae* DPC7161^{T}, which requires further characterisation using molecular cloning approach to confirm its inducible gene expression. The genus-wide distribution of virulence factors, CRISPR-Cas, secondary metabolites and genomic islands (GIEs) were also investigated in this study. Recent reports have identified members of *M. caseolyticus* and *M. canis* harbouring several virulence factors and have been associated in veterinary infections (Cotting et al., 2017b; Li et al., 2018). Our results have also identified a wide distribution of virulence factors genes across the macrococcal genomes. However, to understand and determine the pathogenic potential of these organisms further *in vitro* characterisation is required. Together, the results from this study highlight the high intra-genus diversity associated with high genetic exchange, indicating the importance of HGT as a major driver in the genomic evolution of the genus *Macrococcus*. Further analysis using a functional genomics approach to the genus-wide comparative genomic analysis presented in this study is required to gain a better understanding of the role of the dispensable genome in niche adaptation, pathogenicity and evolution in *Macrococcus*.

In conclusion, the thesis presented a systems wide approach integrating the enzymatic, metabolic and genomic data to unravel the flavour forming potential of *M. caseolyticus* subsp. *caseolyticus* strains. This work has identified a highly active membrane bound protease in dairy-associated *M. caseolyticus* subsp. *caseolyticus* deriving extensive casein hydrolysis, indicating the ability of these strains to generate more substrates to feed the preceding steps in the proteolytic cascade. However, limited peptidase activity correlating with the absence of general peptidases in the genome indicate the potential of these dairy-derived strains to be further studied in combination with high peptidolytic strains which may potentially lead to enhance generation of flavoursome compounds originating from the proteolytic pathway. Additionally, comprehensive genomic and phenotypic comparative analysis has led to the identification of novel species and subspecies of the genus *Macrococcus*. Moreover, genus-wide comparative genomic coupled with phylogenetic and metabolic analysis has revealed the extensive diversity within the genus *Macrococcus*. The evidence of high genomic exchange with the genus-wide distribution of genomic islands suggests horizontal gene transfer to play a key role in the genomic evolution of the genus *Macrococcus*.

7.2 References

(2019). International Code of Nomenclature of Prokaryotes. *International Journal of Systematic and Evolutionary Microbiology* 69**,** S1-S111.

Additives, E.P.O., and Feed, P.O.S.U.I.A. (2012). Guidance on the assessment of bacterial susceptibility to antimicrobials of human and veterinary importance. *EFSA Journal* 10**,** 2740.

Baba, T., Kuwahara-Arai, K., Uchiyama, I., Takeuchi, F., Ito, T., and Hiramatsu, K. (2009). Complete genome sequence of *Macrococcus caseolyticus* strain JSCS5402, reflecting the ancestral genome of the human-pathogenic staphylococci. *Journal of Bacteriology* 191**,** 1180-1190.

Bertuzzi, A. (2017). Modification of cheese flavour through the use of surface microbiota.

Bhowmik, T., and Marth, E.H. (1988). Protease and peptidase activity of *Micrococcus* species. *Journal of Dairy Science* 71**,** 2358-2365.

Bourdichon, F., Casaregola, S., Farrokh, C., Frisvad, J.C., Gerds, M.L., Hammes, W.P., Harnett, J., Huys, G., Laulund, S., and Ouwehand, A. (2012). Food fermentations: microorganisms with technological beneficial use. *International Journal of Food Microbiology* 154**,** 87-97.

Brawand, S.G., Cotting, K., Gómez-Sanz, E., Collaud, A., Thomann, A., Brodard, I., Rodriguez-Campos, S., Strauss, C., and Perreten, V. (2017). *Macrococcus canis* sp. nov., a skin bacterium associated with infections in dogs. *International Journal of Systematic and Evolutionary Microbiology* 67**,** 621-626.

Brawand, S.G., Rychener, L., Schwendener, S., Pantůček, R., and Perreten, V. (2018). Complete Genome Sequence of the Type Strain of *Macrococcus canis*. *Genome Announcements* 6**,** e01507-01517.

Chun, J., Oren, A., Ventosa, A., Christensen, H., Arahal, D.R., Da Costa, M.S., Rooney, A.P., Yi, H., Xu, X.W., De Meyer, S., and Trujillo, M.E. (2018). Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *Int J Syst Evol Microbiol* 68**,** 461-466.

