The roles of *Staphylococcus* and *Corynebacterium* in human malodour formation

Reyme Herman

Doctor of Philosophy

University of York
Department of Biology

September 2025

Abstract

The generation of human body odour is attributed to bacteria of the underarm/axillary microbiome. One of the most distinctive malodours is contributed by thioalcohols, particularly 3-methyl-3-sulfanylhexan-1-ol (3M3SH). Generation of this malodour is thought to involve two specific intracellular processes carried out by cutaneous bacteria. Recent studies have associated this process to a clade (closely related species) of staphylococci, including *Staphylococcus hominis*.

Prior to this research, *S. hominis* proteins responsible for transport of the malodour precursor Cys-Gly-3M3SH and release of 3M3SH from the intermediate Cys-3M3SH had been well characterised. A body of work within this thesis completes the characterisation of the *S. hominis* malodour generation pathway through the identification and characterisation of PepV, a promiscuous dipeptidase capable of converting the imported precursor Cys-Gly-3M3SH to Cys-3M3SH.

The genus *Corynebacterium* was initially associated with malodour formation but due to the lack of evidence for transport of Cys-Gly-3M3SH, research shifted to *Staphylococcus*. Existing phenotypic studies on malodour generation involve representative isolates of common cutaneous species derived from various sources. Here, the limitations of prior studies are addressed through the construction of the York Skin Microbiome (YSM) collection of isolates matched to specific individuals, some of whom were strong malodour producers. The work discussed in Chapters 3 and 4 employs advanced genetic and bioinformatics techniques to assess the diversity of YSM isolates of *Staphylococcus* and *Corynebacterium* respectively.

The YSM isolates then allowed the revisit of the hypothesis that *Corynebacterium* contributes to thioalcohol-based malodour. To test this, genetically diverse YSM isolates of *Staphylococcus* and *Corynebacterium* from a high malodour individual were analysed using malodour formation assays. Work in Chapter 5 revealed roles of *Corynebacterium* in malodour formation either individually or cooperatively with other members of the same microbiome. This project demonstrates the complexities within human microbiomes that can only be captured with a deep characterisation of the microbial community.

Author's Declaration

I, Reyme Herman, declare that this thesis is a presentation of original work and I am the sole author. This work has not been presented for a degree or other qualification at this University or elsewhere. All sources are acknowledged as references.

Chapters 2 – 5 (Results) are presented in a journal-style format where there has been a collaboration in the research conceptualisation, the work and drafting the manuscripts. As per the requirements of the journals these manuscripts have been or will be submitted to, each manuscript and corresponding chapter includes an "Author Contribution" section using the CRediT author statement framework (https://credit.niso.org/).

Chapters 2 – 4 have either been published, accepted or are under review. Chapter 5 is written in a manuscript format and is ready for submission.

Chapter 2:

Title: Identification of a staphylococcal dipeptidase involved in the production of human body odour

Reyme Herman, Bethan Kinniment-Williams, Michelle Rudden, Alexander Gordon James, Anthony J. Wilkinson, Barry Murphy, Gavin H. Thomas

Journal of Biological Chemistry

DOI: 10.1016/j.jbc.2024.107928
Status: Published on 23 October 2024

Chapter 3:

Title: High-resolution *Staphylococcus* profiling reveals intra-species diversity in a single skin niche

Reyme Herman, Daniel T. West, Sean Meaden, Imran Khan, Robert Cornmell, Joanne Hunt, Michelle Rudden, Holly N. Wilkinson, Matthew J. Hardman, Anthony J. Wilkinson, Barry Murphy, Gavin H. Thomas

Microbial Genomics

Preprint DOI: 10.1101/2025.04.01.646391
Status: Manuscript under review with Microbial Genomics. Reviewer's comments have been addressed and incorporated.

Post-viva status: Manuscript published with Microbial Genomics
DOI: 10.1099/mgen.0.001531

Chapter 4:

Title: Revealing the diversity of commensal corynebacteria from a single human skin site

Reyme Herman, Sean Meaden, Michelle Rudden, Robert Cornmell, Holly N. Wilkinson, Matthew J. Hardman, Anthony J. Wilkinson, Barry Murphy, Gavin H. Thomas

mSystems

Preprint DOI: 10.1101/2024.11.28.625817v3
Status: Manuscript under review with mSystems. Reviewer's comments have been addressed and incorporated.
Post-viva status: Manuscript published with mSystems
DOI: 10.1128/msystems.00459-25

Chapter 5:

Title: Malodour formation can be driven by cooperation between *Staphylococcus* sp. and *Corynebacterium* sp. within the same community

Reyme Herman, Lauren Cozens, Robert Cornmell, Anthony J. Wilkinson, Barry Murphy, Gavin H. Thomas

Status: Manuscript prepared with the intention for submission.

Acknowledgements

Firstly, I would like to thank Prof. Gavin Thomas (UoY Biology), Prof. Tony Wilkinson (UoY Chemistry) and Dr Alexander Gordon James (Unilever, retired) for giving me the opportunity to carry out this PhD but also guiding me to develop my skills as a more independent researcher. I extend my thanks to my TAP member Prof. Marjan van der Woude for the support during my PhD and also for continuously being my mentor, especially in the last few years when I began exploring academic routes post-PhD.

I would also like to thank Dr Barry Murphy who did an amazing job taking over from Gordy as my industrial supervisor at the end of 2022. The support I received from everyone at Unilever R&D including Dr. Rob Cornmell not only facilitated many aspects of the PhD project but also made it a very enjoyable experience.

I am who I am today because of the people around me. From my supervisors, to my mentors, my colleagues, my family and my friends, I cannot thank everyone enough.

Finally, I would like to give a special mention to my kids Layla and Rayan. They are my little curious scientists and I hope they stay as curious throughout their lives. They may not pick the academic path but with this, they have their names in at least one thesis like mummy and daddy!

Table of Contents

Abstract 2
Author's Declaration 3
Acknowledgements 6
List of Tables 13
List of Figures ..... **Error! Bookmark not defined.**

Chapter 1:

Thesis introduction 20
1.1 The human skin – structure and composition 21
1.1.1 Epidermis 21
1.1.2 Dermis 22
1.2 The human skin – Glands and associated secretions. 22
1.2.1 The sebaceous gland 22
1.2.2 The eccrine gland 24
1.2.3 The apocrine gland 24
1.3 Human malodour 25
1.3.1 Malodour - Steroids 25
1.3.2 Malodour – Volatile fatty acids (VFAs) 25
1.3.3 Malodours - Thioalcohol-based malodour 28
1.4 The axillary microbiome 32
1.5 Aims of project 33
1.6 Thesis introduction references 35

Chapter 2:

Characterisation of the Staphylococcal Cys-Gly-3M3SH dipeptidase PepV 40
2.1 Manuscript information 41
2.2 Context within thesis 42
2.3 Abstract 43
2.4 Introduction 43

2.5 Results 45
2.5.1 Biochemical identification of a dipeptidase with activity against Cys-Gly-3M3SH 45
2.5.2 ShPepV is a general dipeptidase active on diverse L-amino acid–containing dipeptides 47
2.5.3 ShPepV and SaPepV are efficient Cys-Gly-3M3SH dipeptidases 49
2.5.4 ShPepV is general dipeptidase that can accommodate Cys-Gly-3M3SH 52
2.5.5 Mutagenesis of the ShPepV binding pocket reveals key Cys-Gly-3M3SH accommodating residues 54
2.6 Discussion 58
2.7 Experimental procedures 62
2.7.1 Bacterial strains and plasmids 62
2.7.2 Isolation and identification of ShPepV 62
2.7.3 Cloning, expression, and purification of ShPepV, SaPepV, and CgTpdA 63
2.7.4 Mutagenesis of ShPepV 64
2.7.5 Coupled Cys-Gly-3M3SH DTNB assay 64
2.7.6 Dipeptide activity assayed by TLC 65
2.7.7 Metal cofactor test 65
2.7.8 Kinetics of ShPepV with Cys-Gly 65
2.7.9 Homology modelling of ShPepV and rigid-body refinement modelling of Cys-Gly-3M3SH 66
2.8 Data availability 67
2.9 Conflict of interest 67
2.10 Acknowledgements 67
2.11 Author contributions 67
2.12 Supporting information 69
2.13 References 87

Chapter 3:

Staphylococcal diversity within the axilla of a single individual 92

3.1 Manuscript information 93

3.2 Context within thesis (Abstract) 94

3.3 Abstract 95

3.4 Impact statement 95

3.5 Introduction 96

3.6 Results 97

3.6.1 Isolation and genetic characterisation of axillary *Staphylococcus* 97

3.6.2 Species- and isolate-specific genetic traits reveal potential roles and functional networks within a single environment 100

3.7 Discussion 107

3.8 Materials and methods 108

3.8.1 Sampling procedures 108

3.8.2 Bacterial culturing 108

3.8.3 DNA extraction 109

3.8.4 DNA library preparation and whole genome sequencing 109

3.8.5 Bioinformatics tools 109

3.9 Ethics statement 110

3.10 Data availability 110

3.11 Conflict of interest 110

3.12 Funding 111

3.13 Author contributions 111

3.14 Supporting data 112

3.15 References 126

Chapter 4:

Corynebacterial diversity within the human axilla 130

4.1 Manuscript information 131

4.2 Context within thesis (Abstract) 132
4.3 Abstract 133
4.4 Importance 134
4.5 Introduction 134
4.6 Results 136
4.6.1 High-resolution genomic characterisation reveals corynebacterial diversity in the human axilla 136
4.6.2 Pangenome and metabolic profiling reveal species- and strain-level functional diversity among cutaneous corynebacteria 140
4.6.3 Diverse biosynthetic capacities and antimicrobial resistance in corynebacterial skin isolates 144
4.6.4 Diverse prophage repertoires and conserved phage defence systems in skin-associated *Corynebacterium* isolates 147
4.7 Discussion 150
4.8 Materials and methods 153
4.8.1 Sampling procedures 153
4.8.2 Bacterial culturing 154
4.8.3 DNA extraction 154
4.8.4 DNA library preparation and whole genome sequencing 155
4.8.5 Bioinformatics tools 155
4.8.6 Antibiotic susceptibility testing 156
4.9 Ethics statement 156
4.10 Data availability 157
4.11 Conflict of interest 157
4.12 Funding 157
4.13 Author contributions 157
4.14 Supporting data 158
4.15 References 185

Chapter 5:

The combined roles of staphylococci and corynebacteria in thioalcohol malodour formation 195

5.1 Manuscript information 196

5.2 Context within thesis (Abstract) 197

5.3 Abstract 198

5.4 Introduction 199

5.5 Results 201

5.5.1 Axillary *S. hominis* and *C. amycolatum* isolates are capable of biotransforming both Cys-Gly-3M3SH and Cys-3M3SH 201

5.5.2 Cross-genera cooperation between *Staphylococcus* and *Corynebacterium* in converting Cys-Gly-3M3SH to 3M3SH 203

5.5.3 Pairs of corynebacteria are able to work in tandem to completely biotransform Cys-Gly-3M3SH 206

5.6 Discussion 208

5.7 Materials and methods 211

5.7.1 Bacterial strains and growth conditions 211

5.7.2 Generation of the 3M3SH standard curve 211

5.7.3 Single-organism biotransformation assay 211

5.7.4 Tandem biotransformation assay 212

5.8 Ethics statement 213

5.9 Conflict of interest 213

5.10 Funding 213

5.11 Author contributions 213

5.12 Supporting data 214

5.13 References 221

Chapter 6:

Conclusions and future directions 223

6.1 The potential roles of skin *Corynebacterium* in human malodour formation 224

6.2 Interactions between members of the skin microbiome community. .. 225

6.3 The axillary microbiome and beyond ..... **Error! Bookmark not defined.**

6.4 Future directions 229

6.4.1 Assessing the roles of axillary *Corynebacterium* in other high-odour volunteers 229

6.4.2 Identifying and characterising microbial components that could contribute to the alternative pathways to malodour formation 229

6.4.3 High-resolution genetics to understand the networks within a functional microbiome 229

6.5 Final reflections 231

6.6 Bibliography 232

List of Abbreviations 234

List of Tables

CHAPTER 1: Thesis Introduction

Table 1: Major odour molecules identified to date 27

CHAPTER 2: Characterisation of the Staphylococcal Cys-Gly-3M3SH dipeptidase PepV

Table 1. Kinetics of ShPepV, SaPepV and CgTpdA with Cys-Gly-3M3SH and ShPepV with Cys-Gly 51
Table 2. Kinetics of ShPepV mutants with Cys-Gly-3M3SH determined using the coupled DTNB assay 55
Supplementary Table 1. Bacterial strains used in this study. 82
Supplementary Table 2. Plasmids used in this study. 83
Supplementary Table 3. Primers used in this study 84
Supplementary Table 4. List of dipeptides and controls used for the characterisation of the peptidase activity of ShPepV 85
Supplementary Table 5. Kinetics of ShPepV mutants with Cys-Gly determined using a Cd-ninhydrin method 86

CHAPTER 3: Staphylococcal diversity within the axilla of a single individual

Supplementary Table 1. Multilocus sequence typing (MLST) allelic profiles of representative *S. epidermidis* isolates 118
Supplementary Table 2. Multilocus sequence typing (MLST) allelic profiles of representative *S. hominis* isolates 119
Supplementary Table 3: Non-core gene clusters from isolate cluster SE1 120
Supplementary Table 4: Non-core gene clusters from isolate cluster SE2 121
Supplementary Table 5: Non-core gene clusters from isolate cluster SE3 122
Supplementary Table 6: Non-core gene clusters from isolate cluster SE4 123
Supplementary Table 7: Non-core gene clusters from isolate cluster S 124

Supplementary Table 8. Antimicrobial resistance cassettes identified in all representative isolates using ResFinder 125

CHAPTER 4: Corynebacterial diversity within the human axilla

Table 1. The closest phages and bacteria genomes were identified using a BLASTn analysis on the putative prophages identified by geNomad 148
Supplementary Table 1. Speciating representative genomes using GTDB-tk 118
Supplementary Table 2. Summary of *Corynebacterium* isolates from volunteer A1 119
Supplementary Table 3: Summary of *Corynebacterium* isolates from volunteer A3 120
Supplementary Table 4: Summary of *Corynebacterium* isolates from volunteer A4 121
Supplementary Table 5: Summary of *Corynebacterium* isolates from volunteer A5 122
Supplementary Table 6: Resfinder analysis of the representative isolates 123
Supplementary Table 7: Antibiotic susceptibility disc diffusion assay of the representative isolates 124

CHAPTER 5: The combined roles of staphylococci and corynebacteria in thioalcohol malodour formation

Supplementary Table 1. List of isolates used in this study. 220

List of Figures

CHAPTER 1: Thesis Introduction

Figure 1. The structure of human skin 23
Figure 2. The known pathways to malodour production 30

CHAPTER 2: Characterisation of the Staphylococcal Cys-Gly-3M3SH dipeptidase PepV

Figure 1. Isolation and identification of the dipeptidase ShPepV, a protein capable of Cys-Gly-3M3SH dipeptidase activity in *S. hominis* 46
Figure 2. TLC-based assay of ShPepV activity against a subset of dipeptides tested 48
Figure 3. The dipeptidase activity of ShPepV against Cys-Gly-3M3SH in the presence of various divalent cations reveals a preference for Mn2+ as the metal cofactor 50
Figure 4. The model of ShPepV suggests the canonical PepV fold with a catalytic and lid domain that is observed in the experimentally derived structure of LdPepV 53
Figure 5. Plot of k_{cat} against K_M of ShPepV mutants with the wildtype protein 56
Figure 6. Uptake and enzymatic processing of Cys-Gly-3M3SH in *S. hominis* 59
Supplementary Data S1: The coupled Cys-Gly-3M3SH DTNB assay 69
Supplementary Data S2. Anion exchange chromatography of S. hominis lysates with the Capto Q resin 70
Supplementary Data S3. Hydrophobic interactions chromatography of S. hominis lysates with the Phenyl HP resin 71
Supplementary Data S4. Anion exchange chromatography of S. hominis lysates with the Capto DEAE resin 72
Supplementary Data S5. Size exclusion chromatography of S. hominis lysates with the Sephacryl S200 resin 73
Supplementary Data S6. Cys-Gly-3M3SH dipeptidase activity is detected in SEC elution fractions 12-14 74
Supplementary Data S7. Identification of PepV by peptide mass fingerprinting using trypsin digest and MALDI-MS/MS 75

Supplementary Data S9. TLC assay ShPepV against a range of dipeptides (1) 76
Supplementary Data S10. TLC assay ShPepV against a range of dipeptides (2) .. 77
Supplementary Data S11. Michaelis-Menten kinetics of ShPepV with Cys-Gly, Cys-Gly-3M3SH and SaPepV and CgTpdA with Cys-Gly-3M3SH 78
Supplementary Data S12. Overlay of the binding sites of ShPepV with Cys-Gly-3M3SH and LdPepV with an Asp-Ala phosphinic acid inhibitor 79
Supplementary Data S13. Michaelis-Menten kinetics of ShPepV mutants I413W, M439A, M271R, E175A, D437A and A174V with Cys-Gly-3M3SH 80
Supplementary Data S14. Michaelis-Menten kinetics of ShPepV mutants I413W, M439A, M271R, E175A, D437A and A174V with Cys-Gly 81

CHAPTER 3: Staphylococcal diversity within the axilla of a single individual

Figure 1. Diversity of 120 axillary staphylococci isolates 98
Figure 2. Pangenome analysis of the 7 dereplicated axillary staphylococci isolates 101
Figure 3. COG categories were assigned to singletons from *S. epidermidis* and *S. hominis* isolates using eggNOG-mapper 102
Figure 4. Metabolic pathways in genetically distinct axillary staphylococci reveals species and isolate level variation within the isolates 104
Figure 5. Phage defence systems were predicted in the 7 dereplicated genomes using padloc..... 106
Supplementary Fig 1. Phylogenetic analysis of *S. epidermidis* MLST sequence types using the defined allelic profiles on PubMLST 112
Supplementary Fig 2. Phylogenetic analysis of *S. hominis* MLST sequence types using the defined allelic profiles on PubMLST 113
Supplementary Fig 3. Full 16S rRNA alignment of YSMAA1_1_H3st and *S. capitis* DSM 6717 (type strain). 114
Supplementary Fig 4. KEGG modules shown in Fig 4 are identified. 115
Supplementary Fig 5. KEGG modules shown in Fig 4 are identified 116

CHAPTER 4: Corynebacterial diversity within the human axilla

Figure 1. Capturing the diversity of axillary corynebacteria. 137
Figure 2. The pangenome of the axillary corynebacteria isolates 141
Figure 3. Metabolic pathways in axillary corynebacteria reveals species specific general profiles and some variation within species 143
Figure 4. Multiple putative biosynthetic gene cluster (BGC) types were identified in the representative isolate genomes using antiSMASH 7 145
Figure 5. Various phage defence systems were predicted in the representative genomes using PADLOC 149
Supplementary Figure 1. Pairwise average nucleotide identities (ANI) of all isolates sequenced in this study 158
Supplementary Figure 2. Clustering of corynebacterial isolates using drep 159
Supplementary Figure 4. Alignments of representatives from the two novel species *C. axilliensis* and *C. jamesii* against closest hits from GenBank. 161
Supplementary Figure 5. A core genome alignment of the representative genomes and select RefSeq genomes 162
Supplementary Figure 6. COG classification of *C. axilliensis* genes 163
Supplementary Figure 7. COG classification of genes of the *C. jamesii* YSMAA1_1_F7 164
Supplementary Figure 8. Comparisons of the core genomes of species *C. marquesiae*, *C. tuberculostearicum* and *C. kefirresidentii* derived from the respective isolate genomes 165
Supplementary Figure 9. COG classification of genes specific to *C. kefirresidentii*, *C. tuberculostearicum* and *C. marquesiae* 166
Supplementary Figure 10. Metabolic pathways in axillary corynebacteria of volunteer A1 167
Supplementary Figure 11. Architectures of the 8 distinct putative biosynthetic gene clusters predicted using antiSMASH 7 (1) 168
Supplementary Figure 12. Architectures of the 8 distinct putative biosynthetic gene clusters predicted using antiSMASH 7 (2) 169
Supplementary Figure 13. The distribution of putative biosynthetic gene clusters in the *C. tuberculostearicum* species complex predicted using antiSMASH 7 170

Supplementary Figure 14. The distribution of putative biosynthetic gene clusters in the *C. amycolatum*, *C. gottingense*, *C. axilliensis* and *C. jamesii* isolates predicted using antiSMASH 7. 171
Supplementary Figure 15. Phylogeny of putative prophages 172
Supplementary Figure 16. Global PADLOC analysis reveals profiles of phage defence systems are generally conserved within drep secondary clusters. 173
Supplementary Figure 17. Phage defence systems found in drep secondary clusters of *C. axilliensis*, *C. jamesii*, *C. amycolatum* and *C. gottingense* 174
Supplementary Figure 18. Phage defence systems found in drep secondary clusters of *C. tuberculostearicum* and *C. kefirresidentii* 175
Supplementary Figure 19. Phage defence systems found in drep secondary clusters of *C. marquesiae* 176
Supplementary Figure 20. The proportion of the total number of isolates derived from each of the 4 volunteers which bin to the 7 identified species (*C. tuberculostearicum*, *C. marquesiae*, *C. kefirresidentii*, *C. amycolatum*, *C. axilliensis*, *C. gottingense* and *C. jamesii*) 177

CHAPTER 5: The combined roles of staphylococci and corynebacteria in thioalcohol malodour formation

Figure 1. The pathway to generation of 3M3SH from the precursor Cys-Gly-3M3SH 200
Figure 2. Isolates of *S. hominis* and *C. amycolatum* are able to efficiently biotransform Cys-Gly-3M3SH and Cys-3M3SH 202
Figure 3. Cross-genus tandem biotransformation of Cys-Gly-3M3SH 204
Figure 4. Tandem biotransformation of Cys-Gly-3M3SH by pairs of *Corynebacterium* isolates 207
Supplementary Figure 1. The DTNB-based Cys-Gly-3M3SH/Cys-3M3SH biotransformation assay 214
Supplementary Figure 2. Thiols/3M3SH standard curve 215
Supplementary Figure 3. The experimental workflow to assess the biotransformation capabilities of a single organism in malodour generation. 216
Supplementary Figure 4. Isolates of common cutaneous species were not able to biotransform Cys-Gly-3M3SH or Cys-3M3SH 217

Supplementary Figure 5. The experimental workflow to assess the tandem biotransformation capabilities of pairs of strains in malodour generation. 218
Supplementary Figure 6. Tandem biotransformation of Cys-Gly-3M3SH is possible with pairs of *Corynebacterium* isolates 219

CHAPTER 6: Conclusions and future directions

Figure 1. The proposed network contributing to malodour formation.....228

Chapter 1

Thesis introduction

1.1 The human skin – structure and composition

The integumentary system is widely known to be the largest organ of the human body and forming more than 10% of body mass[1,2]. The outermost component of the integumentary system is the human skin, a complex structure composed of multiple layers and appendages that together provides overall protective barrier between the internal and external environments of the host. One of the skin's main functions is thermoregulation, where the skin detects changes in temperature in the external environment and signals the body to adapt to these changes[3–6]. The skin is also an important barrier to external hazards that could seriously harm the host[1].

Environments between different skin sites of the body can differ depending on the components of the skin itself. The structure of the skin is generally comprised of the dermal and epidermal layer which lies above the hypodermal layer or subcutaneous tissue of the integumentary system (Fig 1).

1.1.1 Epidermis

The epidermis is the outermost layer of the skin and serves as the first line of defence for the host. This layer of skin is stratified with the main constituent cells arranged like bricks of a wall. Within this structure lie sub-layers (from the deepest to the outermost) stratum basale, stratum spinosum, stratum granulosum and stratum corneum. The main constituents of the epidermis are stratified squamous epithelial cells called keratinocytes. Keratinisation is a process where basal cells of the stratum basale divide and gradually migrate up the epithelial layer to the stratum corneum before ultimately shedding and being replaced by new cells.

The fully cornified keratinocytes of the stratum corneum, called corneocytes, are embedded in a mix of nonpolar lipids to form a strong structure in this outermost layer of the skin[7,8]. This is a resilient barrier which provides protection against multiple diverse stresses including chemicals, microbial pathogens and excessive transepidermal water loss[8].

Aside from the barrier function, the stratum corneum is also home to a complex community of commensal microbes called the skin microbiome. These commensals live harmlessly on the human skin and provide some extra protection against stresses, in particular against common microbial pathogens. The axillary/underarm skin microbiome is a major focus of this study (see Chapter 1.4).

1.1.2 Dermis

Beneath the epidermis lies the dermis which is a much thicker layer of connective tissue mainly comprised of collagen fibre bundles[9]. The dermis provides structural support and nutrient delivery to maintain the function of the epidermis. Most importantly, the dermis is home to a variety of appendages including hair, the associated hair-follicles and skin-associated glands. These glands are namely the eccrine, sebaceous, and apocrine glands and the distribution of these glands make a specific skin site either oily (e.g., face), non-sebaceous/dry (e.g., volar forearm) or moist (e.g., underarm/axilla).

1.2 The human skin – Glands and associated secretions.

Skin-associated glands can either be sebaceous (oily) glands or sweat glands (eccrine and apocrine). Each type of skin-associated gland secretes a distinct profile of metabolites that provide each skin site a characteristic environment.

1.2.1 The sebaceous gland

Sebaceous glands are usually associated to the hair follicles and are developed with age up to 17 years, with rapid development occurring during adolescence[10]. While associated with acne vulgaris, sebaceous gland secretions (also known as sebum) are a major source of skin surface lipids including triglycerides, squalene and free fatty acids[11]. Sebum is secreted by specialised glandular cells called sebocytes into the sebaceous duct where it meets the hair follicle and, in some instances, the open skin. Sebum, and in particular its triglyceride and free fatty acid components, aid skin hydration preventing skin desiccation.

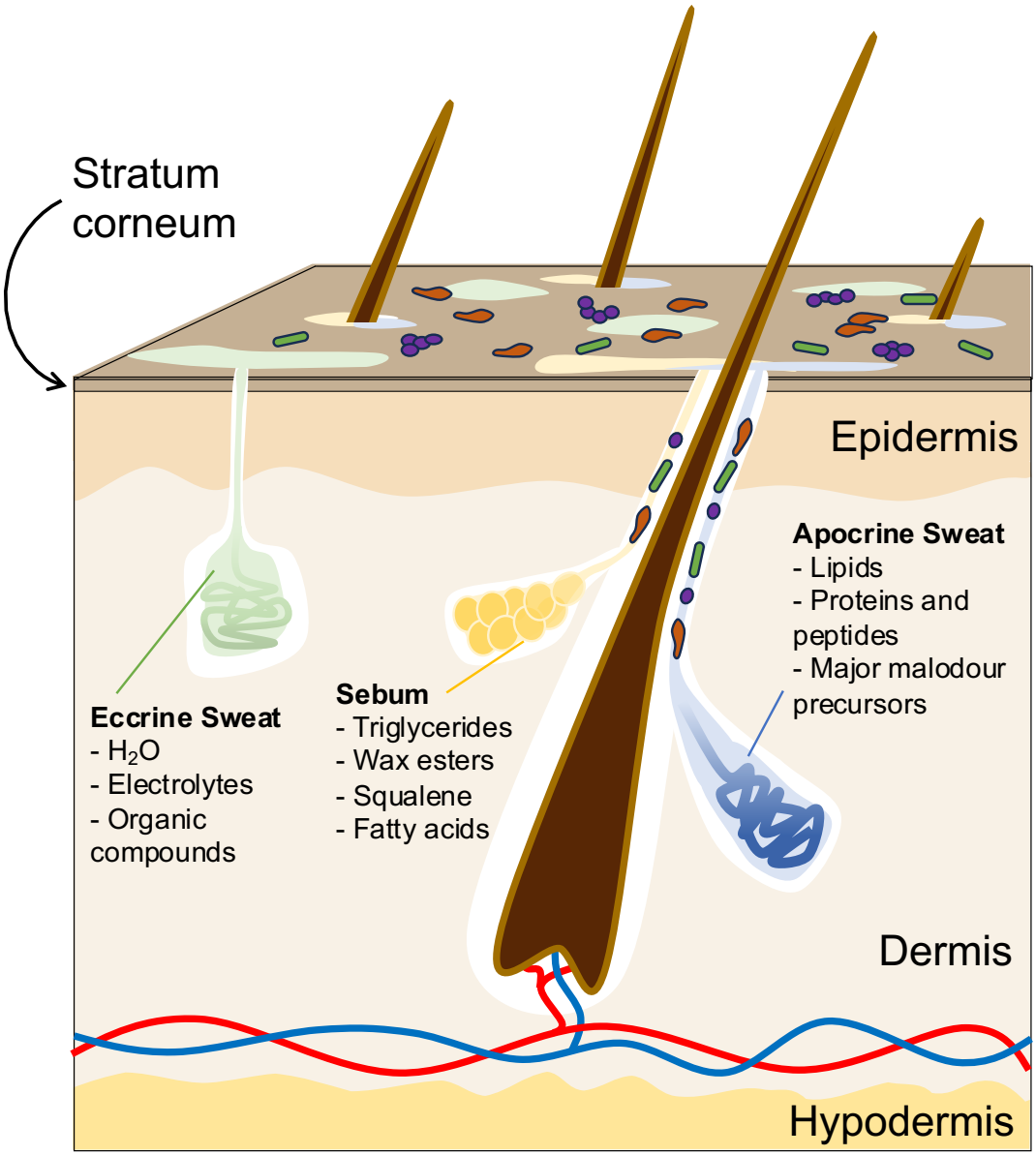


Figure 1. The structure of human skin. The dermal and epidermal layers sit above the hypodermis (also referred to as subcutaneous tissue). Sweat (eccrine and apocrine) and sebaceous glands reside within the dermis and can be associated to hair follicles. The epidermis is home to a community of microbes (represented in the figure by purple cocci, green rods and brown drumsticks) called the skin microbiome. Some members of the skin microbiome can also reside within hair follicles.

1.2.2 The eccrine gland

Eccrine glands are the most abundant skin-associated glands and can be found in a wide variety of skin environments[12]. Secretions from this gland originate from the secretory coil and travel up the sweat duct directly to the skin surface[13]. These secretions include water, salts and other metabolites including lactate and amino acids[14]. Together, these secretions function to maintain the hydration of the skin and are critical in the skin's function as a thermoregulator[15].

1.2.3 The apocrine gland

Unlike the eccrine gland, the apocrine gland is selectively present in some body sites including the anogenital regions (e.g., groin), areola and underarm/axilla[16]. Apocrine glands are usually found in the lower layers of the dermis and are typically associated to the hair follicle. In the axilla, these glands are found in high abundance and open to the follicles of the coarse hair associated to this skin site.

Apocrine glands are commonly associated with odour secretions due to the presence of multiple odour precursors[17]. Human axillary secretions are rich in peptides (including proteins), various lipids and steroids. While these secretions are nearly odourless, some of these metabolites could be acted upon by resident skin bacteria to produce various types of scents that are commonly associated with malodour (see Chapter 1.3). These secretions do not serve a clear purpose in the maintenance of the skin. Rather, it has been hypothesised that these secretions and subsequent processing by skin resident bacteria may be a vestigial scent signalling system previously used in communication. Other primates have also been reported to have scent glands which are strongly associated with communication and behaviour[18,19]. Scent glands of even more evolutionary distinct mammals like the skunk also closely resemble apocrine sweat glands[20]. Hence, it stands to reason that the human apocrine glands could have originated for the purposes of communication but is now a remnant through human evolution[21]. Although humans have developed more advanced communication, the odours associated with apocrine gland secretions can have major effects on human interactions[22].

1.3 Human malodour

Molecules that contribute to human odour tend to be small and volatile. These molecules are detected via the suite of olfactory receptors found in the human nose. Over 400 human olfactory receptors are known to exist and are known to be essential in differentiating over 1 trillion different olfactory simuli[23]. These complex interactions provide the individual with their perception of each odour, be it pleasant or otherwise. What humans perceive as body odour could range between odour descriptors like salty, acidic, oniony and musky. These odour profiles are defined by the presence of specific metabolites that are derived from mainly apocrine secretions (Table 1). In this section, the three reported sources of human malodour are discussed.

1.3.1 Malodour - Steroids

The 16-androstene class of human-derived steroids, mainly 5α-androstenone and 5α-androstenol, are sources of musk-like odour which were first identified in the 1970s[24,25] (Table 1). *Corynebacterium*, a commonly identified genus of the skin microbiome, was first implicated in the generation of odorous steroids through the processing of human derived odourless precursors[26]. However, studies have not found a positive correlation between body odour and detected concentrations of these steroids suggesting that this class of odour molecules does not significantly contribute to malodour[27].

1.3.2 Malodour – Volatile fatty acids (VFAs)

Unlike steroids, volatile fatty acids (VFAs) are a class of molecule that is thought to contribute to human malodour. Short chain (C_2 – C_5) and medium chain (C_6 – C_{11}) volatile fatty acids have previously been attributed to certain characteristic human malodour.

1.3.2.1 Short-chain VFAs

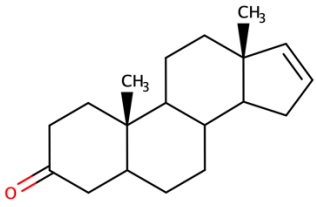
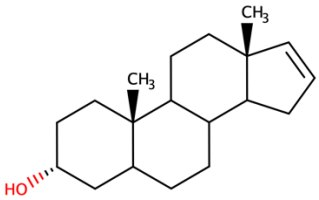
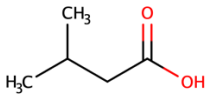
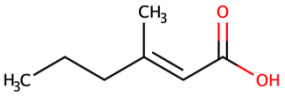
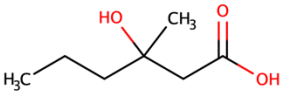
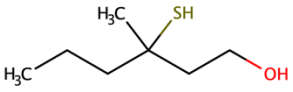
Propionic acid (a C_3 VFA) and acetic acid (a C_2 VFA) share a characteristic sour, vinegar-like odour and are derived from metabolites L-lactic acid and glycerol

respectively[28]. Skin resident organisms like *Cutibacterium* spp. (formerly *Propionibacterium*[29]) and *Staphylococcus* spp. are thought to be responsible for the conversion of these abundant skin metabolites to the respective short-chain VFAs. However, there is little evidence that propionic acid and acetic acid are prominent causes of human malodour. *Staphylococci*, in particular *S. epidermidis*, are also capable of generating highly odorous short-chain branched VFAs like isovaleric acid (Table 1) from the amino acid, L-leucine[30,31]. These methyl-branched VFAs are thought to be the major contributors to the characteristic sour/acidic form of malodour of both the human axilla and the feet.

1.3.2.2 Medium-chain VFAs

Combining chromatography and organoleptic assessments of human sweat, Zeng *et al*. first identified medium-chain (C_6-C_{11}) VFAs, in particular 3-methyl-2-hexanoic acid (3M2H) to be important contributors of human malodour. The odour of the unusual branched VFA 3M2H is described as sweaty, musky and goat-like, which differs significantly from the short-chain VFAs (Table 1). Initial studies revealed that 3M2H and the related odorous axillary molecule 3-hydroxy-3-methylhexanoic acid (HMHA) are both derived from L-glutamine conjugates derived from the apocrine glands of the host[32,33]. Natsch and colleagues hypothesised that Gln-3M2H and Gln-HMHA are hydrolysed to their respective odorous metabolites using an aminoacylase derived from a strain of *Corynebacterium* called *Corynebacterium* sp. Ax20. However, a wider whole-cell study performed by James and colleagues at Unilever failed to capture significant axillary corynebacterial aminoacylase activities against Gln-3M2H and Gln-HMHA[34]. Findings from Natsch *et al*. are based on the aminoacylase from the isolate *C*. sp. Ax20 which was first reported to be *C. striatum* but 16S rRNA analysis performed by James and colleagues places this isolate closer to *C. glaucum*. While both species are thought to exist in the axilla, they are not present in high abundance[34]. The study by James *et al*. could have been limited by the inability to isolate and subsequently test isolates of *Corynebacterium* which are able to biotransform these Gln-conjugated VFAs to the levels observed by Natsch et al. with *C*. sp. Ax20.

Table 1: Major odour molecules identified to date. Each molecule is identified by the name and class of malodour, the chemical structure and their commonly perceived odour. Oxygen atoms and alcohol groups are in red and thiol groups are in olive.

Name (Class of molecule)	Odour molecule	Odour descriptors
5α-androstenone (Steroid)		Urine, sweat, floral
5α-androstenol (Steroid)		Musky
Isovaleric acid (Short-chain VFA)		Sour/acidic, cheesy
3-methyl-2-hexanoic acid (Medium-chain VFA)		Sweaty, musky, goat-like
3-hydroxy-3-methylhexanoic acid (Medium-chain VFA)		Cumin-like/spicy
3-methyl-3-sulfanylhexan-1-ol (Thioalcohol)		Meaty, spicy, oniony

Interestingly, a 2014 study by Fujii *et al*. implicated another genus of skin bacteria, *Anaerococcus*, in the Gln-VFA aminoacylase activity[35]. Using the isolate *Anaerococcus* sp. A20 (later speciated to be *Anaerococcus nagyae*[36]) Fujii and colleagues identified aminoacylase activities in particular against Gln-HMHA. A recent study then characterised an aminoacylase from *A. nagyae* and homologues from other species including *Anaerococcus octavius* which were found to be active against glutamine-conjugates of both 3M2H and HMHA[37], supporting the findings of Fujii *et al*. in implicating this genus to VFA generation.

1.3.3 Malodours - Thioalcohol-based malodour

One of the most distinct and pungent causes of human malodour are molecules which contain thioalcohols. These highly odorous molecules are perceived as meaty, spicy and oniony which are almost completely distinct from the odour profiles of VFAs. In 2004, three independent studies concurrently reported the discovery of a range of thioalcohols that could contribute to malodour, including the major component 3-methyl-3-sulfanylhexan-1-ol (3M3SH)[38–40] (Table 1). Like 3M2H and HMHA, 3M3SH is a branched aliphatic molecule, and as the name suggests, contains thiol groups (-SH) which is major functional group that contributes to the characteristic odour profile.

The study by Natsch *et al*. also suggested the precursor of 3M3SH to be a L-cysteine-(S)-conjugate (i.e., Cys-3M3SH)[40] which was later suggested to be derived from a glycine containing precursor (i.e., Cys-Gly-3M3SH)[41,42] (Fig 2) which itself was found to be derived from the molecule glutathione-(S)-3M3SH[43]. In apocrine secretory cells, glutathione-(S)-3M3SH is broken down to Cys-Gly-3M3SH by γ-glutamyl transferase 1 (GGT1) which is then released by the transporter ABCC11. Axilla resident microbes then import Cys-Gly-3M3SH; first processing it to Cys-3M3SH and then to the odorous 3M3SH. Like in the case of odorous medium-chain VFAs, the studies on 3M3SH by Natsch and colleagues focused on the isolate *Corynebacterium* sp. Ax20 where they identified and characterised the dipeptidase TpdA to remove the glycine from Cys-Gly-3M3SH and then the cysteine-(S)-conjugate β-lyase AecD to complete the biotransformation process to generate 3M3SH[40,41]. However, there was limited evidence for whole cell biotransformation of Cys-Gly-3M3SH to 3M3SH in

Corynebacterium. The early study by Starkenmann *et al*. compared the Cys-Gly-3M3SH biotransformation capabilities of *Corynebacterium xerosis*, *Staphylococcus epidermidis* and *Staphylococcus haemolyticus* and found that the latter is by far the strongest biotransformer of this malodour precursor[42].

Bawdon *et al*. performed a similar assessment using a larger range of axillary *Corynebacterium* and *Staphylococcus* isolates[44]. Like Starkenmann *et al*., Bawdon and colleagues similarly concluded that a group of axillary *Staphylococcus* composing of *S. haemolyticus*, *S. lugdunensis* and *S. hominis* are major biotransformers of Cys-Gly-3M3SH. However, some 3M3SH generation from Cys-Gly-3M3SH was also observed with 7 out of 8 *Corynebacterium amycolatum* isolates, suggesting that some species of this genus could indeed perform this activity and existing studies could be limited by isolates available to test.

Unsurprisingly, the study by Bawdon *et al*. shifted the focus of thioalcohol-based malodour studies from *Corynebacterium* to the identified group of axillary *Staphylococcus* spp. Prior to this PhD project, we attempted to reveal the staphylococcal proteins responsible for the known processes required to generate the odorous 3M3SH from Cys-Gly-3M3SH (Fig 2).

In a collaboration with the Newstead group at the University of Oxford, we characterised a ubiquitous Staphylococcal transporter of the proton-dependent oligopeptide transporter (POT) family called DtpT that is thought to be responsible for the uptake of Cys-Gly-3M3SH into the cell[45]. In this previous study, we focused on the DtpT from the *S. hominis* B10 isolate which was first reported in Bawdon *et al*. Through biochemical and biophysical characterisation of DtpT, we revealed the presence of a protein binding pocket that could accommodate the unique shape of Cys-Gly-3M3SH, a dipeptide with a large modification on the sidechain. This binding pocket is not a common feature amongst POT transporters found in other genera including *Corynebacterium* (data not shown), suggesting that this protein may have adapted to

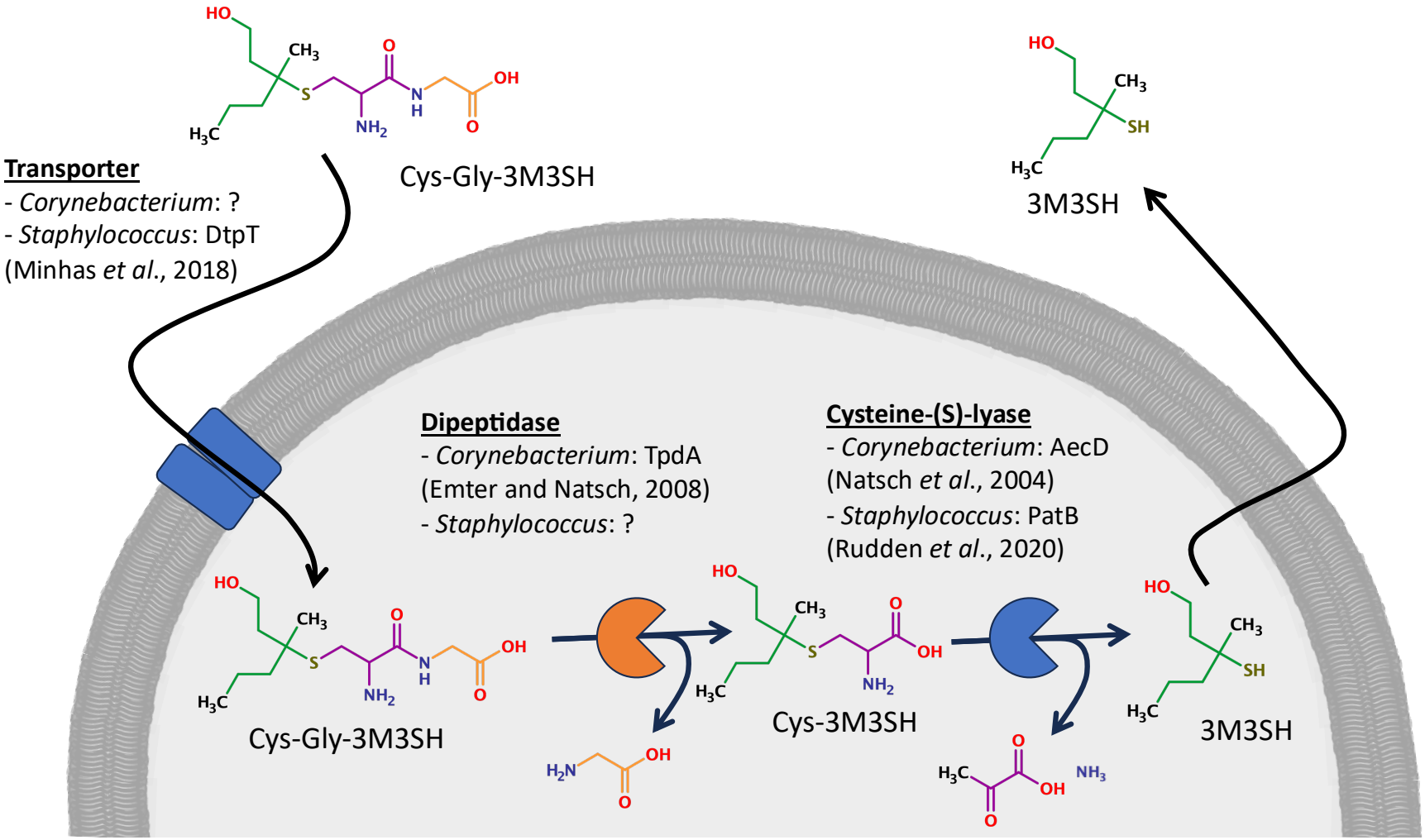


Figure 2. The known pathways to malodour production. Staphylococcal and corynebacterial proteins that are known to contribute to the known process of thioalcohol-based malodour generation were identified. Unknown proteins are indicated by a question mark.

this ligand. Despite the presence of this pocket, the modified dipeptide Cys-Gly-3M3SH still accommodated less ideally than the di- and tri-peptides tested in this study.

The second enzymatic step of this process is thought to involve a cysteine-(S)-conjugate β-lyase. As discussed earlier, Natsch *et al*. first identified the protein AecD of the PatB family of cysteine-(S)-conjugate β-lyases to be highly active against the intermediate molecule Cys-3M3SH[40].

Interestingly, only species from the group of *Staphylococcus* found with 3M3SH generation capabilities have a conserved cysteine-(S)-conjugate β-lyase of the PatB family[46]. Naturally, we proceeded to test this staphylococcal PatB against Cys-3M3SH, again focusing on the protein from *S. hominis* B10. Through comprehensive biochemical testing and structural characterisation, we determined that *S. hominis* PatB is able to selectively target Cys-3M3SH over other cysteine-(S)-conjugates including the common metabolite L-cystathionine. Like DtpT, PatB has a defined binding pocket that almost perfectly accommodates Cys-3M3SH, with the largely hydrophobic 3M3SH accommodated by a region of hydrophobic residues. This could explain the exclusion of L-cystathionine and further suggests this protein could be active against cysteine-(S)-conjugates of other odorous thioalcohols previously identified by Natsch *et al*.[40].

The first enzymatic step involves the conversion of Cys-Gly-3M3SH to Cys-3M3SH which is thought to be performed by a dipeptidase. Natsch *et al*. identified the M20-family dipeptidase TpdA in *Corynebacterium* sp. Ax20 which was able to perform this function[41]. However, we were not able to identify a homologue of TpdA in any staphylococci, including the group of *Staphylococcus* spp. active against Cys-Gly-3M3SH (data not shown). This suggests the presence of another class of enzyme capable to perform the first enzymatic step of 3M3SH generation from Cys-Gly-3M3SH.