Chung, M., Munro, J.B., Tettelin, H., and Hotopp, J.C.D. (2018). Using core genome alignments to assign bacterial species. *MSystems* 3**,** e00236-00218.

Cotting, K., Strauss, C., Rodriguez-Campos, S., Rostaher, A., Fischer, N.M., Roosje, P.J., Favrot, C., and Perreten, V. (2017a). *Macrococcus canis* and *M. caseolyticus* in dogs: occurrence, genetic diversity and antibiotic resistance. *Veterinary Dermatology*.

Cotting, K., Strauss, C., Rodriguez-Campos, S., Rostaher, A., Fischer, N.M., Roosje, P.J., Favrot, C., and Perreten, V. (2017b). *Macrococcus canis* and *M. caseolyticus* in Dogs: Occurrence, Genetic Diversity and Antibiotic Resistance. *Veterinary Dermatology* 28**,** 559-e133.

Hiramatsu, K., Katayama, Y., Matsuo, M., Sasaki, T., Morimoto, Y., Sekiguchi, A., and Baba, T. (2014). Multi-drug-resistant *Staphylococcus aureus* and future chemotherapy. *Journal of Infection and Chemotherapy* 20**,** 593-601.

Kloos, W.E., Ballard, D.N., George, C.G., Webster, J.A., Hubner, R.J., Ludwig, W., Schleifer, K.H., Fiedler, F., and Schubert, K. (1998). Delimiting the genus *Staphylococcus* through description of *Macrococcus caseolyticus* gen. nov., comb. nov. and *Macrococcus equipercicus* sp. nov., and *Macrococcus bovicus* sp. no. and *Macrococcus carouselicus* sp. nov. *International Journal of Systematic and Evolutionary Microbiology* 48 Pt 3**,** 859-877.

Li, G., Du, X., Zhou, D., Li, C., Huang, L., Zheng, Q., and Cheng, Z. (2018). Emergence of pathogenic and multiple-antibiotic-resistant *Macrococcus caseolyticus* in commercial broiler chickens. *Transboundary and Emerging Diseases* 65**,** 1605-1614.

Mašlaňová, I., Wertheimer, Z., Sedláček, I., Švec, P., Indráková, A., Kovařovic, V., Schumann, P., Spröer, C., Králová, S., Šedo, O., Krištofová, L., Vrbovská, V., Füzik, T., Petráš, P., Zdráhal, Z., Ružičková, V., Doškař, J., and Pantuček, R. (2018). Description and comparative genomics of *Macrococcus caseolyticus* subsp. *hominis* subsp. nov., *Macrococcus goetzii* sp. nov., *Macrococcus epidermidis* sp. nov., and *Macrococcus bohemicus* sp. nov., novel macrococci from human clinical material with virulence potential and suspected uptake of foreign DNA by natural transformation. *Frontiers in Microbiology*, 9.

Mcsweeney, P.L. (2017). "Biochemistry of cheese ripening: introduction and overview," in *Cheese: Chemistry, Physics and Microbiology, Third edition*. Elsevier), 379-387.

Meier-Kolthoff, J.P., Hahnke, R.L., Petersen, J., Scheuner, C., Michael, V., Fiebig, A., Rohde, C., Rohde, M., Fartmann, B., Goodwin, L.A., Chertkov, O., Reddy, T.B.K., Pati, A., Ivanova, N.N., Markowitz, V., Kyrpides, N.C., Woyke, T., Göker, M., and Klenk, H.-P. (2014). Complete genome sequence of DSM 30083T, the type strain (U5/41T) of *Escherichia coli*, and a proposal for delineating subspecies in microbial taxonomy. *Standards in Genomic Sciences* 9**,** 2.

Schwendener, S., and Perreten, V. (2018). The integrase of the *Macrococcus caseolyticus* resistance island *mecD* (McRImecD) inserts DNA site-specifically into *Staphylococcus* and *Bacillus* chromosomes. *Molecular microbiology* 110**,** 455-468.

Wu, Y., Cui, C., Sun, W., Yang, B., and Zhao, M. (2009). Effects of *Staphylococcus condimenti* and *Micrococcus caseolyticus* on the volatile compounds of Cantonese sausage. *Journal of Food Process Engineering* 32**,** 844-854.