1.4 The axillary microbiome

Some major human body sites and organs, like the gut and the skin, are home to communities of microbes called the microbiome. These organisms usually live symbiotically with the healthy host, contributing to various factors like immunomodulation and protection against pathogens[47,48]. Microbiomes are complex communities of microbes including bacteria, fungi and viruses. The diversity within each microbiome is thought to be determined by the nature of the specific body site. The skin is a nutrition limited environment and discrete skin sites can have different characteristics (i.e., moist, dry, sebaceous/oily) depending on the structure of the overall skin, and the types and number of glands present in these sites[49]. The sebaceous regions of the skin like the forehead and face are thought to have poor microbial diversity. It has been suggested that these sebaceous regions favour the growth of organisms like the skin abundant lipid auxotrophic *Cutibacterium* spp. and *Corynebacterium* spp.[49–51].

Moist sites like the axilla, on the other hand, are more nutrient rich and home to a more diverse community of microbes dominated by *Staphylococcus* spp., *Corynebacterium* spp. and *Cutibacterium* spp.[52,53]. As discussed in Chapter 1.3, the axillary microbes are thought to play major roles in human malodour generation from multiple host-derived precursors. The abundance of *S. hominis*, a major cause of thioalcohol-based malodour (e.g., 3M3SH), is known to correlate with malodour intensity, unlike other species of *Staphylococcus*, including the highly abundant skin species *S. epidermidis* [54].

Interestingly, abundance of the genus *Corynebacterium* also strongly correlates with human malodour[53–55] even though whole cell assessments of malodour formation by isolates of this genus have yet to provide significant evidence of Cys-Gly-3M3SH biotransformation activity[44]. Assessments of microbial malodour generation, like the study performed by Bawdon *et al*., tend to utilise collections of axillary isolates. After harvesting, these isolates are commonly assessed using 16S rRNA gene sequencing which only sometimes allows for the species level resolution. For example, the cutaneous species *Corynebacterium tuberculostearicum* has an identical 16S rRNA sequence to two other closely related species of skin *Corynebacterium* that form the

C. tuberculostearicum species complex (average nucleotide identity ~93-95%)[56]. Hence, phenotypic studies of skin isolates are limited by the resolution of the genetic characterisation technique used to classify each isolate.

1.5 Aims of project

The overarching goal of this thesis is to uncover the roles of *Corynebacterium* spp. in malodour formation alongside the group of malodour producing *Staphylococcus* spp. highlighted by Bawdon *et al*.[44]. These roles may involve single organism pathways or in conjunction with other members of the skin microbiome, including non-malodour producing *Staphylococcus* spp. The work presented in this thesis attempts to satisfy the main objective through three different aims:

1. To complete the pathway to malodour formation in *Staphylococcus*.
2. To assess the diversity and assemble isolates of corynebacteria and staphylococci from the same axilla of a high odour producing individual.
3. To test the capabilities of the assembled isolates in thioalcohol-based malodour production either individually or together with other members of the same axillary community.

Prior to this work, the *Staphylococcus* pathway to malodour formation has largely been elucidated, with the exception of the dipeptidase responsible for converting Cys-Gly-3M3SH to Cys-3M3SH. Aim 1 is addressed in Chapter 2 of this thesis through the identification of the ubiquitous staphylococcal dipeptidase PepV which is capable of converting Cys-Gly-3M3SH to Cys-3M3SH as efficiently as the corynebacterial dipeptidase TpdA.

The unclear and potentially limited genetic diversity of isolates used in previous whole cell malodour studies was highlighted as a major limitation previously. Aim 2 is explored in the work reported in Chapters 3 and 4, where a large number of isolates from volunteers with strong malodour were collected and genetically characterised using a high-throughput long-read sequencing workflow to capture the diversity within each axillary region. These isolates establish the York Skin Microbiome (YSM) collection: a growing collection of genetically characterised skin microbiome isolates

from defined skin sites of matched individuals to enable more realistic community-based phenotypic studies. Chapter 3 discusses the diversity of axillary *Staphylococcus* from a single volunteer using the genetic characterisation workflow. Chapter 4 explores the diversity of axillary *Corynebacterium* from four volunteers in this body of work. Both studies employ a comprehensive bioinformatics pipeline to predict the roles of these isolates within their respective source environments.

Aim 3 is investigated in Chapter 5 through the characterisation of 23 *Staphylococcus* and *Corynebacterium* isolates from a single high malodour individual. This work aims to provides evidence that *Corynebacterium* spp. may still play a role in thioalcohol-based malodour formation and explore the “tandem biotransformation” route to malodour formation where two isolates from the same microbial community could together convert Cys-Gly-3M3SH to 3M3SH in cross-species or maybe even cross-genus interactions.

1.6 Thesis introduction references

1. The Structure and Function of Skin. *Dermatological and Transdermal Formulations* 19–58 (2002) doi:10.1201/9780824743239-4.
2. Kim, J. Y. & Dao, H. Physiology, Integument. *StatPearls* (2023).
3. Nakamura, K. Central circuitries for body temperature regulation and fever. *Am J Physiol Regul Integr Comp Physiol* **301**, 1207–1228 (2011).
4. Kanosue, K., Crawshaw, L. I., Nagashima, K. & Yoda, T. Concepts to utilize in describing thermoregulation and neurophysiological evidence for how the system works. *Eur J Appl Physiol* **109**, 5–11 (2010).
5. Nakamura, K. & Morrison, S. F. A thermosensory pathway mediating heat-defense responses. *Proc Natl Acad Sci U S A* **107**, 8848–8853 (2010).
6. Nakamura, K. & Morrison, S. F. A thermosensory pathway that controls body temperature. *Nat Neurosci* **11**, 62 (2007).
7. Murphy, B. *et al.* Alteration of barrier properties, stratum corneum ceramides and microbiome composition in response to lotion application on cosmetic dry skin. *Scientific Reports 2022 12:1* **12**, 1–11 (2022).
8. Marks, R. The Stratum Corneum Barrier: The Final Frontier. *J Nutr* **134**, 2017S-2021S (2004).
9. Yousef, H., Alhajj, M., Fakoya, A. O. & Sharma, S. Anatomy, Skin (Integument), Epidermis. *StatPearls* (2024).
10. Zouboulis, C. C. Acne and sebaceous gland function. *Clin Dermatol* **22**, 360–366 (2004).
11. Picardo, M., Ottaviani, M., Camera, E. & Mastrofrancesco, A. Sebaceous gland lipids. *Dermatoendocrinol* **1**, 68 (2009).
12. Sato, K., Kang, W. H., Saga, K. & Sato, K. T. Biology of sweat glands and their disorders. I. Normal sweat gland function. *J Am Acad Dermatol* **20**, 537–563 (1989).
13. Collins, K. J., Sargent, F. & Weiner, J. S. The effect of arterial occlusion on sweat-gland responses in the human forearm. *J Physiol* **148**, 615 (1959).
14. Harker, M., Coulson, H., Fairweather, I., Taylor, D. & Daykin, C. A. Study of metabolite composition of eccrine sweat from healthy male and female human subjects by 1H NMR spectroscopy. *Metabolomics* **2**, 105–112 (2006).
15. Best, A., Lieberman, D. E. & Kamilar, J. M. Diversity and evolution of human eccrine sweat gland density. *J Therm Biol* **84**, 331–338 (2019).
16. Hodge, B. D. & Brodell, R. T. *Anatomy, Skin, Sweat Glands*. *StatPearls* (2018).

17. Shelley, W. B., Hurley, H. J. & Nichols, A. C. Axillary odor; experimental study of the role of bacteria, apocrine sweat, and deodorants. *AMA Arch Derm Syphilol* **68**, 430–46 (1953).
18. Zeller, U., Epple, G., Küderling, I. & Kuhn, H. -J. The anatomy of the circumgenital scent gland of Saguinus fuscicollis (Callitrichidae, Primates). *J Zool* **214**, 141–156 (1988).
19. Suárez-Bonnet, A., Priestnall, S. L., Ramírez, G. A., González-Sánchez, C. & Jaber, J. R. Scent (Apocrine) Gland Adenocarcinoma in a Wedge-Capped Capuchin Monkey (Cebus olivaceus): Histological and Immunohistochemical Features. *J Comp Pathol* **179**, 1–6 (2020).
20. Heidt, G. A. & Morgans, L. F. Comparative histology of the scent glands of five species of skunks. *J Anat* **134**, 121 (1982).
21. Russell, M. J. Human olfactory communication. *Nature* **260**, 520–522 (1976).
22. Stern, K. & McClintock, M. K. Regulation of ovulation by human pheromones. *Nature* **392**, 177–179 (1998).
23. Bushdid, C., Magnasco, M. O., Vosshall, L. B. & Keller, A. Humans can discriminate more than 1 trillion olfactory stimuli. *Science (1979)* **343**, 1370–1372 (2014).
24. Gower, D. B. 16-Unsaturated C19 steroids a review of their chemistry, biochemistry and possible physiological role. *J Steroid Biochem* **3**, 45–103 (1972).
25. Brooksbank, B. W. L., Brown, R. & Gustafsson, J. A. The detection of 5 α-androst-16-en-3 α-ol in human male axillary sweat. *Experientia* **30**, 864–865 (1974).
26. Gower, D. B., Holland, K. T., Mallet, A. I., Rennie, P. J. & Watkins, W. J. Comparison of 16-Androstene steroid concentrations in sterile apocrine sweat and axillary secretions: Interconversions of 16-Androstenes by the axillary microflora—a mechanism for axillary odour production in man? *J Steroid Biochem Mol Biol* **48**, 409–418 (1994).
27. Zeng, X. nong *et al.* Analysis of characteristic odors from human male axillae. *J Chem Ecol* **17**, 1469–1492 (1991).
28. James, A. G., Casey, J., Hyliands, D. & Mycock, G. Fatty acid metabolism by cutaneous bacteria and its role in axillary malodour. *World J Microbiol Biotechnol* **20**, 787–793 (2004).
29. Scholz, C. F. P. & Kilian, M. The natural history of cutaneous propionibacteria, and reclassification of selected species within the genus propionibacterium to the proposed novel genera acidipropionibacterium gen. Nov., cutibacterium gen. nov.

and pseudopropionibacterium gen. nov. *Int J Syst Evol Microbiol* **66**, 4422–4432 (2016).

30. James, A. G., Hyliands, D. & Johnston, H. Generation of volatile fatty acids by axillary bacteria1. *Int J Cosmet Sci* **26**, 149–156 (2004).
31. Ara, K. *et al.* Foot odor due to microbial metabolism and its control. *https://doi.org/10.1139/w05-130* **52**, 357–364 (2011).
32. Natsch, A., Gfeller, H., Gygax, P., Schmid, J. & Acuna, G. A Specific Bacterial Aminoacylase Cleaves Odorant Precursors Secreted in the Human Axilla *. *Journal of Biological Chemistry* **278**, 5718–5727 (2003).
33. Troccaz, M. *et al.* Gender-specific differences between the concentrations of nonvolatile (R)/(S)-3-Methyl-3-Sulfanylhexan-1-Ol and (R)/(S)-3-Hydroxy-3-Methyl-Hexanoic acid odor precursors in axillary secretions. *Chem Senses* **34**, 203–210 (2009).
34. James, A. G. The Axillary Microbiome and its Relationship with Underarm Odor. *Skin Microbiome Handbook: From Basic Research to Product Development* 79–129 (2020) doi:10.1002/9781119593058.CH5.
35. Fujii, T., Shinozaki, J., Kajiura, T., Iwasaki, K. & Fudou, R. A newly discovered Anaerococcus strain responsible for axillary odor and a new axillary odor inhibitor, pentagalloyl glucose. *FEMS Microbiol Ecol* **89**, 198–207 (2014).
36. Veloo, A. C. M., de Vries, E. D., Jean-Pierre, H. & van Winkelhoff, A. J. Anaerococcus nagyae sp. nov., isolated from human clinical specimens. *Anaerobe* **38**, 111–115 (2016).
37. Rose, M. Characterisation of anaerococcal enzymes required for production of volatile fatty acids in human body odour. (2023).
38. Hasegawa, Y., Yabuki, M. & Matsukane, M. Identification of new odoriferous compounds in human axillary sweat. *Chem Biodivers* **1**, 2042–2050 (2004).
39. Troccaz, M., Starkenmann, C., Niclass, Y., Van De Waal, M. & Clark, A. J. 3-Methyl-3-sulfanylhexan-1-ol as a Major Descriptor for the Human Axilla-Sweat Odour Profile. *Chem Biodivers* **1**, 1022–1035 (2004).
40. Natsch, A., Schmid, J. & Flachsmann, F. Identification of Odoriferous Sulfanylalkanols in Human Axilla Secretions and Their Formation through Cleavage of Cysteine Precursors by a C□S Lyase Isolated from Axilla bacteria. *Chem Biodivers* **1**, 1058–1072 (2004).

41. Emter, R. & Natsch, A. The sequential action of a dipeptidase and a beta-lyase is required for the release of the human body odorant 3-methyl-3-sulfanylhexan-1-ol from a secreted Cys-Gly-(S) conjugate by Corynebacteria. *J Biol Chem* **283**, 20645–52 (2008).
42. Starkenmann, C., Niclass, Y., Troccaz, M. & Clark, A. J. Identification of the precursor of (S)-3-methyl-3-sulfanylhexan-1-ol, the sulfury malodour of human axilla sweat. *Chem Biodivers* **2**, 705–716 (2005).
43. Baumann, T. *et al.* Glutathione-conjugated sulfanylalkanols are substrates for ABCC11 and γ-glutamyl transferase 1: A potential new pathway for the formation of odorant precursors in the apocrine sweat gland. *Exp Dermatol* **23**, 247–252 (2014).
44. Bawdon, D., Cox, D. S., Ashford, D., James, A. G. & Thomas, G. H. Identification of axillary Staphylococcus sp. involved in the production of the malodorous thioalcohol 3-methyl-3-sufanylhexan-1-ol. *FEMS Microbiol Lett* **362**, 1–10 (2015).
45. Minhas, G. S. *et al.* Structural basis of malodour precursor transport in the human axilla. *Elife* **7**, (2018).
46. Rudden, M. *et al.* The molecular basis of thioalcohol production in human body odour. *Sci Rep* **10**, 12500 (2020).
47. Sanz, Y. *et al.* The gut microbiome connects nutrition and human health. *Nature Reviews Gastroenterology & Hepatology 2025* 1–22 (2025) doi:10.1038/s41575-025-01077-5.
48. Byrd, A. L., Belkaid, Y. & Segre, J. A. The human skin microbiome. *Nature Reviews Microbiology 2018 16:3* **16**, 143–155 (2018).
49. Grice, E. A. & Segre, J. A. The skin microbiome. *Nature Reviews Microbiology 2011 9:4* **9**, 244–253 (2011).
50. Jensen, M. G. *et al.* Species- and strain-level diversity of Corynebacteria isolated from human facial skin. *BMC Microbiol* **23**, 1–14 (2023).
51. Costello, E. K. *et al.* Bacterial community variation in human body habitats across space and time. *Science (1979)* **326**, 1694–1697 (2009).
52. Grice, E. A. *et al.* Topographical and temporal diversity of the human skin microbiome. *Science (1979)* **324**, 1190–1192 (2009).
53. Taylor, D. *et al.* Characterization of the microflora of the human axilla. *Int J Cosmet Sci* **25**, 137–145 (2003).
54. Troccaz, M. *et al.* Mapping axillary microbiota responsible for body odours using a culture-independent approach. *Microbiome* **3**, 1–15 (2015).

55. Leyden, J. J., McGinley, K. J., Holzle, E., Labows, J. N. & Kligman, A. M. The Microbiology of the Human Axilla and Its Relationship to Axillary Odor. *Journal of Investigative Dermatology* **77**, 413–416 (1981).
56. Ahmed, N. *et al.* Genomic characterization of the C. tuberculostearicum species complex, a prominent member of the human skin microbiome . *mSystems* **8**, (2023).

Chapter 2

Characterisation of the Staphylococcal Cys-Gly-3M3SH dipeptidase PepV

2.1 Manuscript information

Identification of a staphylococcal dipeptidase involved in the production of human body odour

Reyme Herman[1], Bethan Kinniment-Williams[1,2], Michelle Rudden[1,3], Alexander Gordon James[3], Anthony J. Wilkinson[4], Barry Murphy[5], Gavin H. Thomas[1*]

[1] Department of Biology, University of York, Wentworth Way, York, YO10 5DD, UK
[2] Hull York Medical School, University of York, Heslington, York, YO10 5DD, UK
[3] School of Life Sciences, University of Hull, Hull, HU6 7RX, UK
[4] Unilever, Colworth Science Park, Sharnbrook, Bedfordshire, MK44 1LQ, UK
[5] York Structural Biology Laboratory, Department of Chemistry, University of York, Wentworth Way, York, YO10 5DD, UK
[6] Unilever Research & Development, Port Sunlight Laboratory, Quarry Road East, Bebington, Wirral, Merseyside, CH63 3JW, UK

Journal of Biological Chemistry

DOI: 10.1016/j.jbc.2024.107928

Published on 23 October 2024

Co-authors have given permission for this manuscript to be used in this thesis.

2.2 Context within thesis

While the two enzymatic steps in thioalcohol-based malodour production have been characterised in *Corynebacterium*, prior to this study, only the second enzymatic step had been characterised in *Staphylococcus* with the dipeptidase step being performed by an unknown enzyme. Chapter 2 discusses the identification of the ubiquitous Staphylococcal dipeptidase PepV which is capable of performing the dipeptidase step as efficiently as the *Corynebacterium* Cys-Gly-3M3SH dipeptidase, TpdA. With this study, we complete the identification of possible enzymatic pathways to malodour formation in both genus of interest to this study. Revealing these processes contributes to the understanding of the biochemical steps that could occur in the axillary microbial communities to contribute to malodour, be it via individual strains or with other members of the same microbiome.

2.3 Abstract

The production of human body odour is the result of the action of commensal skin bacteria, including *Staphylococcus hominis*, acting to biotransform odourless apocrine gland secretions into volatile chemicals like thioalcohols such as 3-methyl-3-sulphanylhexan-1-ol (3M3SH). As the secreted odour precursor Cys-Gly-3M3SH contains a dipeptide, yet the final enzyme in the biotransformation pathway only functions on Cys-3M3SH, we sought to identify the remaining step in this human-adapted biochemical pathway using a novel coupled enzyme assay. Purification of this activity from *S. hominis* extracts led to the identification of the M20A-family PepV peptidase (ShPepV) as the primary Cys-Gly-3M3SH dipeptidase. To establish whether this was a primary substrate for PepV, the recombinant protein was purified and demonstrated broad activity against diverse dipeptides. The binding site for Cys-Gly-3M3SH was predicted using modelling, which suggested mutations that might accommodate this ligand more favourably. Indeed, a D437A resulted in an almost 6-fold increase in the k_{cat}/K_M, while other introduced mutations reduced or abolished function. Together these data identify an enzyme capable of catalysing the missing step in an ancient human-specific biochemical transformation and suggest that the production of 3M3SH neither uses a dedicated transporter nor peptidase for its breakdown, with only the final cleavage step, catalysed by PatB C-S β-lyase, being a unique enzyme.

2.4 Introduction

The production of body odour or malodour in humans is a process in which microbes on the skin play a central role through their bioconversion of odourless secreted molecules to volatile odourants (1–3). These precursor molecules are produced by the apocrine glands, which are associated exclusively with hair follicles in specific sites like the underarm and the genital regions. Unlike eccrine sweat glands that are found in all human skin and have roles in thermoregulation, the physiological functions of apocrine gland secretions are less well understood (4). However, it is known that human malodour precursors like S-(1-hydroxy-3-methylhexan-3-yl)-L-cysteinylglycine (Cys-Gly-3M3SH) (5), volatile fatty acid conjugated glutamine (6, 7), steroids and long-chain fatty acids (8, 9) are secreted onto the human skin via the apocrine gland, and

there is now extensive evidence that many of these compounds are immediately metabolised by the resident skin microbiome living within the hair follicle or the surrounding surface (9, 10).

The human skin is stably colonised by a unique community of microbes collectively known as the skin microbiome. Taxonomic characterisation of the human skin microbiome (11, 12) has identified *Staphylococcus* as key genera that persistently colonise the skin, alongside *Cutibacterium* and *Corynebacteria* species. The axilla hosts its own unique microbiota (13) which is dominated by both staphylococci and corynebacteria (14). Both these microbial genera play a key role in axillary malodour formation. Bacterial biotransformation of the odourless Cys-Gly-3M3SH precursor involves its uptake via a peptide transporter and subsequent enzymatic action by a peptidase followed by a cysteine-S-conjugate β-lyase (C-S β-lyase) which generates the highly pungent volatile molecule 3M3SH (15–19) (Fig. 1A). Initially, the dipeptidase and the C-S β-lyase capable of the complete enzymatic processing of Cys-Gly-3M3SH were identified in an axillary isolate of corynebacteria (17, 18). Subsequently work from us and others focussed on the more predominant staphylococci and identified key roles for specific *Staphylococcal* species in 3M3SH production using whole cell *in vivo* assays with the physiological Cys-Gly-3M3SH precursor. We identified a host-restricted clade of axillary staphylococci including *S. hominis* and *Staphylococcus haemolyticus* (5, 16, 20) as the key microbial species capable of metabolising Cys-Gly-3M3SH directly to volatile 3M3SH. We identified the molecular basis for the biotransformation in *S. hominis* for two of the three steps: (1) a proton-dependent oligopeptide transporter (POT) required for uptake of Cys-Gly-3M3SH (15), and (2) a uniquely modified PLP-dependent PatB enzyme that cleaves Cys-3M3SH leading to the production of 3M3SH (16). Importantly this enzyme specifically recognises the Cys-3M3SH precursor and not the dipeptide conjugated Cys-Gly-3M3SH, suggesting that an intracellular staphylococcal dipeptidase is essential for production of human malodour.

Through a biochemical approach, the Cys-Gly-3M3SH active corynebacterial M20A family dipeptidase TpdA was identified in the skin isolate *Corynebacterium* sp. Ax20 (17). This protein, which is thought to be a metallopeptidase of the succinyldiaminopimelate desuccinylase (DapE) subfamily, was found to be active

against Cys-Gly-3M3SH, but not against the non-conjugated dipeptide of Cys-Gly. Here, we set out to identify the dipeptidase responsible for this function in the staphylococcal malodour producer, *S. hominis*. We report the identification and biochemical characterisation of the Cys-Gly-3M3SH active promiscuous metallopeptidase PepV from *S. hominis*.

2.5 Results

2.5.1 Biochemical identification of a dipeptidase with activity against Cys-Gly-3M3SH

To identify the enzyme(s) catalysing the conversion of Cys-Gly-3M3SH to Cys-3M3SH in staphylococci, we first looked for homologues of the *Corynebacterium* M20-family dipeptidase TpdA, that is known to catalyse this reaction (17) but no homologues are present in any cutaneous *Staphylococcus* genomes. Therefore, we widened our computational analysis of putative peptidases using the MEROPS database (21), which predicted 40 such enzymes encoded in *S. hominis*, 10 of which were classified to be of the M20 family that contains TpdA. Previously reported examples include the tripeptidase PepT (22, 23) of the M20B subfamily and the promiscuous dipeptidase PepV (24, 25) of the M20A subfamily. Alternatively, the M17 subfamily leucyl aminopeptidase PepA might also possess Cys-Gly dipeptidase activity with previous homologues shown to have this activity (26–28). Due to the high number of potential enzymes and the lack of genetic tools to manipulate *S. hominis*, a biochemical approach was undertaken to isolate the major Cys-Gly-3M3SH dipeptidase from cell extracts.

To enable us to screen for the desired activity in cellular fractions, we modified a DTNB assay that detects 3M3SH, which we used previously to study the function of the C-S β-lyase ShPatB *in vitro* (16) and the transporter DtpT *in vivo* (15). Here, we exploit the fact the ShPatB will not work on the substrate of the dipeptidase reaction, Cys-Gly-3M3SH, but will act on its product, Cys-3M3SH to produce the DTNB reactive 3M3SH (Supplementary Data S1). Whole cell extracts of *S. hominis* B10 were prepared and fractionated using a series of chromatographic methods using the coupled ShPatB-DTNB assay to measure biochemical activity across fractions. Proteins were first

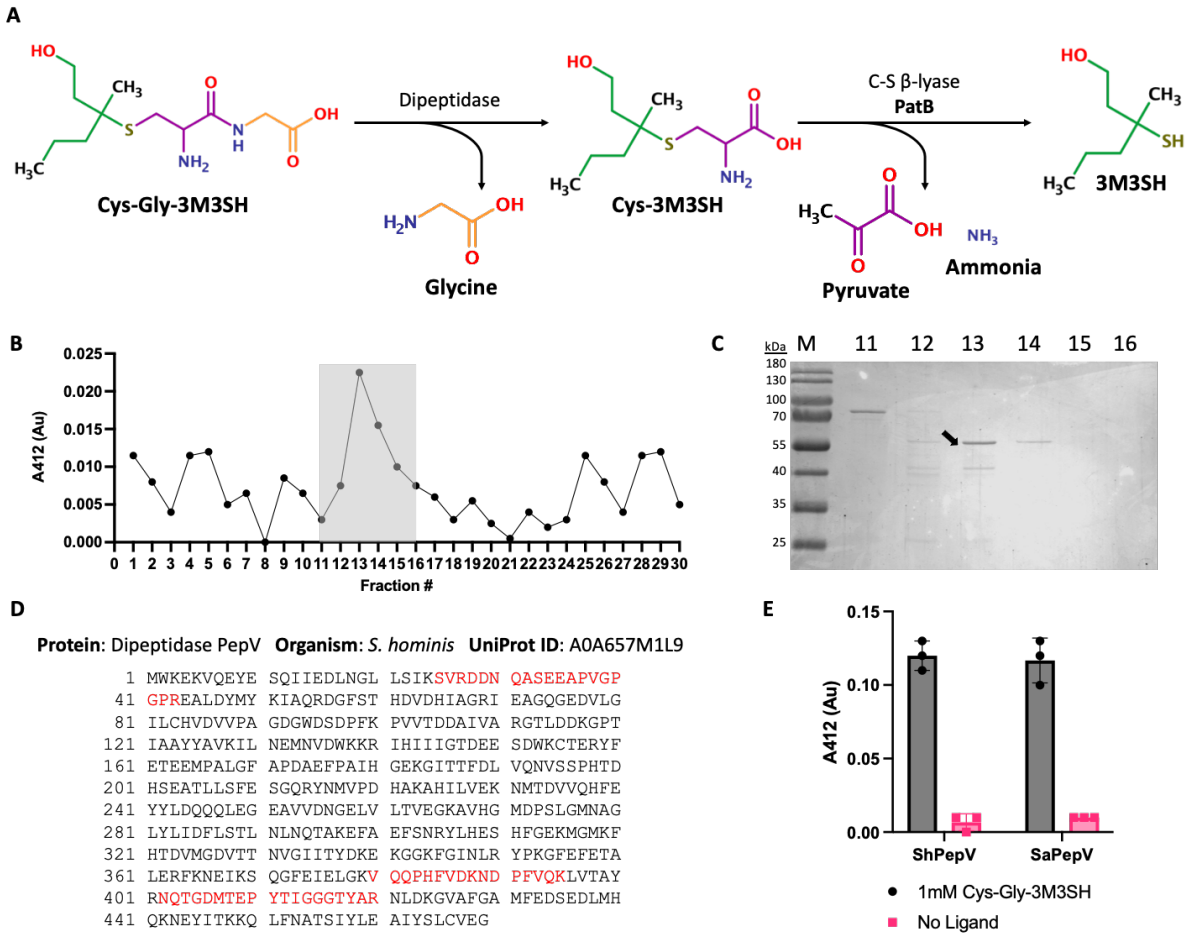


Figure 1. (A) The generation of the odorous molecule 3M3SH from Cys-Gly-3M3SH in *S. hominis* is catalysed by a dipeptidase and PatB, a C-S β-lyase. (B, C) Isolation and (D) identification of the dipeptidase ShPepV, a protein (E) capable of Cys-Gly-3M3SH dipeptidase activity in *S. hominis*. A coupled DTNB assay was used to colorimetrically detect Cys-Gly-3M3SH dipeptidase function in ShPepV and *S. aureus* PepV (SaPepV). (B, shaded region indicates the fractions further analysed. C, SDS-PAGE analysis of column fractions 11-16 - arrow denotes the protein band excised for analysis by mass spectrometry. D, sequence of PepV from *S. hominis* – the segments in red denote peptides which were identified by mass spectrometry. E, Results of coupled DTNB assays suggest the abilities of both ShPepV and SaPepV in contributing to the biotransformation of Cys-Gly-3M3SH (data in technical triplicates with errors (s.d)).

separated on a strong anion exchange (AEX) column followed by a hydrophobic interactions chromatography (HIC) column then a weak AEX column and finally by size-exclusion chromatography (SEC) (Supplementary Data S2-5). Following the final step, fractions with detectable activity were pooled (Fig. 1B), concentrated and analysed by SDS-PAGE revealing a prominent species with an apparent molecular mass of ~55 kDa in fraction 13 (Fig. 1C, Supplementary Data S6). This band was excised and the protein within it was unambiguously determined to be the PepV protein from *S. hominis* by peptide mass fingerprinting (ShPepV) through the identification of three distinct and one modified peptide (Fig. 1D, Supplementary Data S7).

To confirm that ShPepV has dipeptidase activity against Cys-Gly-3M3SH, we recombinantly expressed and purified ShPepV and the structurally characterised *S. aureus* homologue SaPepV (29). Using the coupled DTNB assay it was determined that both enzymes, when present at 1 μM, had activity on 1 mM Cys-Gly-3M3SH (Fig. 1E). These data suggest that PepV is the major peptidase catalysing this reaction in extracts of *S. hominis*. Since PepV from non-thioalcohol producing *S. aureus* (which lacks the C-S lyase PatB) can also catalyse this reaction, we suggest that PepV dipeptidase activity is most likely conserved among other staphylococci and has not specifically evolved for malodour biotransformation.

2.5.2 ShPepV is a general dipeptidase active on diverse L-amino acid–containing dipeptides

Having established that ShPepV catalyses the production of Cys-3M3SH from Cys-Gly-3M3SH, and that the corresponding enzyme from a non-thioalcohol-producing staphylococci also catalyses this reaction, we tested the hypothesis that this protein has a more general cellular function in dipeptide turnover, rather than a specialised role in malodour production. Indeed, SaPepV is known to cleave various dipeptides and also to hydrolyse β-lactams (30). Similar activity has been observed for other bacterial M20A enzymes such as the PepV from *Lactobacillus delbrueckii* subsp. *lactis* DSM 7290 (LdPepV) (24).

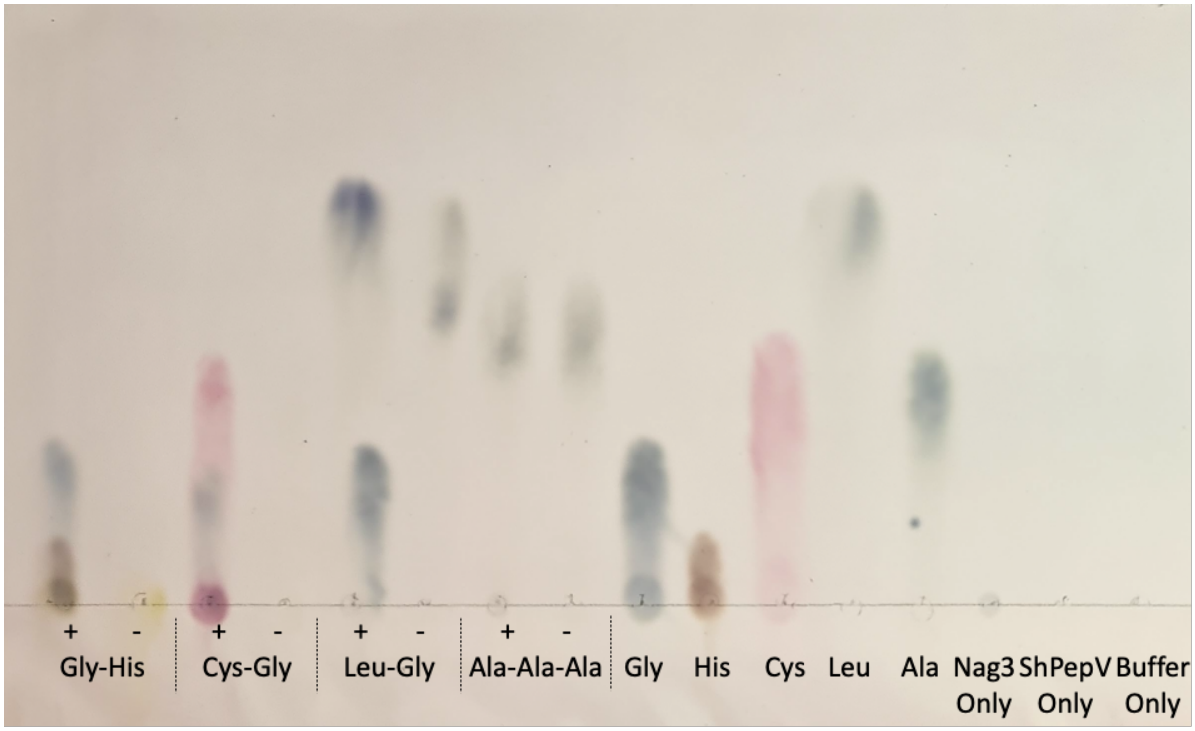


Figure 2. TLC-based assay of ShPepV activity against a subset of dipeptides tested. Peptides were incubated with (+) or without (-) PepV. Relevant protein and amino acid controls were included. All substrates and controls were solubilised in dH_2O. Other TLC plates are shown in Supplementary Data S9-10.

The substrate specificity of ShPepV was assessed using a range of commercially available dipeptides with varying chemistries and sizes at the P1 and P1' positions (Supplementary Table 4). With the exception of D-Ala-D-Ala, which has a very different stereochemistry to the other L-amino acid containing peptides, ShPepV was capable of cleaving all dipeptides tested (Fig. 2 and Supplementary Data S9-10). However, it was unable to cleave the tripeptide L-Ala-L-Ala-L-Ala, suggesting preference and selectivity for dipeptide substrates (Fig. 2).

2.5.3 ShPepV and SaPepV are efficient Cys-Gly-3M3SH dipeptidases

We next monitored the kinetics of PepV-catalysis against malodour substrates. The study of SaPepV by Girish and Gopal (29) revealed that the protein is metal-dependent with highest catalytic activity when purified in the presence of Mn^{2+}. To determine if ShPepV behaves similarly, we removed co-purifying metals from the recombinantly purified protein through dialysis against an EDTA-containing buffer (apo-ShPepV) and added a range of divalent metals to the protein before testing activity against Cys-Gly-3M3SH. We observe striking resemblance to the findings of Girish and Gopal (29) with Mn^{2+} bound ShPepV exhibiting the highest activity with Co^{2+} having the second highest (Fig. 3). In light of this, downstream enzyme assays of ShPepV and SaPepV were carried out with proteins purified in the presence of Mn^{2+} and TCEP.

Using the coupled DTNB assay, we determined the kinetic parameters of the two staphylococcal PepV proteins with Cys-Gly-3M3SH as the substrate (Table 1). For comparison, we expressed and purified the previously characterised Cys-Gly-3M3SH dipeptidase TpdA (17) from *Corynebacterium glaucum* DSM 30827 (31) as a known positive control (CgTpdA).

ShPepV displayed the highest catalytic efficiency (k_{cat}/K_M) of the three enzymes with TpdA exhibiting comparable activity against Cys-Gly-3M3SH (Table 1, Supplementary Data S11 B-D). Interestingly, the catalytic efficiency (k_{cat}/K_M) of SaPepV with Cys-Gly-3M3SH is approximately 3-fold lower than that of ShPepV. While the k_{cat} remains comparable, the weaker K_M values derived for the *S. aureus* homolog seems to be contributing to the lower catalytic efficiency suggesting the enzyme is slightly less capable of accommodating Cys-Gly-3M3SH.

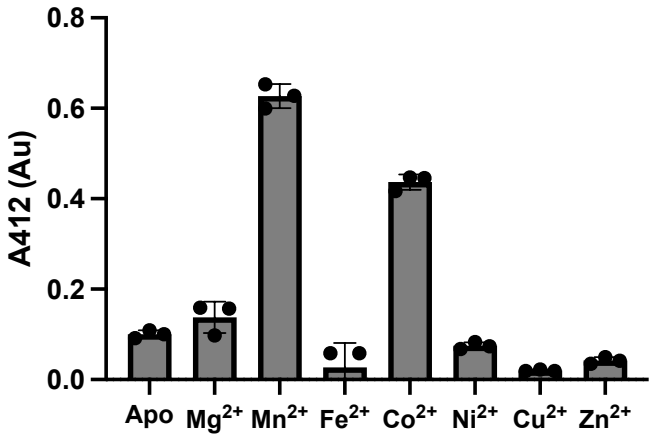


Figure 3. The dipeptidase activity of ShPepV against Cys-Gly-3M3SH in the presence of various divalent cations reveals a preference for Mn2+ as the metal cofactor. A negative control (Apo) containing no added divalent cations was included. Enzyme activity was determined using the coupled DTNB assay (data in technical triplicates with errors (s.d)).

Table 1. Kinetics of ShPepV, SaPepV and CgTpdA with Cys-Gly-3M3SH and ShPepV with Cys-Gly. The assay used for kinetic parameter determination is also included. Kinetic values and errors (s.d) reported were corrected to 2 s.f.

Protein	Substrate	Assay Used	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (s^{-1} mM^{-1})
ShPepV	Cys-Gly-3M3SH	Coupled DTNB	3.41 ± 0.07	1.29 ± 0.04	2.64 ± 0.10
SaPepV	Cys-Gly-3M3SH	Coupled DTNB	2.96 ± 0.14	3.34 ± 0.34	0.89 ± 0.10
CgTpdA	Cys-Gly-3M3SH	Coupled DTNB	1.20 ± 0.07	0.48 ± 0.03	2.50 ± 0.85
ShPepV	Cys-Gly	Cd-ninhydrin	405.40 ± 23.56	2.19 ± 0.22	186.00 ± 21.76

To determine the effects of the large 3M3SH group in the catalytic activity of ShPepV with Cys-Gly-3M3SH, we compared its activity to that against the dipeptide Cys-Gly. To measure this, we used a previously reported Cd-ninhydrin method (32) to detect free amino acids (Table 1, Supplementary Data S11 A). Surprisingly, the K_m was ~2-fold more than with Cys-Gly-3M3SH suggesting a slightly less preferable binding. However, the enzyme turnover number (k_{cat}) with Cys-Gly was almost 100-fold higher than with Cys-Gly-3M3SH which resulted in an overall efficiency increase of ~70-fold. We have demonstrated that ShPepV is capable of cleaving Cys-Gly-3M3SH at similar rates exhibited *in vitro* by the TpdA enzymes from *Corynebacterium glaucum* DSM 30827 (29) but is far superior with the simpler Cys-Gly dipeptide.

2.5.4 ShPepV is general dipeptidase that can accommodate Cys-Gly-3M3SH

The ability of ShPepV to function using Cys-Gly-3M3SH as a substrate, suggests that the active site accommodates the bulky (S)-conjugate even though it is unlikely to have evolved primarily for this function. To address this hypothesis, we exploited the high sequence identity between ShPepV and SaPepV (79.4% identity, 90.2% similarity) and used the existing structure of the latter (PDB: 3KI9) (29) as a template for homology modelling of ShPepV. The model of ShPepV was generated with high confidence. The GMQE (Global Model Quality Estimation) score was determined to be 0.91 which suggests high confidence of the generated tertiary structure. The QMEANDisCo Global value was calculated to be 0.86 ± 0.05 which suggests high per-residue quality for the model against the template. The overall structure of ShPepV conforms to the canonical PepV fold consisting of a catalytic domain and a 'lid' domain enclosing a putative metal containing active site (Fig. 4A). As expected, both staphylococcal PepVs share identical residues in their putative active sites and hence, the two metal cofactors expected in the binding site (Fig. 4B). However, despite the high sequence identities of ShPepV and SaPepV, the latter exhibited weaker dipeptidase activity with Cys-Gly-3M3SH (Table 1, column k_{cat}/K_M).

Our previous structural studies demonstrated the importance of largely hydrophobic cavities in the transporter DtpT (15) and the C-S β-lyase PatB (16) for the selective accommodation of the largely hydrophobic 3M3SH moiety. The model of ShPepV reveals an interdomain cleft which passes through the protein and is likely to be the

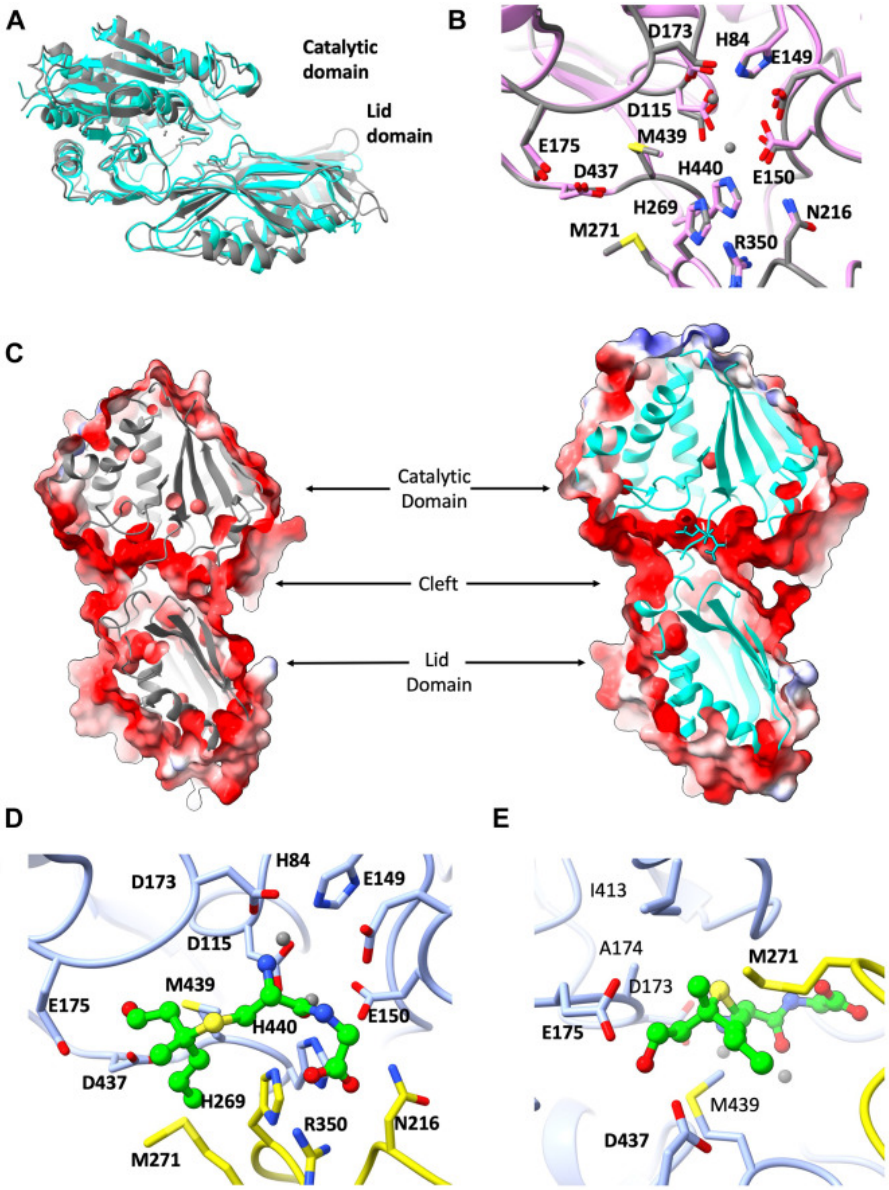


Figure 4. (A) The model of ShPepV (grey) suggests the canonical PepV fold with a catalytic and lid domain that is observed in the experimentally derived structure of LdPepV (cyan) (PDB: 1LFW, r.m.s.d 1.66Å over 447 residues). (B) The putative binding site of ShPepV (grey) shares the same residues as SaPepV (pink). (C) Surface representations of ShPepV (grey ribbon) and LdPepV (cyan ribbon) reveal a common substrate binding cleft. (D) Using a rigid-body refinement approach, Cys-Gly-3M3SH (ball and stick representations) was modelled into the binding site of ShPepV using the known positions of the Asp-Ala inhibitor in LdPepV. (E) A rotated view of the modelled ligand binding site indicating various buried and surface (bold) residues that were mutagenized. (D, E: Carbons of the catalytic domain of ShPepV are represented in light blue and that of the lid domain is in gold). Images were generated with ChimeraX 1.8 (33).

peptide binding site (Fig. 4C). A similar cleft was observed in the cross-section of the experimentally derived structure of *Lactobacillus delbrueckii* PepV (LdPepV, PDB: 1LFW), containing the inhibitor Asp-Ala phosphinic acid (24). Using this structure and the position of the dipeptide mimic, we performed a manual rigid body refinement to suggest a fit for Cys-Gly-3M3SH into the model of ShPepV. We observed highly conserved residues with roles in the binding of the dipeptide backbone of both structures (Supplementary Data S12). The proposed binding cleft accommodates Cys-Gly-3M3SH with the major protein interactions expected to be with the dipeptide backbone (Fig. 4D). The 3M3SH moiety of the fitted ligand extends out of the binding site into a large cavity which can accommodate the bulky hydrophobic group within the allowed molecular restraints.

2.5.5 Mutagenesis of the ShPepV binding pocket reveals key Cys-Gly-3M3SH accommodating residues

When fitting the hydrophobic part of 3M3SH, which includes a quaternary carbon, there were very few possible positions it could adopt in the binding cavity. The 3M3SH moiety seem to fit best in a pocket circumscribed by residues A174, I413, M439, D437, E175 and M271. This allowed us to identify residues that might modulate the quality of the fit of the ligand into the binding pocket (Fig. 4E). To test this, we recombinantly expressed ShPepV with these residues replaced by either smaller or larger residues to increase or decrease the size of the cavity, respectively, at its base or the surface. Consistent with the wildtype (WT) protein, these mutants were also purified in the presence of Mn^{2+} and TCEP.