Appendix 1

Draft Genome Sequences of the Type Strains of Six *Macrococcus* Species

All of this appendix has been published in Microbiology Resource Announcements (2019). 8(19).

DOI: 10.1128/MRA.00344-19

8.1 Abstract

We report here the draft genome sequences of *Macrococcus bovicus* ATCC51825^{T}, *Macrococcus carouselicus* ATCC51828^{T}, *Macrococcus equipercicus* ATCC51831^{T}, *Macrococcus brunensis* CCM4811^{T}, *Macrococcus hajekii* CCM4809^{T}and *Macrococcus lamae* CCM4815^{T}. The availability of the genome sequences of these species will enable cross –species comparison which could lead to a more comprehensive understanding of organisms of the *Macrococcus* genus.

8.2 Genome Announcement

The Gram-positive genus *Macrococcus* contains a total of 11 species - *Macrococcus bovicus, Macrococcus carouselicus, Macrococcus equipercicus, Macrococcus brunensis, Macrococcus hajekii, Macrococcus lamae, Macrococcus goetzii, Macrococcus epidermidis, Macrococcus bohemicus, Macrococcus caseolyticus* and *Macrococcus canis* (Brawand et al., 2018; Mašlaňová et al., 2018). These species are disseminated in nature as animal commensals and are indicated to be the immediate antecedent of the *Staphylococcus* species (Hiramatsu et al., 2014). While staphylococci are widespread as human pathogens, macrococci are defined to be avirulent (Baba et al., 2009). However, recent publications have indicated possible pathogenic potential of *Macrococcus* strains isolated from human clinical samples (Mašlaňová et al., 2018). The draft genome sequence reported here are of the type strains of six *Macrococcus* species that were isolated from artiodactyl and perissodactyl hosts (Kloos et al., 1998; Mannerová et al., 2003).

Genomic DNA was isolated from overnight cultures grown at 37°C in tryptic soy broth (TSB; Becton, Dickinson and Company, Berkshire, England) using the UltraClean microbial DNA isolation kit (Mo Bio Laboratories, Cambridge, United Kingdom) per the included protocol. Genomic libraries were prepared with the Nextera XT DNA library preparation kit (Illumina, Inc., San Diego, CA, USA). The 2 × 250 bp paired-end read sequencing was performed on an Illumina HiSeq 2500 platform (MicrobesNG, University of Birmingham, UK). Reads were adapter trimmed using Trimmomatic 0.30 with a sliding window quality cutoff of Q15 (Bolger et al., 2014). De novo assembly was performed on each sample using SPAdes version 3.7 using the program's default parameters (Bankevich et al., 2012). The genome sequences were annotated with the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (Haft et al., 2017). *In silico* analysis of acquired antimicrobial resistance genes and virulence genes was conducted using ResFinder version 3.4, VirulenceFinder version 2.0, PathogenFinder version 1.1, and the Virulence Factor Database (VFDB)(Chen et al., 2005; Lindsey et al., 2016). CheckM was used to calculate the completeness (all illustrated 99.6% completeness) and purity of the reads (ranged from 1.1- 2.04 % impurity) (Parks et al., 2015).

In all sequenced genomes, capsule (cap)-associated genes that are involved in phagocytosis evasion were identified. These genes involved in capsule biosynthesis illustrated DNA sequence identities in the range of 70 to 92% to those present in pathogenic strains of Staphylococcus aureus. Other putative virulence factors found were hemolysin III (*hly-III*), aureolysin (*aur*), and fibronectin-binding protein A (*fbpA*), among others. The sequencing and assembly statistics of the draft genome sequences are shown in Table 1. The genome sequences of these species could facilitate a better understanding of the biology of these organisms and a comprehensive understanding of the genus *Macrococcus*, which in turn could contribute to a greater understanding of antibiotic resistance acquisition and the pathogenic potential of the genus *Staphylococcus*.

Data availability. Genome sequence data of the six Macrococcus species were deposited into NCBI GenBank and the Sequence Read Archive (SRA) under BioProject number PRJNA515496. The accession numbers are listed in Table 1.