Enlarging buried residues in the cavity through the substitutions A174V and I413W should severely restrict the volume around the quaternary carbon. The mutants retained 55% and 14.6% activity respectively when compared with the wildtype protein (Table 2). The K_M derived for the I413W mutant was almost 8-fold worse than the wildtype protein suggesting that ligand accommodation is significantly hampered by the much larger tryptophan residue (Fig. 5, Table 2, Supplementary Data S13A). A complete loss of activity against Cys-Gly-3M3SH is seen with a mutant of a metal coordinating residue D173A (Table 2). This was expected as the metal cofactor is

Table 2. Kinetics of ShPepV mutants with Cys-Gly-3M3SH determined using the coupled DTNB assay. The fold change in catalytic efficiencies ($(k_{cat}/K_M)_{mut}/(k_{cat}/K_M)_{WT}$) and differences in transition state binding energies ($\Delta\Delta G^{\neq}$) of the mutants against that of the wildtype (WT) protein were also calculated. (N.D.: not detectable)

	Purpose	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (s^{-1} mM^{-1})	$(k_{cat}/K_M)_{mut}/(k_{cat}/K_M)_{WT}$	$\Delta\Delta G^{\neq}$ ($kJ.mol^{-1}$)
WT	Native	3.41 ± 0.07	1.29 ± 0.04	2.64 ± 0.10	-	-
D173A	Metal coordinating	N.D.	N.D.	N.D.	-	-
A174V	Smaller cavity (Buried)	2.30 ± 0.05	1.58 ± 0.14	1.46 ± 0.13	0.55	-1.54
I413W	Smaller cavity (Buried)	3.01 ± 0.11	7.79 ± 0.67	0.39 ± 0.04	0.146	-4.96
M439A	Bigger cavity (Buried)	4.71 ± 0.05	0.70 ± 0.01	6.77 ± 0.13	2.56	2.42
D437A	Bigger cavity (Surface)	5.79 ± 0.06	0.37 ± 0.03	15.7 ± 1.27	5.96	4.60
E175A	Bigger cavity (Surface)	4.73 ± 0.21	0.85 ± 0.11	5.59 ± 0.75	2.11	1.93
M271R	Smaller cavity (Surface)	2.96 ± 0.03	2.31 ± 0.08	1.28 ± 0.05	0.486	-1.86

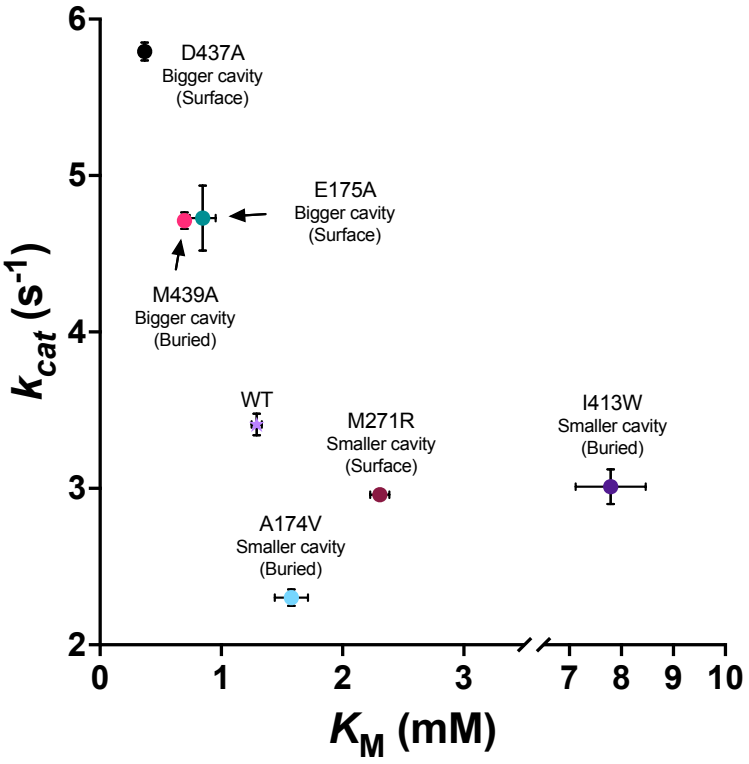


Figure 5. Plot of k_{cat} against K_M of ShPepV mutants with the wildtype protein. Mutants that constrict the binding site (A174V, M271R and I413W) result in varying reductions in affinity for Cys-Gly-3M3SH (higher K_M). Conversely, ShPepV mutants resulting in a bigger cavity (D437A, M439A and E175A) have improved affinity values for Cys-Gly-3M3SH.

thought to activate water for nucleophilic attack and/or polarise the peptide carbonyl for this attack.

Conversely, the M439A mutant which is expected to result in a larger binding site and retention of activity, leads to a marginal increase in k_{cat} and a ~2-fold decrease in the K_M providing a 2-fold improvement in k_{cat}/K_M when compared with the wildtype protein (Fig. 5, Table 2, Supplementary Data S13B). Continuing this logic, we also reduced the size of other residues that our model predicted might restrict the active site, specifically E175 and D437. Altering the first two of these residues to an alanine slightly improved the k_{cat} of the resulting enzymes with K_M increasing by ~1.5-fold and ~3.5-fold respectively (Fig. 5, Table 2, Supplementary Data S13C-D). Similarly, the residue enlarging M271R mutation resulted in a ~2-fold increase in K_M while retaining a similar turnover number (k_{cat}). Together these data suggest our structural model for the enzyme has some validity and by being able to actually increase the catalytic efficiency of ShPepV against Cys-Gly-3M3SH, that the enzyme clearly has not evolved to recognise this as its primary substrate.

We generally observe similar patterns of fold changes in the catalytic efficiencies of ShPepV mutants against the wildtype protein ($(k_{cat}/K_M)_{mut}/(k_{cat}/K_M)_{WT}$) with the dipeptide Cys-Gly (Supplementary Table 5, Supplementary Data 14). The activity is unchanged in the A174V mutant when compared to the wildtype protein which suggests this mutant is less important in the accommodation of the thiol group of the cysteine moiety of Cys-Gly than in the accommodation of the quaternary carbon conjugated sulphur of Cys-Gly-3M3SH. However, the I413W mutant only retained 0.0025% of the wildtype protein activity with the k_{cat} largely contributing to the decrease. In comparison, the same mutant led to only a ~7-fold decrease in enzymatic efficiency (14.6% activity retention) when compared with the wildtype protein with Cys-Gly-3M3SH (Table 2). Clearly but unexpectedly, this mutation has drastically affected the reaction involving the simpler dipeptide. Overall, our data suggest that the binding site could not only be improved for the complex malodour precursor Cys-Gly-3M3SH but also the simpler dipeptide Cys-Gly.

2.6 Discussion

Since the discovery of the important role staphylococci play in the generation of thioalcohol-based human malodour derived from the axillary metabolite Cys-Gly-3M3SH [5,19], a number of studies have helped identify and characterise proteins involved in the biotransformation of this molecule (15–18). The POT family transporter DtpT is involved in the cellular uptake of the odourless Cys-Gly-3M3SH into the cytoplasm (15), where a hitherto unknown dipeptidase and the C-S β-lyase PatB (16) cleave this molecule in tandem to produce the pungent 3M3SH (Fig 6). Here, we identified the Cys-Gly-3M3SH active dipeptidase ShPepV empirically through chromatography techniques with the aid of a coupled DTNB assay. Homologues of this protein were previously reported in the literature to be non-specific L-amino acid dipeptidases with no apparent selectivity for or against dipeptides with large sidechains on the N-terminus (24, 25, 29). Assays of the purified recombinant enzyme confirmed the dipeptidase activity of ShPepV with Cys-Gly-3M3SH. Using structural modelling, we identified a clear interdomain substrate binding cleft and suggested a cavity capable of accommodating the bulky 3M3SH group. This possible location was supported by mutagenic studies on the putative pocket which when enlarged, allowed for a more conducive fit as seen by an increase in apparent affinity (K_M). In contrast when the pocket size is reduced there is reduction of activity (Table 2). In addition, we also unexpectedly noted increases in k_{cat} of mutants that enlarges the binding site. These residues could play a role in the interdomain interactions which in turn controls the enzyme turnover. Clearly, our mutations created a more favourable environment for the dipeptidase activity with Cys-Gly-3M3SH by having improved transition state binding energies when compared to the wildtype protein (Table 2, column $\Delta\Delta G^{\neq}$). This suggest that PepV from *S. hominis* could have easily evolved to improve Cys-Gly-3M3SH turnover but this was not observed. We suggest that the dipeptidase step may not be the rate-limiting step of the reaction and hence, does not need to be improved. Alternatively, the general ability for *S. hominis* to biotransform Cys-Gly-3M3SH might not be an essential for cell survival in the axillary environment.

Despite the apparent ability for accommodation of an extended dipeptide like Cys-Gly-3M3SH, ShPepV was not able to cleave L-Ala-L-Ala-L-Ala. Vongerichten *et al.* (25) also observed this with LdPepV and found only a handful of other tripeptides that are

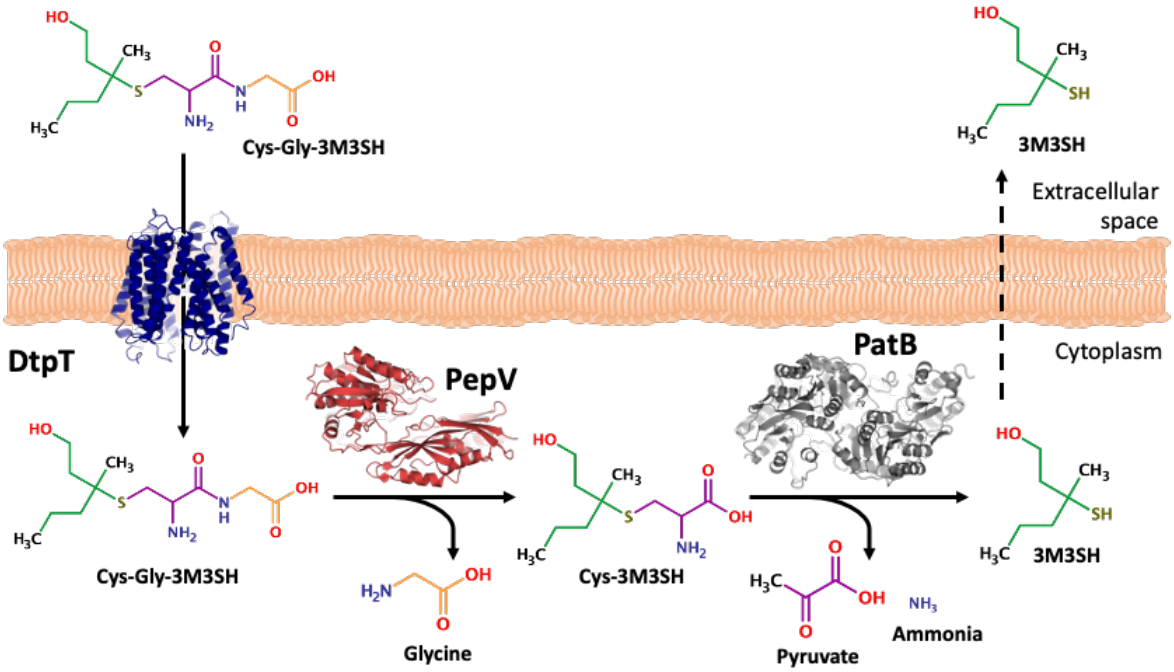


Figure 6. Uptake and enzymatic processing of Cys-Gly-3M3SH in *S. hominis*. The odourless human derived metabolite Cys-Gly-3M3SH is taken into the bacterial cell by the POT family transporter DtpT. The general dipeptidase PepV cleaves the glycine off before the C-S β-lyase activity of PatB releases the pungent molecule 3M3SH which escapes the cell through an unknown mechanism.

able to be hydrolysed by this dipeptide preferring enzyme. In our model, we observe the N-terminus of the dipeptide backbone facing D173 and H84 and the same was observed with the inhibitor bound structure of LdPepV (Fig. 4D and Supplementary Data S12). Together this suggest that PepV is a true dipeptidase.

The staphylococcal PepV enzymes are highly conserved with likely roles in metabolising exogenous and endogenous peptides and potentially in antibiotic resistance (29, 30). We have identified PepVs in a wide range of skin-associated staphylococci including the commonly found opportunistic pathogens *S. epidermidis* and *S. aureus*. Alongside ShPepV, we followed on from previous characterisation studies of the *S. aureus* homologue SaPepV (29, 30) to characterise its potential Cys-Gly-3M3SH dipeptidase ability. The residues that form the likely active site of both ShPepV and SaPepV were found to be identical in both the primary sequence and the tertiary structure (Fig. 4B). Consistent with our observations with ShPepV, we found SaPepV to be enzymatically active against Cys-Gly-3M3SH.

Clearly, proteins possessing functional capabilities in the biotransformation of Cys-Gly-3M3SH can exist in multiple species and strains that are found in the underarm, potentially giving them access to useful nutrients for growth using host secretions. However, as of yet it is still not clear as to the selective advantage, if any, of metabolising these malodour precursors. Malodour precursors like Cys-Gly-3M3SH and Gln-3M2H are released from the apocrine gland via the protein ABCC11 (34). Multiple genetic studies have identified SNPs at amino acid position 180 of ABCC11 which limits its ability to export these malodour precursors in people in some Asian communities (34–36). These individuals with lower axillary odour seemed to have a change in the distribution of genera within the axillary microbiome, shifting slightly from staphylococci and anaerococci to corynebacteria (37). This shift in population could be attributed to a change in the metabolites available due to the absence of a functioning ABCC11. Most skin resident staphylococci would stand to benefit from the products of the biotransformation of Cys-Gly-3M3SH, as resultant glycine from the PepV reaction can be used directly for protein synthesis (38) but also forms the major part of the crosslink between the glycan chains of the staphylococcal peptidoglycan(39, 40).

To date, we are unaware of how the expression of DtpT, PepV and PatB are regulated. The genes encoding these proteins are not clustered on the genome of *S. hominis* suggesting they may have alternative roles within the cell. Unlike the PatB enzyme, neither the transporter DtpT nor the peptidase PepV seem to be uniquely adapted for Cys-Gly-3M3SH biotransformation, further supporting the notion that this pathway piggybacks on other general cellular functions. In addition, while we purified PepV from an activity assay and were able to clearly measure its activity on Cys-Gly-3M3SH, we cannot say if this is the only peptidase in the cell that is active on this substrate. We are not able to truly replicate the conditions in the underarm *in vitro* nor is it feasible for us to perform genetic manipulations on this genetically intractable species of *Staphylococcus*. Hence, it is possible that other related enzymes also make some contribution to this process *in vivo*. Nonetheless, the ability to transport short peptides and to then cleave them into free amino acids in the cytoplasm is likely an important core nutritional function for skin dwelling staphylococci and indeed there is evidence that the peptide transporters Opp and DtpT are required for host colonisation (41–44). The function of peptidases must also be essential to use these transported peptides, but the multiplicity of these systems in the cell means that it is unlikely that removal of a single system would result in a strong cellular phenotype. However, there is some evidence for specialisation of peptidase function as PepV itself has been linked to β-lactam resistance (30), presumably by cleaving β-lactams intracellularly, while other more specific aminopeptidases are known to be important for *S. aureus* virulence (45).

In conclusion, this study provides evidence for the missing mechanistic step in the full pathway for the production of thioalcohol-based malodour in *S. hominis*. As DtpT and PepV are found widely in staphylococci, it supports our previous finding that the occurrence of *patB*, encoding the cysteine-(S) β-lyase, correlates very strongly with malodour production in staphylococci (16).

2.7 Experimental procedures

2.7.1 Bacterial strains and plasmids

Bacterial strains and plasmids used in this study are as listed in Supplementary Tables 1 and 2. In this study, all growth of *Escherichia coli* was performed using LB media (Luria Bertani) while staphylococci and corynebacteria were grown in TSB media (Tryptic Soy Broth).

2.7.2 Isolation and identification of ShPepV

A 1l overnight culture of *S. hominis* B10 was grown at 37°C, with shaking in a baffled flask. The cells were then pelleted and resuspended in 20 mM Tris-HCl, pH 8, supplemented with 5 µg lysostaphin (Merck). The suspension was incubated for an hour at 37°C before being sonicated to lyse the cells further. The lysate was clarified by centrifugation and the supernatant was loaded onto a 4.7ml HiScreen CaptoQ AEX column (Cytiva). Proteins were then fractionated using a 0 – 1 M NaCl linear gradient in 20 mM Tris-HCl pH 8.0. Each fraction from this and all subsequent purification steps was analysed for activity using a coupled DTNB assay as described below (Supplementary Data S1). Fractions with the highest activity were pooled and buffer exchanged into 50 mM potassium phosphate (KPi) pH 7.8, 1 M $(NH_4)_2SO_4$.
The proteins were then resolved by hydrophobic interaction chromatography on a HiScreen Phenyl HP HIC column (Cytiva). After washing protein were eluted using a descending (1 – 0 M) $(NH_4)_2SO_4$ linear gradient in 50 mM KPi pH 7.8 (Supplementary Data S2). Fractions with the highest activity were buffer exchanged into 20 mM Bis-Tris pH 7.0.

The proteins were further fractionated by a gradient elution into 20 mM Bis-Tris pH 7, 1 M NaCl using a HiScreen CaptoDEAE AEX column (Cytiva). The coupled DTNB assay was repeated as above but incubated for an hour (Supplementary Data S3). Fractions with the highest activity were finally loaded onto a HiPrep S200 Sephacryl SEC column (Cytiva) (Supplementary Data S4).

The fraction with the highest activity, the two fractions preceding and the two fractions succeeding were concentrated, activity confirmed by the coupled DTNB assay as above (Supplementary Data S6) and the remaining protein were loaded onto a 12% SDS-PAGE gel. The resulting gel was Coomassie stained and image was captured on a gel-doc system. The protein band consistent with the activity seen in the fraction was excised, eluted from the gel and digested with trypsin. The resulting peptide fragments were analysed by MALDI-MS/MS and searched against all UniProt entries within the species *Staphylococcus hominis*. Three distinct and one modified peptide fragments matched to the dipeptidase PepV (Accession number: A0A657M1L9).

2.7.3 Cloning, expression, and purification of ShPepV, SaPepV, and CgTpdA

Genomic DNA of *Staphylococcus hominis* B10, *Staphylococcus aureus* USA300 JE2 and *Corynebacterium glaucum* DSM30827 were isolated using the DNeasy Blood and Tissue Kit (Qiagen) with either lysostaphin or lysozyme used to digest the cell walls. The *pepV* genes from the staphylococci and *tpdA/dapE3* from the corynebacteria were amplified by PCR (Supplementary Table 3) and cloned into pBADcLIC using the ClonExpress II One Step Cloning Kit (Vazyme). This recombinant expression system introduces a C-terminal tag (-ENLYFQGHHHHHHHHHH). For protein production, these plasmids were introduced into *E. coli* MC1061. Overnight cultures were used for inoculation into fresh 1l LB supplemented with 100 μg mL^{-1} ampicillin. Cultures were grown to an OD_{600} of ~0.5 at 37°C before the induction of recombinant protein production with 0.01% L-arabinose. The culture was allowed to grow for a further 5 hours at 30°C. Cells were then harvested by centrifugation and the resulting pellet was resuspended in purification buffer A (50 mM KPi pH 7.8, 200 mM NaCl, 20% glycerol, 40 mM imidazole). The cells were then sonicated at 250W with a 30% duty cycle for 10 minutes. Clarification of the lysate was then performed at 27,000 x *g* for 30 minutes. Soluble protein fractions were then passed through a pre-equilibrated HisTrap FF crude column (Cytiva). The column was washed with purification buffer A over 10 column volumes and the bound proteins were eluted with purification buffer B (50 mM KPi pH 7.8, 200 mM NaCl, 20% glycerol, 500 mM imidazole). Prior to downstream experiments, the proteins were buffer exchanged into 20 mM Tris-HCl pH 8 and 50 mM NaCl. The Mn^{2+} bound protein used for the kinetics experiments were prepared

as described above but with 1 mM $MnCl_2$ and 3 mM TCEP added to all purification buffers.

2.7.4 Mutagenesis of ShPepV

Mutants of the recombinantly expressed ShPepV was generated by inverse PCR using the mutagenic primer pairs mentioned on Supplementary Table 3 with pBcL::Sh*pepV* as the template. PCR products were circularised by blunt end ligation using 100 ng of DNA, 1 µL T4 DNA Ligase (NEB), 1 µL T4 Polynucleotide Kinase (NEB) in T4 DNA Ligase buffer (NEB) and incubated overnight at 16°C. The ligated products were transformed into *E. coli* XL1-Blue and then sequence verified. Verified plasmids were then transformed into *E. coli* MC1061 for protein production and purification in the presence of Mn^{2+} as mentioned in the previous section.

2.7.5 Coupled Cys-Gly-3M3SH DTNB assay

A coupled 5,5′-dithiobis-(2-nitrobenzoic acid) (DTNB) assay was used to detect peptidase activity against Cys-Gly-3M3SH (Supplementary Data S5). Reaction mixes contained 1 mM Cys-Gly-3M3SH (Concept Life Sciences), 1 µM purified ShPatB, 0.4 mM DTNB and 0.1 M Tris-HCl pH 8. Assays performed for the identification of the peptidase included the elution fractions at 25% (v/v). *In vitro* characterisation of ShPepV, its homologues and its mutants were performed with the same assay mixture containing 1 µM purified enzyme. Reaction mixtures were incubated at 37°C for 30 minutes and the absorbance at 412 nm (A_{412}) was measured using an Epoch 2 Spectrophotometer with a minimum detectable change of 0.0001 OD (BioTek/Agilent).

The steady-state kinetics of the PepV dipeptidase activity on Cys-Gly-3M3SH was monitored using the coupled DTNB assay with varying concentrations of Cys-Gly-3M3SH in a 96-well plate. The A_{412} values were measured every 15 seconds on a microplate spectrophotometer to determine the initial rate of substrate turnover for each concentration of ligand. The measured A_{412} values correspond to TNB which is generated in a 1:1 ratio to every 3M3SH generated. Michaelis-Menten plots were generated to derive the k_{cat} and K_M using Prism 10 (GraphPad). The catalytic efficiencies were then derived by calculating k_{cat} upon K_M.

2.7.6 Dipeptide activity assayed by TLC

100 mM stocks of peptides and amino acids (Supplementary Table 4) were prepared in dH_2O with the exception of L-tyrosine which was prepared in 1 M HCl. 20 µM ShPepV was incubated with 20 mM peptides in 20 mM Tris-HCl pH 8 and 50 mM NaCl for 30 minutes at 37°C. Nag3, an M20 peptidase family aminohydrolase from *Anaerococcus nagyae* (Accession: WP_117521798) was used as a negative control enzyme. 3 µl of each reaction mixture and equivalent amounts of the controls were dropped onto glass-backed silica TLC plates. A 90:10 mixture of ethanol:acetic acid was used as the mobile phase in a pre-equilibrated chamber. Following chromatography, the plates were allowed to dry, and a 1% ninhydrin solution (Merck) was sprayed over the plates. The plates were allowed to dry again before colour development using a heat gun.

2.7.7 Metal cofactor test

The Apo-ShPepV used in the metal cofactor test was prepared by dialysing the purified protein with 10 mM EDTA on ice for 5 hours followed by a dialysis over 18 hours into a buffer containing 20 mM Tris-HCl pH 8, 50 mM and 10 mM EDTA. The resulting protein was then buffer exchanged into 20 mM Tris-HCl pH 8 and 50 mM NaCl. For the assay, 1µM Apo-ShPepV was added into a buffer containing 20 mM Tris-HCl pH 8, 50 mM NaCl, 1mM Cys-Gly-3M3SH, 1 µM purified ShPatB and 1 mM of either of the following divalent cation salts ($MgCl_2$, $MnCl_2$, $FeCl_2$, $CoCl_2$, $NiCl_2$, $CuSO_4$, $ZnCl_2$) and incubated at 37°C for 30 minutes. The Apo-ShPepV without any added divalent cation salts was added as a negative control.

2.7.8 Kinetics of ShPepV with Cys-Gly

1 µM ShPepV was mixed with a range of concentrations of Cys-Gly between 0.125 mM and 2 mM in 50 mM KPi and 200 mM NaCl in a 96-well microtitre plate. Reactions were incubated at 37°C for 10 minutes. Samples were taken at 0, 5 and 10 minutes and the reactions were quenched by adding 10 mM of EDTA. 20 µl of the samples were mixed with 30µl of 2% ninhydrin and 50 µl H_2O. The colour was developed by

incubating the samples at 80°C for 10 minutes. The absorbance at 570 nm was measured on a microplate spectrophotometer. A standard curve was generated by mixing equimolar concentrations of L-cysteine and glycine between 0.125 mM and 2 mM, incubating the mixture for 10 minutes at 37°C and developed with ninhydrin before measurement at 570 nm as above. The A_{570} values measured from the dipeptidase reaction were then compared with those of the standard curve to derive the concentration of Cys-Gly that has been cleaved. A Michaelis-Menten plot was generated to derive the k_{cat} and K_m using Prism 9 (GraphPad). The catalytic efficiency was then derived by calculating k_{cat} upon K_m. The fold changes in (k_{cat} /K_m) and the changes in transition state energies ($\Delta\Delta G^{\neq}$) of ShPepV mutants against the wildtype (WT) were calculated (46)

2.7.9 Homology modelling of ShPepV and rigid-body refinement modelling of Cys-Gly-3M3SH

A model of ShPepV was generated on SWISS-MODEL (47–49) using the known structure of SaPepV (PDB: 3KI9) (29) as the template due to its high sequence identity at 80%. The GMQE (Global Model Quality Estimation) score was determined to be 0.91 which suggests high confidence of the generated tertiary structure. The QMEANDisCo Global value was calculated to be 0.86 ± 0.05 which suggests high per-residue quality for the model against the template. To suggest the likely binding cavity of ShPepV for Cys-Gly-3M3SH, the structure of *Lactobacillus delbrueckii* PepV bound to an Asp-Ala phosphinic acid inhibitor (PDB: 1LFW) (24) was aligned with the model of ShPepV. Cys-Gly-3M3SH was fitted into the position of the inhibitor by manual rigid-body modelling in the highly conserved binding sites starting with the carboxy terminus of the peptide, followed by the peptide bond, then the terminal amino group. This positioned the 3M3SH moiety into an adjacent cavity and a suggested fit of this bulky group was modelled, avoiding clashes with neighbouring residues. In order to estimate the goodness-of-fit between the modelled Cys-Gly-3M3SH and ShPepV, the buried surface area between the protein and ligand was calculated using ChimeraX 1.8 (33). Solvent accessible surfaces were determined using a probe radius of 1.4 Å which is used to approximate a water molecule. The binding site was defined as ShPepV protein residues within 5 Å from the ligand. This yielded the following results; buried surface area (area of contact between ligand and protein) = 323.52 Å², ligand surface

area = 509.81 Å², surface area of protein residues within 5 Å from the ligand = 1592.9 Å², combined surface area of the ligand and nearby protein residues which are 5 Å away = 1455.7 Å². Approximately 63% (323.52/509.81 Å²) of Cys-Gly-3M3SH surface is in contact with ShPepV and there is a decrease in the total surface area (1592.9 Å² +509.81 Å² > 1455.7 Å²) suggesting there is significant/notable interaction.

2.8 Data availability

All the data that support the finding in this study are reported either within the article or in the supporting information.

2.9 Conflict of interest

The authors declare that B. M. is a current employee of Unilever, whereas A. G. J. was an employee of Unilever when this study was carried out. All other authors declare that they have no conflicts of interest with the contents of this article.

2.10 Acknowledgements

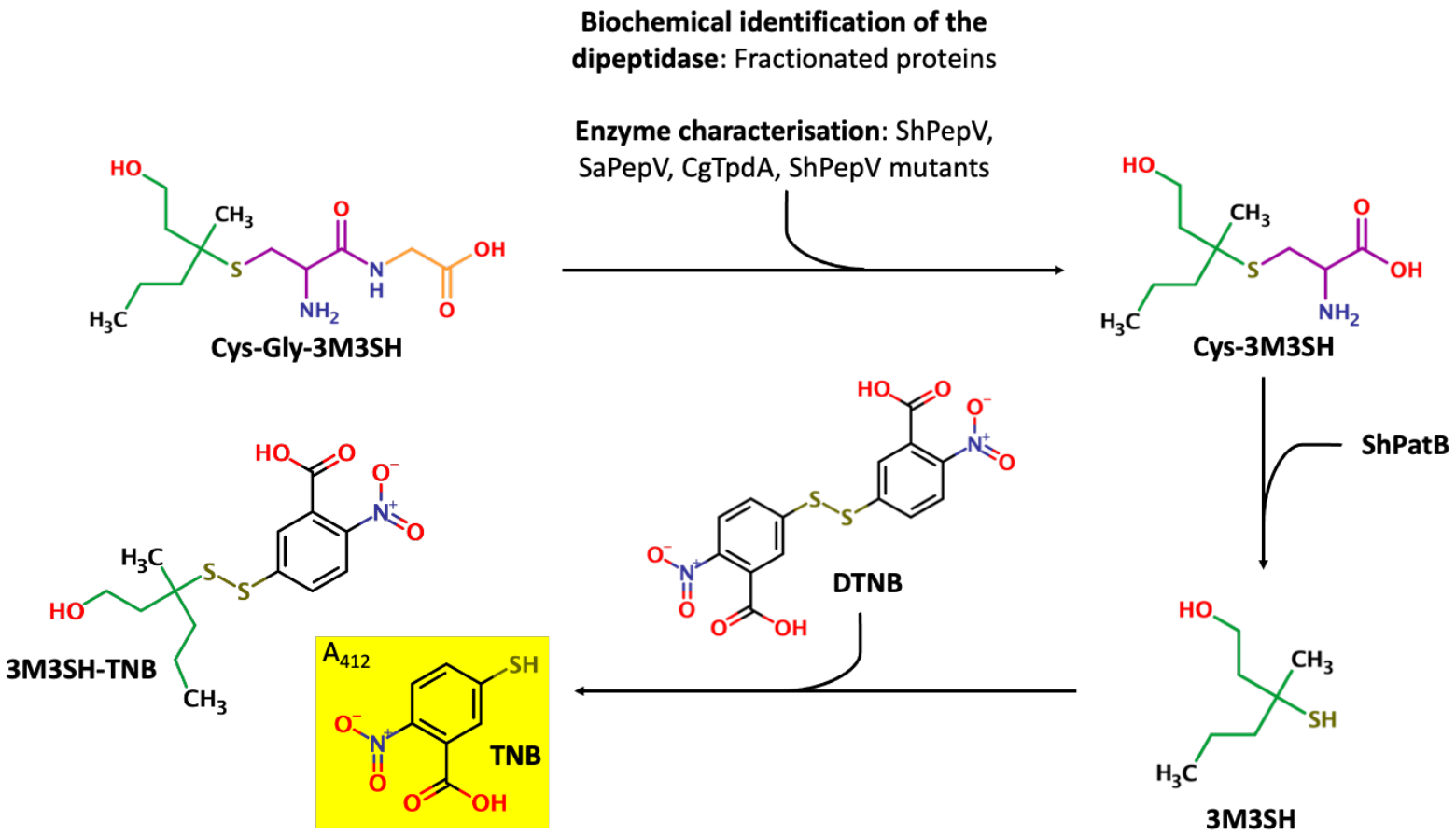
This work was supported by the Biotechnology and Biological Sciences Research Council grant BB/W510531/1 supporting R. H. and grant BB/N006615/1 that supported R. H. and M. R. Molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from the National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases.

2.11 Author contributions

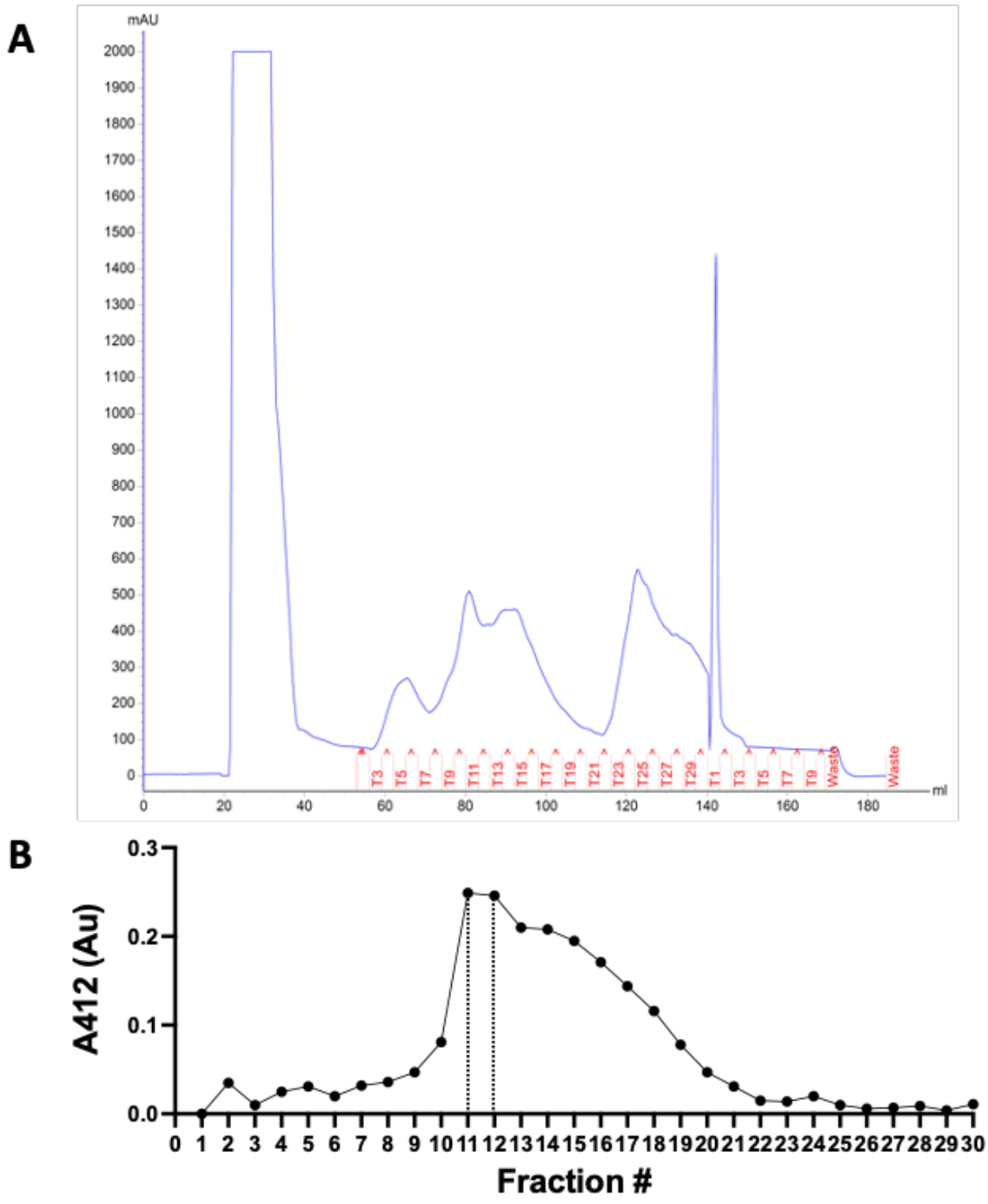
R. H., M. R., A. G. J., A. J. W., and G. H. T. conceptualization; R. H., M. R., A. J. W., and B. M. validation; R. H., B. K.-W., and G. H. T. formal analysis; R. H. investigation; R. H. data curation; R. H. writing–original draft; R. H., B. K.-W., M. R., A. J. W., B. M., and G. H. T. writing–review & editing; R. H. and B. K.-W. visualization; A. J. W., B. M.,

and G. H. T. supervision; R. H., B. M., and G. H. T. project administration; A. G. J. and G. H. T. funding acquisition.

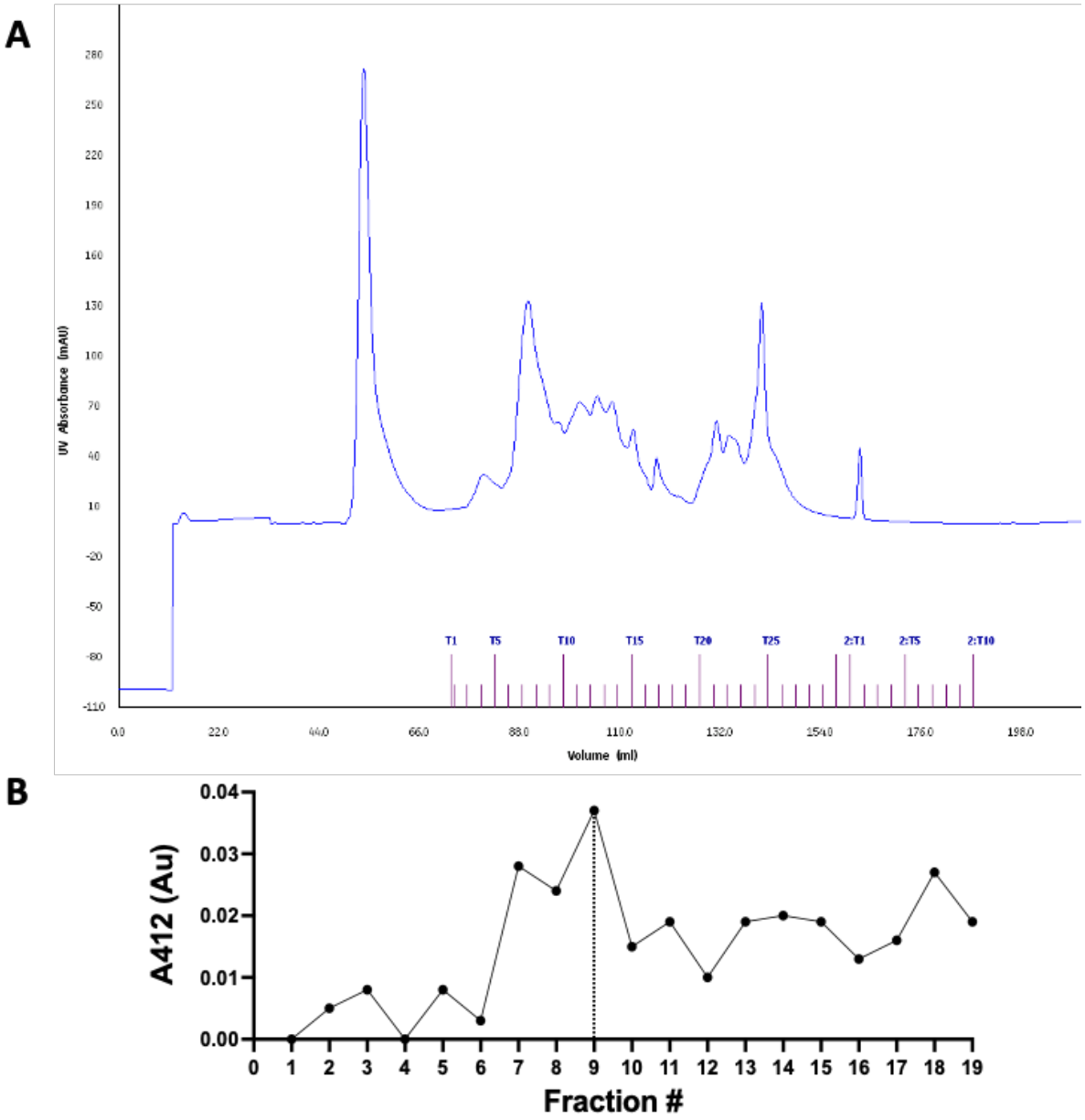
2.12 Supporting information



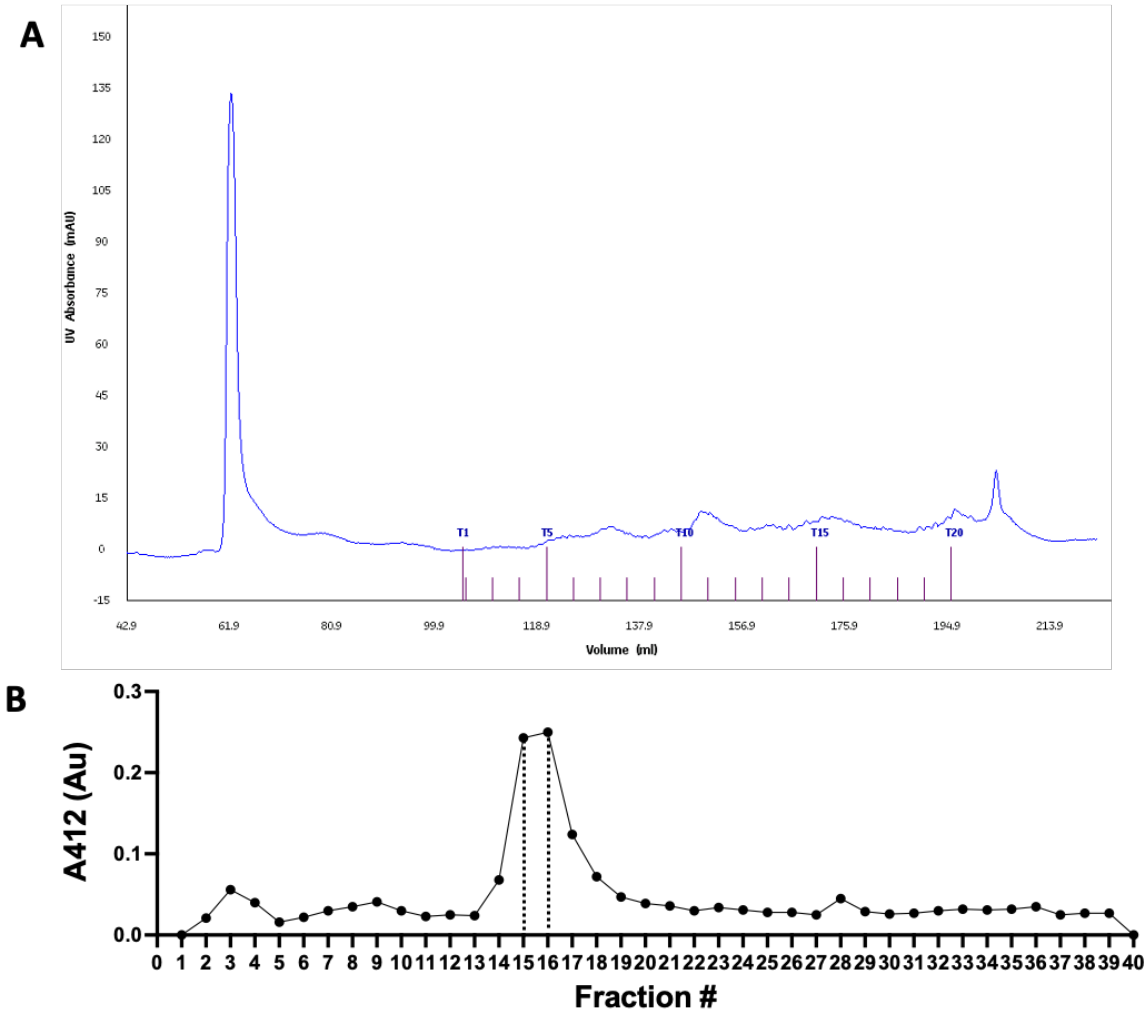
Supplementary Data S1: The coupled Cys-Gly-3M3SH DTNB assay. To test the dipeptidase activity, Cys-Gly-3M3SH was either incubated with the fractionated proteins from the various steps of the biochemical identification of the dipeptidase or subsequent recombinantly expressed proteins. ShPatB was added to process any resultant Cys-3M3SH to form 3M3SH which in turn interacts with labelling molecule 5,5′-dithiobis-(2-nitrobenzoic acid) (DTNB). This results in the formation of TNB which absorbs at 412 nm.



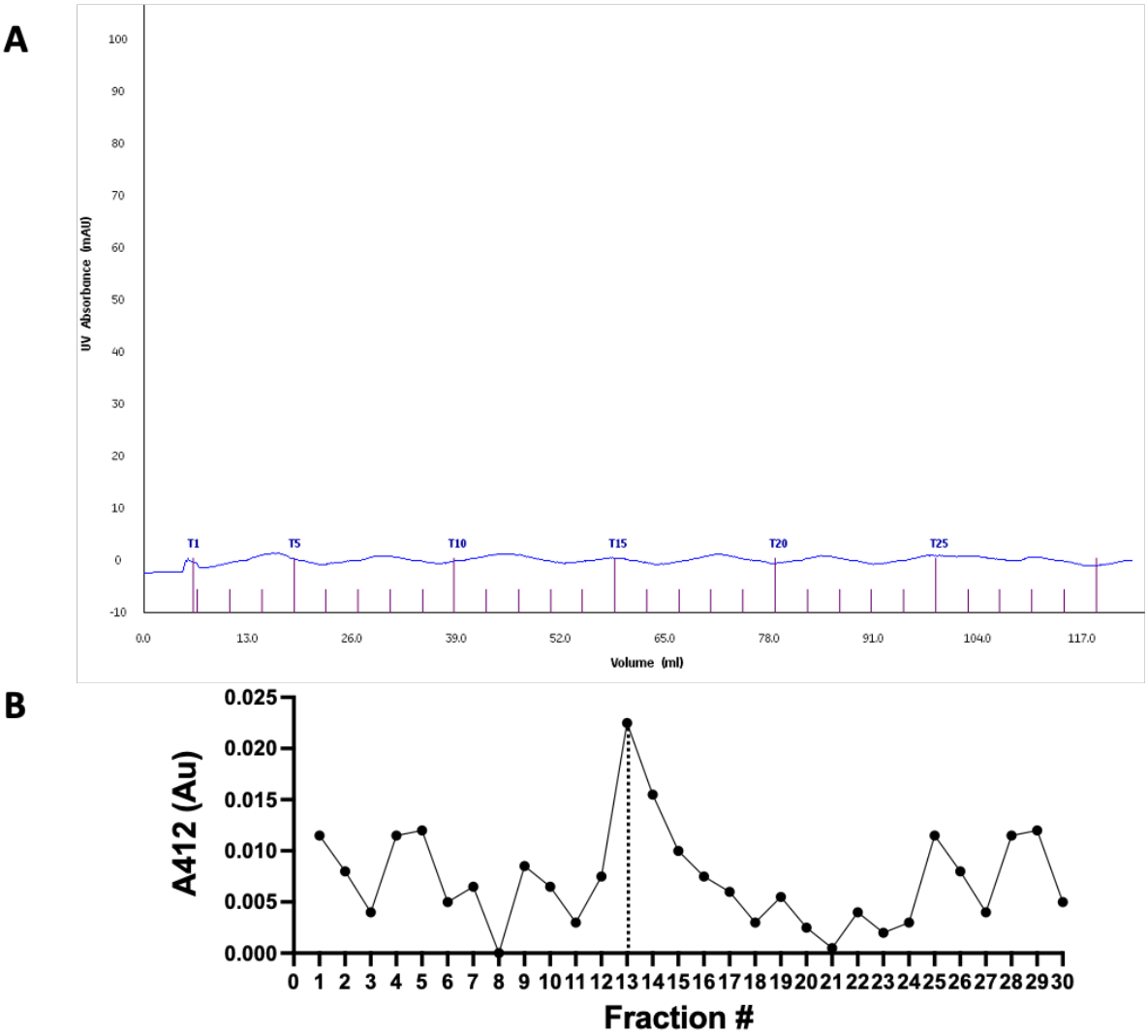
Supplementary Data S2. (A) Absorbance at 280 nm measured during anion exchange chromatography with the Capto Q resin (Cytiva). Elution fractions are labelled in red on the X-axis. (B) Elution fractions were tested for Cys-Gly-3M3SH peptidase activity using a coupled DTNB assay. Activity was indirectly measured through the spectrophotometric detection of TNB at 412 nm. Fractions 11 and 12 (black dotted lines) proceeded to the next purification step.