8.3 Acknowledgements

This research was funded by Teagasc (reference 6697) and the Teagasc Walsh Fellowship Programme (reference 2015055)

8.4 References

Baba, T., Kuwahara-Arai, K., Uchiyama, I., Takeuchi, F., Ito, T., and Hiramatsu, K. (2009). Complete genome sequence of *Macrococcus caseolyticus* strain JSCS5402, reflecting the ancestral genome of the human-pathogenic staphylococci. *Journal of Bacteriology* 191**,** 1180-1190.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., and Prjibelski, A.D. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology* 19**,** 455-477.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30**,** 2114-2120.

Brawand, S.G., Rychener, L., Schwendener, S., Pantůček, R., and Perreten, V. (2018). Complete Genome Sequence of the Type Strain of *Macrococcus canis*. *Genome Announcements* 6**,** e01507-01517.

Chen, L., Yang, J., Yu, J., Yao, Z., Sun, L., Shen, Y., and Jin, Q. (2005). VFDB: a reference database for bacterial virulence factors. *Nucleic acids research* 33**,** D325-D328.

Haft, D.H., Dicuccio, M., Badretdin, A., Brover, V., Chetvernin, V., O'neill, K., Li, W., Chitsaz, F., Derbyshire, M.K., and Gonzales, N.R. (2017). RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic acids research* 46**,** D851-D860.

Hiramatsu, K., Katayama, Y., Matsuo, M., Sasaki, T., Morimoto, Y., Sekiguchi, A., and Baba, T. (2014). Multi-drug-resistant *Staphylococcus aureus* and future chemotherapy. *Journal of Infection and Chemotherapy* 20**,** 593-601.

Kloos, W.E., Ballard, D.N., George, C.G., Webster, J.A., Hubner, R.J., Ludwig, W., Schleifer, K.H., Fiedler, F., and Schubert, K. (1998). Delimiting the genus *Staphylococcus* through description of *Macrococcus caseolyticus* gen. nov., comb. nov. and *Macrococcus equipercicus* sp. nov., and *Macrococcus bovicus* sp. no. and *Macrococcus carouselicus* sp. nov. *International Journal of Systematic and Evolutionary Microbiology* 48 Pt 3**,** 859-877.

Lindsey, R.L., Pouseele, H., Chen, J.C., Strockbine, N.A., and Carleton, H.A. (2016). Implementation of whole genome sequencing (WGS) for

identification and characterization of Shiga toxin-producing *Escherichia coli* (STEC) in the United States. *Frontiers in Microbiology* 7**,** 766.

Mannerová, S., Pantůček, R., Doškař, J., Švec, P., Snauwaert, C., Vancanneyt, M., Swings, J., and Sedláček, I. (2003). *Macrococcus brunensis* sp. nov., *Macrococcus hajekii* sp. nov. and *Macrococcus lamae* sp. nov., from the skin of llamas. *International Journal of Systematic and Evolutionary Microbiology* 53**,** 1647-1654.

Mašlaňová, I., Wertheimer, Z., Sedláček, I., Švec, P., Indráková, A., Kovařovic, V., Schumann, P., Spröer, C., Králová, S., Šedo, O., Krištofová, L., Vrbovská, V., Füzik, T., Petráš, P., Zdráhal, Z., Ružičková, V., Doškař, J., and Pantuček, R. (2018). Description and comparative genomics of *Macrococcus caseolyticus* subsp. *hominis* subsp. nov., *Macrococcus goetzii* sp. nov., *Macrococcus epidermidis* sp. nov., and *Macrococcus bohemicus* sp. nov., novel macrococci from human clinical material with virulence potential and suspected uptake of foreign DNA by natural transformation. *Frontiers in Microbiology*, 9.

Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome research* 25**,** 1043-1055.

TABLE 1 Genome characteristics of the six type strains of the *Macrococcus* genus

Strain	GenBank accession no.	SRA accession no.	Draft genome size (bp)	G+C content (%)	No. of contigs	Coverage (×)	N_{50} (bp)	No. of CDS(a)
M. bovicus ATCC51825^T	SDQM00000000	SRR8448136	2,087,234	44.25	46	216	146,194	2,191
M. hajekii CCM4809^T	SDQI000000000	SRR8448137	2,052,566	40.05	61	88	1,051,403	2,195
M. carouselicus ATCC51828^T	SDGN00000000	SRR8448134	2,069,165	43.83	34	82	1,168,712	2,153
M. equipercicus ATCC51831^T	SDQF00000000	SRR8448135	2,154,579	43.58	205	131	1,168,712	2,326
M. lamae CCM4815^T	SDQG00000000	SRR8448132	2,031,524	40.28	73	156	102,971	2,095
M. brunensis CCM4811^T	SDQH00000000	SRR8448133	2,089,39	41.59	83	122	999,97	2,184

(a)CDS, coding sequences.

Appendix 2

Draft Genome Sequences of *Macrococcus caseolyticus*, *Macrococcus canis*, *Macrococcus bohemicus*, and *Macrococcus goetzii*

All of this appendix has been published in Microbiology Resource Announcements (2019).8(19).

DOI: 10.1128/MRA.00343-19

9.1Abstract

Here, we present the draft genome sequences of 14 strains of 4 species of the genus *Macrococcus*. These strains were isolated from bovine milk and tongue samples obtained during a screening program.

9.2 Genome Announcement

Fourteen strains belonging to four members of the *Macrococcus* genus, namely, 3 *Macrococcus caseolyticus* strains (DPC6291, DPC7170, and DPC7171), 7 *Macrococcus canis* strains (DPC7158, DPC7160, DPC7162, DPC7163, DPC7165, DPC7168, and DPC7169), 3 *Macrococcus goetzii strains* (DPC7159, DPC7164, and DPC7166), and 1 *Macrococcus bohemicus* strain (DPC7215), were isolated from bovine milk and tongue by utilizing a ctaC PCR, as described previously (Mazhar et al., 2018). Recently emerging information regarding multidrug resistance and putative virulence genes present in species belonging to this genus prompted us to perform whole-genome sequencing (WGS) to investigate the presence of such genes in these *Macrococcus* strains (Wu et al., 2009; Wang et al., 2012; Schwendener et al., 2017).

The genomic DNA was isolated from overnight cultures grown at 37°C in tryptic soy broth (TSB; Becton, Dickinson and Company, Berkshire, England) using the UltraClean microbial DNA isolation kit (Mo Bio Laboratories, Cambridge, UK) as per the included protocol. Genomic libraries were prepared with a Nextera XT DNA library preparation kit (Illumina, Inc., San Diego, CA, USA). The 2 250-bp paired read sequencing was performed on an Illumina HiSeq 2500 platform (MicrobesNG, University of Birmingham, UK). Reads were adapter trimmed using Trimmomatic version 0.30, with a sliding window quality cutoff of Q15 (Bolger et al., 2014). *De novo* assembly was performed on each sample using SPAdes version 3.7 with the program's default parameters (Bankevich et al., 2012). Detection of acquired antimicrobial resistance genes in the assembled genomes was analyzed using ResFinder version 3.4 and Resistance Gene Identifier (RGI) version 4.2.2 to search against the Comprehensive Antibiotic Resistance Database (CARD). Virulence genes were identified using VirulenceFinder version 2.0, PathogenFinder version 1.1, and the Virulence Factors Database (VFDB) (Chen et al., 2005; Zankari et al., 2012; Jia et al., 2016; Lindsey et al., 2016). The genome sequences were annotated using the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) (Haft et al., 2017). The final draft genomes were estimated using CheckM to be 96% complete with 2.5% contamination (Parks et al., 2015). All sequenced genomes illustrated the presence of putative virulence factors, namely, hemolysin III (hlyIII), aureolysin (aur), and capsule (cap) genes. An RGI search of the homology models in CARD identified a total of 86 different antibiotic resistance genes, most of which are predicted to confer resistance

to fluoroquinolone (n = 19), macrolides (n = 26), and tetracycline (n = 24). The sequencing and assembly statistics of the draft genome sequences of the above-mentioned *Macrococcus* strains are shownin Table 1. The sequencing data contribute to the pool of available *Macrococcus* genomes and enable further generation of information regarding the presence of antibiotic resistance determinants and other virulence factors present in *Macrococcus* species.

Data availability. The draft WGS data were deposited into NCBI GenBank and the Sequence Read Archive (SRA) under the BioProject no. PRJNA515496. The accession numbers are listed in Table 1.

9.3 Acknowledgments

This research was funded by Teagasc (reference no. 6697) and the Teagasc Walsh Fellowship Programme (reference no. 2015055).

9.4 References

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., and Prjibelski, A.D. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of computational biology* 19**,** 455-477.

Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30**,** 2114-2120.