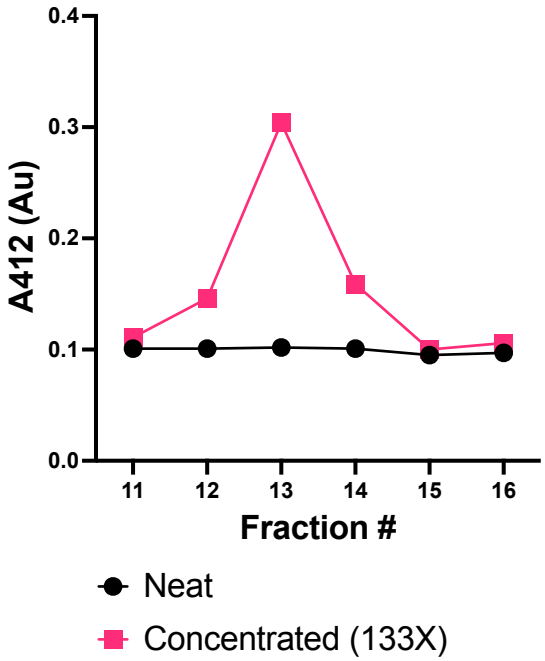
Supplementary Data S3. (A) Absorbance at 280 nm measured during hydrophobic interactions chromatography with the Phenyl HP resin (Cytiva). Elution fractions are labelled in red on the X-axis. (B) Elution fractions were tested for Cys-Gly-3M3SH peptidase activity using a coupled DTNB assay. Activity was indirectly measured through the spectrophotometric detection of TNB at 412 nm. Fraction 9 (black dotted line) proceeded to the next purification step.



Supplementary Data S4. (A) Absorbance at 280 nm measured during anion exchange chromatography with the Capto DEAE resin (Cytiva). Elution fractions are labelled in red on the X-axis. (B) Elution fractions were tested for Cys-Gly-3M3SH peptidase activity using a coupled DTNB assay. Activity was indirectly measured through the spectrophotometric detection of TNB at 412 nm. Fractions 15 and 16 (black dotted lines) proceeded to the next purification step.



Supplementary Data S5. (A) Absorbance at 280 nm measured during size exclusion chromatography with the Sephacryl S200 resin (Cytiva). Elution fractions are labelled in red on the X-axis. (B) Elution fractions were tested for Cys-Gly-3M3SH peptidase activity using a coupled DTNB assay. Activity was indirectly measured through the spectrophotometric detection of TNB at 412 nm. Fraction 13 (black dotted lines) was analysed for the identification of the protein.



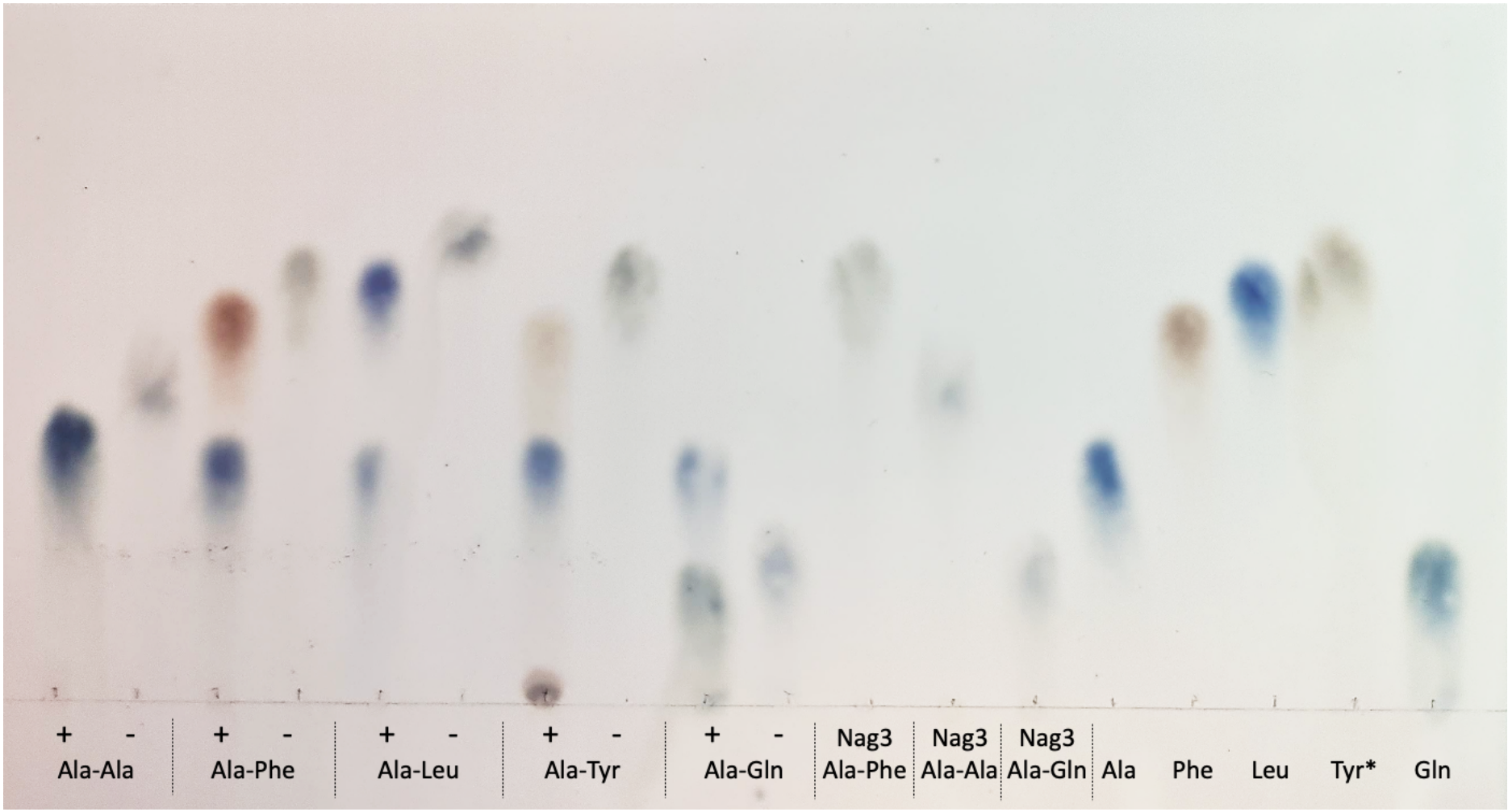
Supplementary Data S6. Cys-Gly-3M3SH dipeptidase activity is detected in SEC elution fractions 12-14 with fraction 13 having the highest signal. Samples from the elution fractions were tested neat (black dots) or concentrated by 133X (pink squares) in the coupled DTNB assay. Proteins were incubated with 1 mM Cys-Gly-3M3SH and 1 μM ShPatB for 30 minutes in the presence of DTNB. Dipeptidase activity was indirectly measured through the spectrophotometric detection of TNB at 412 nm.

Protein: Dipeptidase PepV
Organism: *S. hominis*
UniProt ID: A0A657M1L9
Score: 288

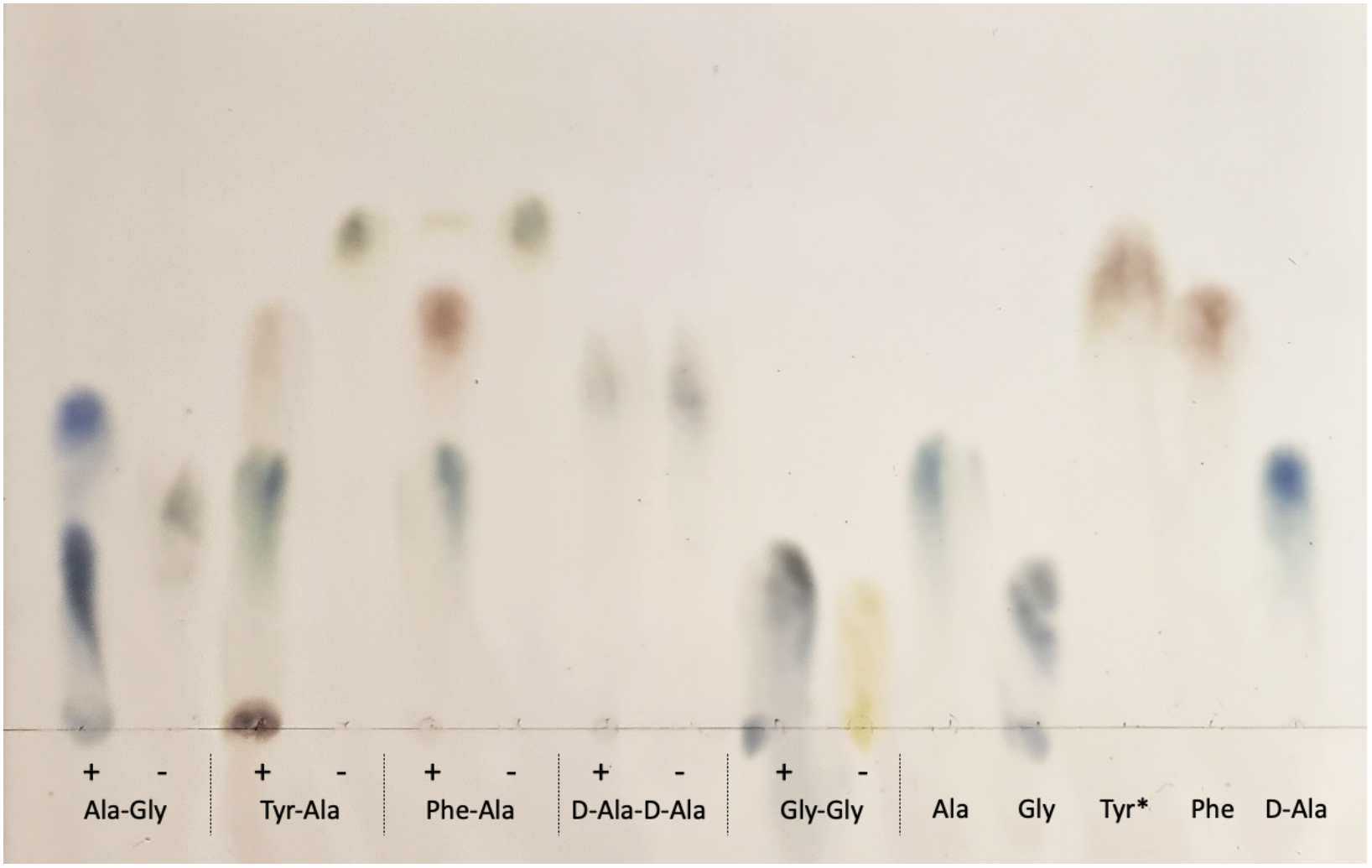
```
  1 MWKEKVQEYE SQIIEDLNGL LSIKSVRDDN QASEEAPVGP
 41 GPREALDYMY KIAQRDGFST HDVDHIAGRI EAGQGEDVLG
 81 ILCHVDVVPA GDGWDSDPFK PVVTDDAIVA RGTLDDKGPT
121 IAAYYAVKIL NEMNVDWKKR IHIIIGTDEE SDWKCTERYF
161 ETEEMPALGF APDAEFPAIH GEKGITTFDL VQNVSSPHTD
201 HSEATLLSFE SGQRYNMVPD HAKAHILVEK NMTDVVQHFE
241 YYLDQQQLEG EAVVDNGELV LTVEGKAVHG MDPSLGMNAG
281 LYLIDFLSTL NLNQTAKEFA EFSNRYLHES HFGEKMGMKF
321 HTDVMGDVTT NVGIITYDKE KGGKFGINLR YPKGFEFETA
361 LERFKNEIKS QGFEIELGKV QQPHFVDKND PFVQKLVTAY
401 RNQTGDMTEP YTIGGGTYAR NLDKGVAFGA MFEDSEDLMH
441 QKNEYITKKQ LFNATSIYLE AIYSLCVEG
```

#	Sequence position	Experimental m/z	Molecular mass derived from m/z	Expected molecular mass	Score	Expectation value	Sequence	Protein sequence percentage coverage (%)
1	25 - 43	1981.0874	1980.0801	1979.9297	93	2.8e-08	KSVRDDNQASEEAPVGPGPRE	4.47
2	380 - 395	1926.1489	1925.1416	1924.9796	95	1.4e-09	KVQQPHFVDKNDPFVQKL	3.84
3	402 - 420	2032.0428	2031.0355	2030.9004	99	1.3e-08	RNQTGDMTEPYTIGGGTYARN	4.47
4	402 - 420	2048.0278	2047.0205	2046.8953	75	9.6e-08	RNQTGDMTEPYTIGGGTYARN + Oxidation (M)	4.47

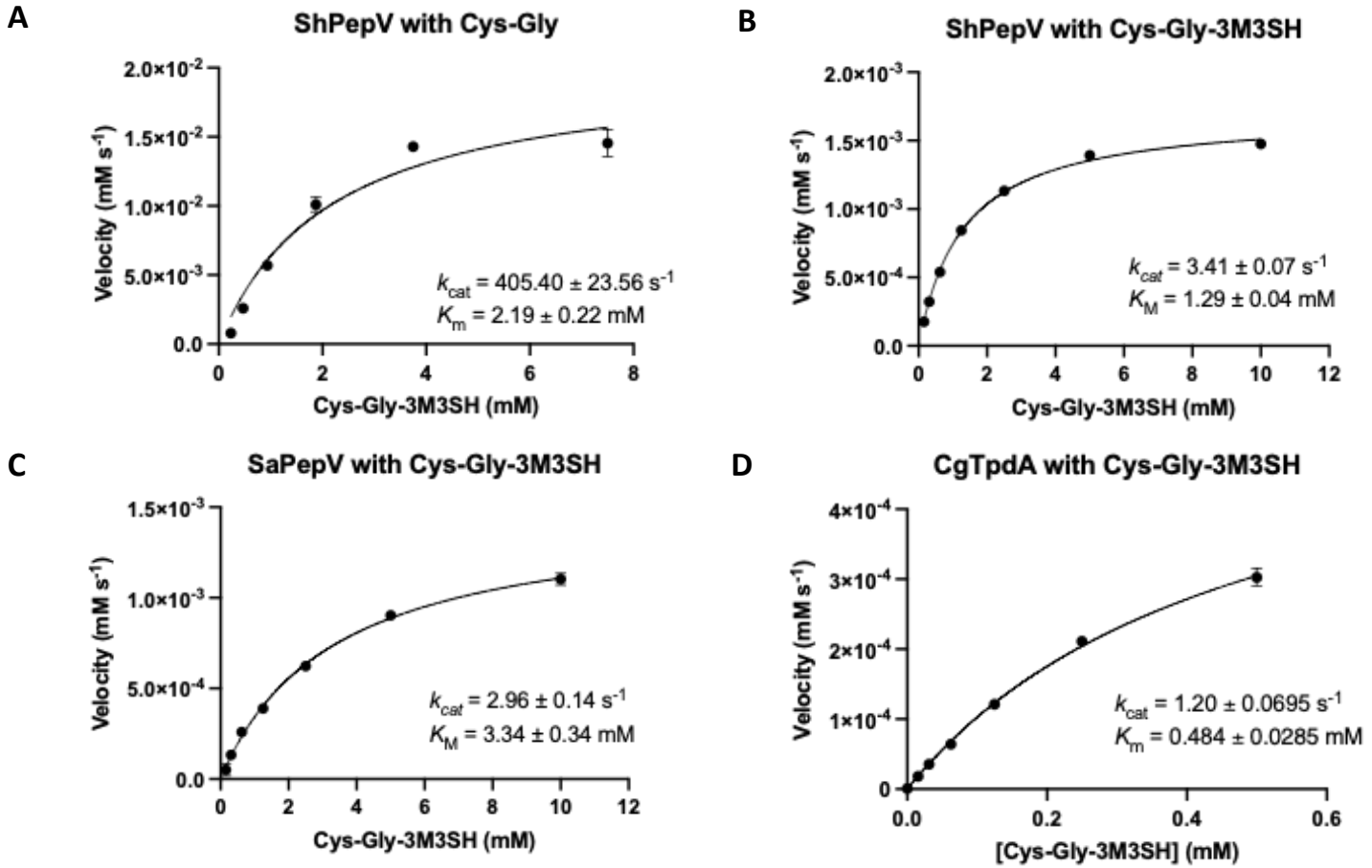
Supplementary Data S7. Identification of PepV by peptide mass fingerprinting using trypsin digest and MALDI-MS/MS. The total protein score is derived from the ions scores for each distinct sequence. No other proteins were matched with the experimentally isolated protein fragments.



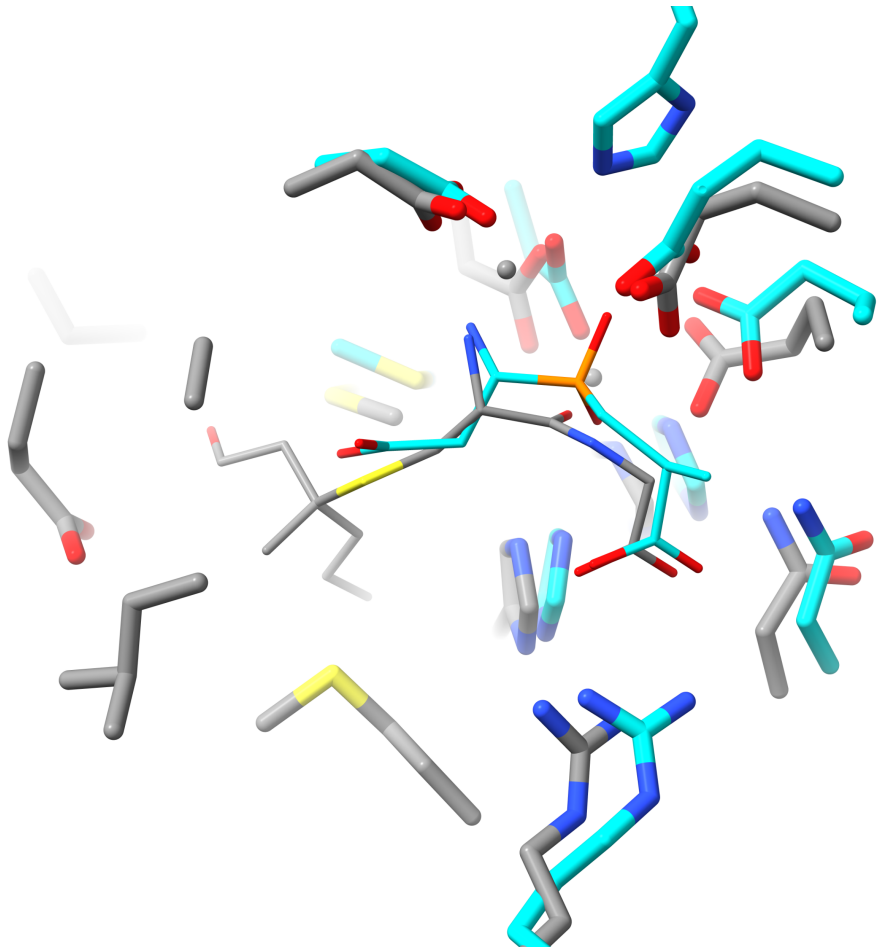
Supplementary Data S9. TLC assay ShPepV against a range of dipeptides. Peptides were incubated with (+) or without (-) PepV or the control enzyme Nag3 as indicated. Relevant peptide and amino acid controls were included. *L-Tyrosine was solubilised in 1M HCl. All other substrates and controls were solubilised in dH2O.



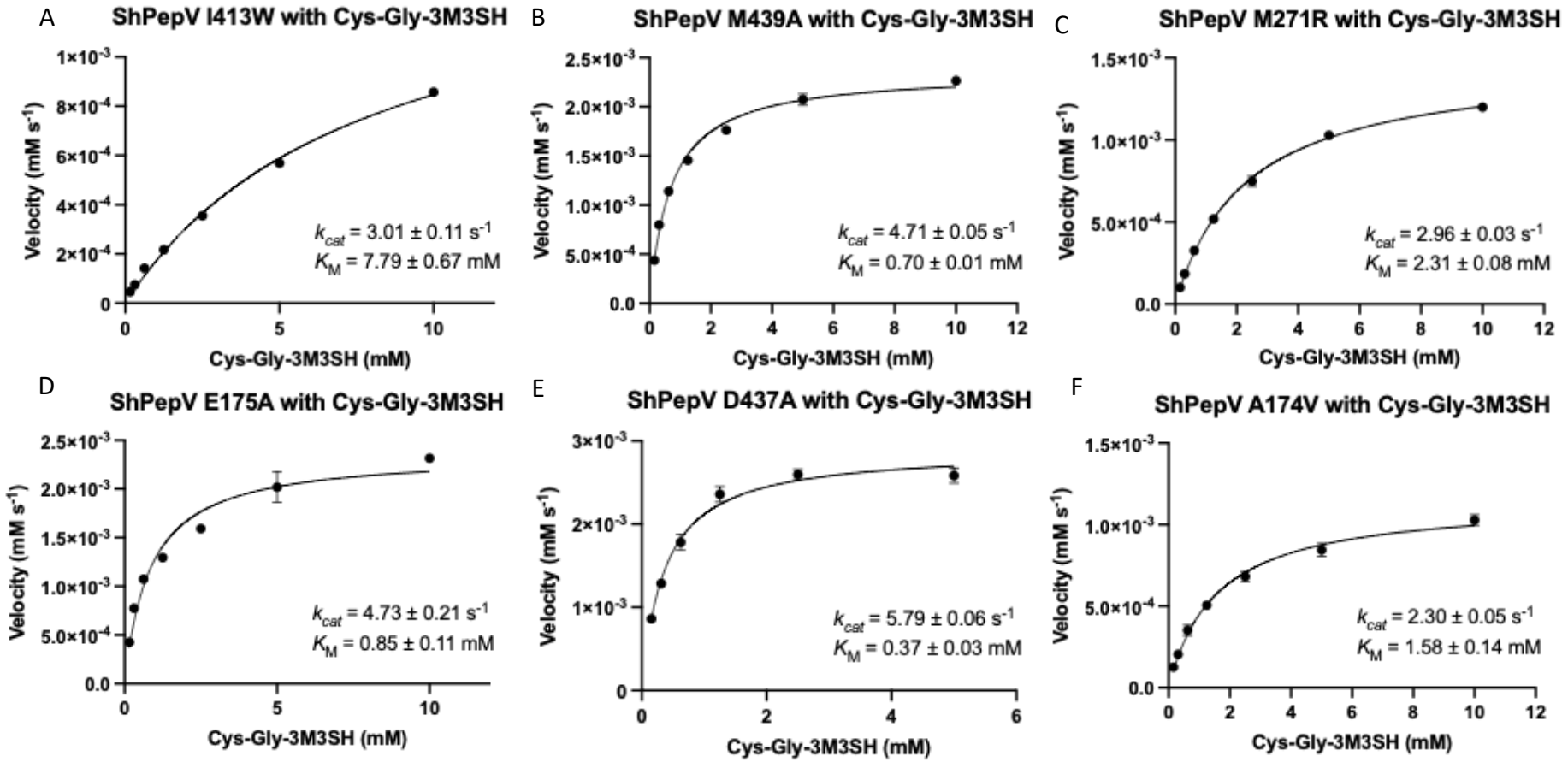
Supplementary Data S10. TLC assay ShPepV against a range of dipeptides. Peptides were incubated with (+) or without (-) PepV. Relevant peptide and amino acid controls were included. *L-Tyrosine was solubilised in 1M HCl. All other substrates and controls were solubilised in dH_2O.