Chen, L., Yang, J., Yu, J., Yao, Z., Sun, L., Shen, Y., and Jin, Q. (2005). VFDB: a reference database for bacterial virulence factors. *Nucleic acids research* 33**,** D325-D328.

Haft, D.H., Dicuccio, M., Badretdin, A., Brover, V., Chetvernin, V., O'neill, K., Li, W., Chitsaz, F., Derbyshire, M.K., and Gonzales, N.R. (2017). RefSeq: an update on prokaryotic genome annotation and curation. *Nucleic acids research* 46**,** D851-D860.

Jia, B., Raphenya, A.R., Alcock, B., Waglechner, N., Guo, P., Tsang, K.K., Lago, B.A., Dave, B.M., Pereira, S., and Sharma, A.N. (2016). CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic acids research***,** gkw1004.

Lindsey, R.L., Pouseele, H., Chen, J.C., Strockbine, N.A., and Carleton, H.A. (2016). Implementation of whole genome sequencing (WGS) for identification and characterization of Shiga toxin-producing *Escherichia coli* (STEC) in the United States. *Frontiers in Microbiology* 7**,** 766.

Mazhar, S., Hill, C., and Mcauliffe, O. (2018). A rapid PCR-based method to discriminate *Macrococcus caseolyticus* and *Macrococcus canis* from closely-related *Staphylococcus* species based on the *ctaC* gene sequence. *Journal of Microbiological Methods* 152**,** 36-38.

Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015). CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome research* 25**,** 1043-1055.

Schwendener, S., Cotting, K., and Perreten, V. (2017). Novel methicillin resistance gene *mecD* in clinical *Macrococcus caseolyticus* strains from bovine and canine sources. *Scientific Reports* 7**,** 43797.

Wang, Y., Wang, Y., Schwarz, S., Shen, Z., Zhou, N., Lin, J., Wu, C., and Shen, J. (2012). Detection of the staphylococcal multiresistance gene *cfr* in *Macrococcus caseolyticus* and *Jeotgalicoccus pinnipedialis*. *Journal of antimicrobial chemotherapy* 67**,** 1824-1827.

Wu, Y., Cui, C., Sun, W., Yang, B., and Zhao, M. (2009). Effects of Staphylococcus condimenti and Micrococcus caseolyticus on the volatile compounds of Cantonese sausage. *Journal of food process engineering* 32**,** 844-854.

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., Aarestrup, F.M., and Larsen, M.V. (2012). Identification of acquired antimicrobial resistance genes. *Journal of antimicrobial chemotherapy* 67**,** 2640-2644.

TABLE 1 Genome characteristics of the *Macrococcus* strains used in this study

Organisim	**SRA accession no.**	**GenBank accession no.**	**Draft genome size (bp)**	**G+C content (%)**	**No. of contigs**	**Coverage (×)**	**N_{50} (bp)**
M. caseolyticus DPC6291	SRR8868656	SDQM00000000	2,171,480	36.68	74	70	229,924
M. canis DPC7158	SRR8868660	SDQI000000000	2,179,466	36.75	69	197	578,934
M. goetzii DPC7159	SRR8868665	SDGN00000000	2,530,812	34.06	93	184	275,573
M. canis DPC7160	SRR8868666	SDQF00000000	2,148,516	36.58	37	136	413,516
M. canis DPC7162	SRR8868667	SDQG00000000	2,139,904	36.62	44	107	353,259
M. canis DPC7163	SRR8868668	SDQH00000000	2,167,812	36.63	79	146	417,178
M. goetzii DPC7164	SRR8868659	SDGO00000000	2,563,253	34.07	61	137	458,326
M. canis DPC7165	SRR8868658	SDGP00000000	2,165,327	36.68	72	157	1,280,134
M. goetzii DPC7166	SRR8868662	SDGQ00000000	2,591,067	34.16	95	201	466,093
M. canis DPC7168	SRR8868661	SDGR00000000	2,134,151	36.68	41	95	397,880
M. canis DPC7169	SRR8868664	SDGS00000000	2,160,199	36.56	89	264	1,113,524
M. caseolyticus DPC7170	SRR8868655	SDQK00000000	2,106,646	36.77	67	48	147,285
M. caseolyticus DPC7171	SRR8868657	SDQJ00000000	2,110,528	36.77	99	231	108,839
M. bohemicus DPC7215	SRR8868663	SELR00000000	2,555,877	33.98	55	160	234,144