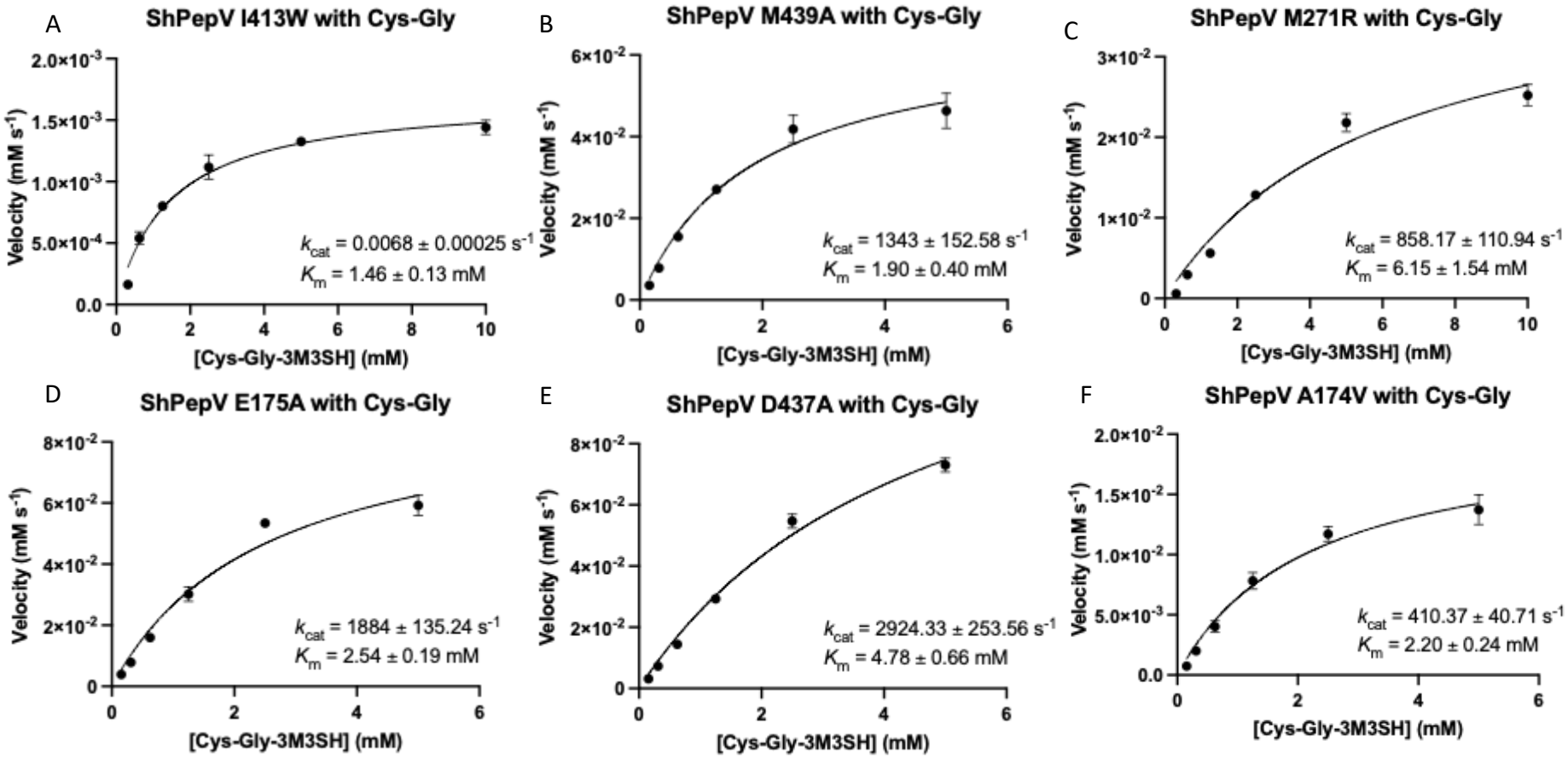
Supplementary Data S11. Michaelis-Menten kinetics of ShPepV with (A) Cys-Gly, (B) Cys-Gly-3M3SH and (C) SaPepV and (D) CgTpdA with Cys-Gly-3M3SH



Supplementary Data S12. Overlay of the binding sites of ShPepV with Cys-Gly-3M3SH (Grey) and LdPepV with an Asp-Ala phosphinic acid inhibitor (Blue). Ligands are represented in thin cylinder form while amino acid residues are represented in thick cylinder form. Metals are represented by grey balls.



Supplementary Data S13. Michaelis-Menten kinetics of ShPepV mutants (A) I413W, (B) M439A, (C) M271R, (D) E175A, (E) D437A and (F) A174V with Cys-Gly-3M3SH.



Supplementary Data S14. Michaelis-Menten kinetics of ShPepV mutants (A) I413W, (B) M439A, (C) M271R, (D) E175A, (E) D437A and (F) A174V with Cys-Gly.

Supplementary Table 1. Bacterial strains used in this study.

Strains	Purpose	Source
S. hominis B10	Skin isolate. Source of ShPepV	University of Liverpool
S. aureus		
RN4220	Restriction-deficient shuttle strain	Nair *et al.*, 2011
USA300 JE2	Source of SaPepV	Kennedy *et al.*, 2008
E. coli		
XL1-Blue	Cloning strain	Agilent
MC1061	Expression strain	Casadaban and Cohen, 1980
C. glaucum DSM 30827	Source of CgTpdA	DSMZ

Supplementary Table 2. Plasmids used in this study.

Plasmids	Purpose	Source
pBADcLIC (pBcL)	*E. coli* expression vector	Geertsma *et al.*, 2008
pBcL::Sh*pepV*	Expression of ShPepV	This study
pBcL::Sa*pepV*	Expression of SaPepV	This study
pBcL::Cg*tpdA*	Expression of CgTpdA	This study
pBcL::Sh*pepV* E175A	Expression of ShPepV E175A	This study
pBcL::Sh*pepV* D437A	Expression of ShPepV D437A	This study
pBcL::Sh*pepV* D437E	Expression of ShPepV D437E	This study
pBcL::Sh*pepV* M271R	Expression of ShPepV M271R	This study
pBcL::Sh*pepV* A174V	Expression of ShPepV A174V	This study
pBcL::Sh*pepV* D173A	Expression of ShPepV D173A	This study
pBcL::Sh*pepV* I413W	Expression of ShPepV I413W	This study

Supplementary Table 3. Primers used in this study

Proteins	Primers
ShPepV	TAACAGGAGGAATTAACCATGTGGAAAGAAAAAGTACAAGAATATGAAAG
	TTGGAAGTATAAATTTTCTCCCTCCACACATAGTGAATAAATTG
SaPepV	TAACAGGAGGAATTAACCATGTGGAAAGAAAAAGTTCAACAATAC
	TTGGAAGTATAAATTTTCTTCCTCCACGCATAATGAATAAATTG
CgTpdA	TAACAGGAGGAATTAACCATGAGTAACGACAACGCAGTAAAC
	TTGGAAGTATAAATTTTCTTTTCCGTAGGTGAGCAGGAATG
ShPepV Mutants	**Primers**
D173A	TTTTGCACCAGCGGCTGAATTCC
	CCTAAAGCAGGCATTTCTTC
A174V	TGCACCAGATGTTGAATTCCCTG
	AAACCTAAAGCAGGCATTTC
I413W	ACCTTATACATGGGGTGGAGGCACTTATG
	TCTGTCATATCACCCGTTTG
M439A	AGAGGATTTAGCGCATCAAAAGAATGAATATATTAC
	GAATCTTCAAACATAGCTC
E175A	ACCAGATGCTGCGTTCCCTGCTATTC
	GCAAAACCTAAAGCAGGC
D437A	AGATTCAGAGGCGTTAATGCATCAAAAG
	TCAAACATAGCTCCAAATG
D437E	AGATTCAGAGGAATTAATGCATC
	TCAAACATAGCTCCAAATG
M271R	CGTCCATGGTCGTGATCCTTCACTTGG
	GCTTTCCCTTCAACAGTTAATAC

Supplementary Table 4. List of dipeptides and controls used for the characterisation of the peptidase activity of ShPepV. All non-glycine amino acids are in the L- form unless indicated. Activity of PepV against the di-/tripeptides are also included.

Dipeptides	PepV Activity?	Tripeptides	PepV Activity?
Ala-Ala	Active	Ala-Ala-Ala	Not active
Ala-Phe	Active		
Ala-Leu	Active	**Amino acids**	
Ala-Tyr	Active	Ala	
Ala-Gln	Active	Phe	
Ala-Gly	Active	Leu	
Tyr-Ala	Active	Tyr	
Phe-Ala	Active	Gln	
D-Ala-D-Ala	Not active	Gly	
Gly-Gly	Active	D-Ala	
Gly-His	Active	His	
Cys-Gly	Active	Cys	
Leu-Gly	Active		

Supplementary Table 5. Kinetics of ShPepV mutants with Cys-Gly determined using a Cd-ninhydrin method. The fold change in catalytic efficiencies ($(k_{cat}/K_M)_{mut}/(k_{cat}/K_M)_{WT}$) and differences in transition state binding energies ($\Delta\Delta G^{\neq}$) of the mutants against that of the wildtype (WT) protein were also calculated. (N.D.: not detectable)

	Purpose	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (s^{-1} mM^{-1})	$(k_{cat}/K_M)_{mut}/(k_{cat}/K_M)_{WT}$	$\Delta\Delta G^{\neq}$
WT	Native	405.40 ± 23.56	2.19 ± 0.22	186.00 ± 21.76	-	-
D173A	Metal coordinating	N.D.	N.D.	N.D.	-	-
A174V	Smaller cavity (Buried)	410.37 ± 40.71	2.20 ± 0.24	186.89 ± 27.33	1.00	0
I413W	Smaller cavity (Buried)	0.0068 ± 0.00025	1.46 ± 0.13	0.0046 ± 0.00044	0.000025	-27.33
M439A	Bigger cavity (Buried)	1343 ± 152.58	1.90 ± 0.40	714.67 ± 170.80	3.84	3.50
D437A	Bigger cavity (Surface)	2924.33 ± 253.56	4.78 ± 0.66	614.91 ± 99.91	3.31	3.09
E175A	Bigger cavity (Surface)	1884 ± 135.24	2.54 ± 0.19	743.49 ± 77.73	4.00	3.58
M271R	Smaller cavity (Surface)	858.17 ± 110.94	6.15 ± 1.54	142.02 ± 40.03	0.76	-0.71

2.13 References

1. James, A. G. (2020) The Axillary Microbiome and its Relationship with Underarm Odor. *Skin Microbiome Handbook: From Basic Research to Product Development*. 10.1002/9781119593058.CH5
2. Shelley, W. B., Hurley, H. J., and Nichols, A. C. (1953) Axillary odor; experimental study of the role of bacteria, apocrine sweat, and deodorants. *AMA Arch Derm Syphilol*. **68**, 430–46
3. Natsch, A., and Emter, R. (2020) The specific biochemistry of human axilla odour formation viewed in an evolutionary context. *Philosophical Transactions of the Royal Society B*. 10.1098/RSTB.2019.0269
4. Hodge, B. D., and Brodell, R. T. (2018) *Anatomy, Skin, Sweat Glands*, [online] http://europepmc.org/books/NBK482278 (Accessed October 31, 2022)
5. Troccaz, M., Starkenmann, C., Niclass, Y., van de Waal, M., Clark, A. J. J., Van De Waal, M., and Clark, A. J. J. (2004) 3-Methyl-3-sulfanylhexan-1-ol as a Major Descriptor for the Human Axilla-Sweat Odour Profile. *Chem Biodivers*. **1**, 1022–1035
6. Zeng, X. nong, Leyden, J. J., Lawley, H. J., Sawano, K., Nohara, I., and Preti, G. (1991) Analysis of characteristic odors from human male axillae. *Journal of Chemical Ecology 1991 17:7*. **17**, 1469–1492
7. Natsch, A., Gfeller, H., Gygax, P., Schmid, J., and Acuna, G. (2003) A Specific Bacterial Aminoacylase Cleaves Odorant Precursors Secreted in the Human Axilla *. *Journal of Biological Chemistry*. **278**, 5718–5727
8. Barzantny, H., Brune, I., and Tauch, A. (2012) Molecular basis of human body odour formation: insights deduced from corynebacterial genome sequences. *Int J Cosmet Sci*. **34**, 2–11
9. James, A. G., Austin, C. J., Cox, D. S., Taylor, D., and Calvert, R. (2013) Microbiological and biochemical origins of human axillary odour. *FEMS Microbiol Ecol*. **83**, 527–40
10. Fredrich, E., Barzantny, H., Brune, I., and Tauch, A. (2013) Daily battle against body odor: towards the activity of the axillary microbiota. *Trends Microbiol*. **21**, 305–312
11. Grice, E. A., Kong, H. H., Conlan, S., Deming, C. B., Davis, J., Young, A. C., Bouffard, G. G., Blakesley, R. W., Murray, P. R., Green, E. D., Turner, M. L., and Segre, J. A. (2009) Topographical and temporal diversity of the human skin microbiome. *Science (1979)*. **324**, 1190–1192

12. Grice, E. A., and Segre, J. A. (2011) The skin microbiome. *Nature Reviews Microbiology 2011 9:4*. **9**, 244–253
13. Troccaz, M., Gaïa, N., Beccucci, S., Schrenzel, J., Cayeux, I., Starkenmann, C., and Lazarevic, V. (2015) Mapping axillary microbiota responsible for body odours using a culture-independent approach. *Microbiome*. 10.1186/s40168-014-0064-3
14. Taylor, D., Daulby, A., Grimshaw, S., James, G., Mercer, J., and Vaziri, S. (2003) Characterization of the microflora of the human axilla. *Int J Cosmet Sci*. **25**, 137–145
15. Minhas, G. S., Bawdon, D., Herman, R., Rudden, M., Stone, A. P., Gordon James, A., Thomas, G. H., and Newstead, S. Structural basis of malodour precursor transport in the human axilla. 10.7554/eLife.34995.001
16. Rudden, M., Herman, R., Rose, M., Bawdon, D., Cox, D. S., Dodson, E., Holden, M. T. G., Wilkinson, A. J., James, A. G., and Thomas, G. H. (2020) The molecular basis of thioalcohol production in human body odour. *Sci Rep*. **10**, 12500
17. Emter, R., and Natsch, A. (2008) The sequential action of a dipeptidase and a beta-lyase is required for the release of the human body odorant 3-methyl-3-sulfanylhexan-1-ol from a secreted Cys-Gly-(S) conjugate by Corynebacteria. *J Biol Chem*. **283**, 20645–52
18. Natsch, A., Schmid, J., and Flachsmann, F. (2004) Identification of Odoriferous Sulfanylalkanols in Human Axilla Secretions and Their Formation through Cleavage of Cysteine Precursors by a C□S Lyase Isolated from Axilla bacteria. *Chem Biodivers*. **1**, 1058–1072
19. Natsch, A., Gfeller, H., Gygax, P., and Schmid, J. (2005) Isolation of a bacterial enzyme releasing axillary malodor and its use as a screening target for novel deodorant formulations1. *Int J Cosmet Sci*. **27**, 115–122
20. Bawdon, D., Cox, D. S., Ashford, D., James, A. G., and Thomas, G. H. (2015) Identification of axillary Staphylococcus sp. involved in the production of the malodorous thioalcohol 3-methyl-3-sufanylhexan-1-ol. *FEMS Microbiol Lett*. **362**, 1–10
21. Rawlings, N. D., Barrett, A. J., Thomas, P. D., Huang, X., Bateman, A., and Finn, R. D. (2018) The MEROPS database of proteolytic enzymes, their substrates and inhibitors in 2017 and a comparison with peptidases in the PANTHER database. *Nucleic Acids Res*. **46**, D624–D632
22. Hermsdorf, C. L. (1978) Tripeptide-specific aminopeptidase from Escherichia coli AJ005. *Biochemistry*. **17**, 3370–3376

23. Sanz, Y., Mulholland, F., and Toldrá, F. (1998) Purification and Characterization of a Tripeptidase from Lactobacillus sake. *J Agric Food Chem*. **46**, 349–353
24. Jozic, D., Bourenkow, G., Bartunik, H., Scholze, H., Dive, V., Henrich, B., Huber, R., Bode, W., and Maskos, K. (2002) Crystal Structure of the Dinuclear Zinc Aminopeptidase PepV from Lactobacillus delbrueckii Unravels Its Preference for Dipeptides. *Structure*. **10**, 1097–1106
25. Vongerichten, K. F., Klein, J. R., Matern, H., and Plapp, R. (1994) Cloning and nucleotide sequence analysis of pepV, a carnosinase gene from Lactobacillus delbrueckii subsp. lactis DSM 7290, and partial characterization of the enzyme. *Microbiology (N Y)*. **140**, 2591–2600
26. Chu, L., Lai, Y., Xu, X., Eddy, S., Yang, S., Song, L., and Kolodrubetz, D. (2008) A 52-kDa leucyl aminopeptidase from treponema denticola is a cysteinylglycinase that mediates the second step of glutathione metabolism. *J Biol Chem*. **283**, 19351–19358
27. Cappiello, M., Spinelli, M., Mormino, M., Renzone, G., Scaloni, A., Moschini, R., Balestri, F., Mura, U., and Del-Corso, A. (2016) Purification and characterization of a Cys-Gly hydrolase from the gastropod mollusk, Patella caerulea. *J Enzyme Inhib Med Chem*. **31**, 1560–1565
28. Cappiello, M., Lazzarotti, A., Buono, F., Scaloni, A., D'Ambrosio, C., Amodeo, P., Méndez, B. L., Pelosi, P., Del Corso, A., and Mura, U. (2004) New role for leucyl aminopeptidase in glutathione turnover. *Biochemical Journal*. **378**, 35
29. Girish, T. S., and Gopal, B. (2010) Crystal structure of Staphylococcus aureus metallopeptidase (Sapep) reveals large domain motions between the manganese-bound and apo-states. *Journal of Biological Chemistry*. **285**, 29406–29415
30. Staub, I., and Sieber, S. A. (2009) β-Lactam probes as selective chemical-proteomic tools for the identification and functional characterization of resistance associated enzymes in MRSA. *J Am Chem Soc*. **131**, 6271–6276
31. Moissl-Eichinger, C., Rettberg, P., and Pukall, R. (2012) The First Collection of Spacecraft-Associated Microorganisms: A Public Source for Extremotolerant Microorganisms from Spacecraft Assembly Clean Rooms. *https://home.liebertpub.com/ast*. **12**, 1024–1034
32. Doi, E., Shibata, D., and Matoba, T. (1981) Modified colorimetric ninhydrin methods for peptidase assay. *Anal Biochem*. **118**, 173–184

33. Pettersen, E. F., Goddard, T. D., Huang, C. C., Meng, E. C., Couch, G. S., Croll, T. I., Morris, J. H., and Ferrin, T. E. (2021) UCSF ChimeraX: Structure visualization for researchers, educators, and developers. *Protein Science*. **30**, 70–82
34. Martin, A., Saathoff, M., Kuhn, F., Max, H., Terstegen, L., and Natsch, A. (2010) A functional ABCC11 allele is essential in the biochemical formation of human axillary odor. *Journal of Investigative Dermatology*. **130**, 529–540
35. Prokop-Prigge, K. A., Greene, K., Varallo, L., Wysocki, C. J., and Preti, G. (2016) The Effect of Ethnicity on Human Axillary Odorant Production. *J Chem Ecol*. **42**, 33–39
36. Baumann, T., Bergmann, S., Schmidt-Rose, T., Max, H., Martin, A., Enthaler, B., Terstegen, L., Schweiger, D., Kalbacher, H., Wenck, H., Jedlitschky, G., and Jovanovic, Z. (2014) Glutathione-conjugated sulfanylalkanols are substrates for ABCC11 and γ-glutamyl transferase 1: a potential new pathway for the formation of odorant precursors in the apocrine sweat gland. *Exp Dermatol*. **23**, 247
37. Harker, M., Carvell, A. M., Marti, V. P. J., Riazanskaia, S., Kelso, H., Taylor, D., Grimshaw, S., Arnold, D. S., Zillmer, R., Shaw, J., Kirk, J. M., Alcasid, Z. M., Gonzales-Tanon, S., Chan, G. P., Rosing, E. A. E., and Smith, A. M. (2014) Functional characterisation of a SNP in the ABCC11 allele—Effects on axillary skin metabolism, odour generation and associated behaviours. *J Dermatol Sci*. **73**, 23–30
38. Halsey, C. R., Lei, S., Wax, J. K., Lehman, M. K., Nuxoll, A. S., Steinke, L., Sadykov, M., Powers, R., and Fey, P. D. (2017) Amino acid catabolism in Staphylococcus aureus and the function of carbon catabolite repression. *mBio*. 10.1128/MBIO.01434-16/SUPPL_FILE/MBO001173181ST2.DOCX
39. Berger-B, B., Barberis-Maino, L., Str, A., and Kayser, F. H. (1989) FemA, a host-mediated factor essential for methicillin resistance in Staphylococcus aureus: Molecular cloning and characterization. **219**, 263–269
40. Rohrer, S., Ehlert, K., Tschierske, M., Labischinski, H., and Berger-Bachi, B. (1999) The essential Staphylococcus aureus gene fmhB is involved in the first step of peptidoglycan pentaglycine interpeptide formation. *Proceedings of the National Academy of Sciences*. **96**, 9351–9356
41. Zeden, M. S., Burke, Ó., Vallely, M., Fingleton, C., and O'Gara, J. P. (2021) Exploring amino acid and peptide transporters as therapeutic targets to attenuate virulence and antibiotic resistance in Staphylococcus aureus. *PLoS Pathog*. **17**, e1009093

42. Halsey, C. R., Lei, S., Wax, J. K., Lehman, M. K., Nuxoll, A. S., Steinke, L., Sadykov, M., Powers, R., and Fey, P. D. (2017) Amino acid catabolism in Staphylococcus aureus and the function of carbon catabolite repression. *mBio*. 10.1128/MBIO.01434-16/SUPPL_FILE/MBO001173181ST2.DOCX
43. Mei, J.-M., Nourbakhsh, F., Ford, C. W., and Holden, D. W. Identification of Staphylococcus aureus virulence genes in a murine model of bacteraemia using signature-tagged mutagenesis. 10.1046/j.1365-2958.1997.5911966.x
44. Begun, J., Sifri, C. D., Goldman, S., Calderwood, S. B., and Ausubel, F. M. (2005) Staphylococcus aureus virulence factors identified by using a high-throughput Caenorhabditis elegans-killing model. *Infect Immun*. **73**, 872–877
45. Carroll, R. K., Robison, T. M., Rivera, F. E., Davenport, J. E., Jonsson, I. M., Florczyk, D., Tarkowski, A., Potempa, J., Koziel, J., and Shaw, L. N. (2012) Identification of an intracellular M17 family leucine aminopeptidase that is required for virulence in Staphylococcus aureus. *Microbes Infect*. **14**, 989–999
46. Wilkinson, A. J., Fersht, A. R., Blow, D. M., and Winter, G. (1983) Site-Directed Mutagenesis as a Probe of Enzyme Structure and Catalysis: Tyrosyl-tRNA Synthetase Cysteine-35 to Glycine-35 Mutation. *Biochemistry*. **22**, 3581–3586
47. Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T. A. P., Rempfer, C., Bordoli, L., Lepore, R., and Schwede, T. (2018) SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res*. **46**, W296–W303
48. Bienert, S., Waterhouse, A., de Beer, T. A. P., Tauriello, G., Studer, G., Bordoli, L., and Schwede, T. (2017) The SWISS-MODEL Repository—new features and functionality. *Nucleic Acids Res*. **45**, D313–D319
49. Guex, N., Peitsch, M. C., and Schwede, T. (2009) Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: A historical perspective. *Electrophoresis*. **30**, S162–S173

Chapter 3

Staphylococcal diversity within the axilla of a single individual

3.1 Manuscript information

High-resolution *Staphylococcus* profiling reveals intra-species diversity in a single skin niche

Reyme Herman[1], Daniel T. West[1], Sean Meaden[1], Imran Khan[1], Robert Cornmell[2], Joanne Hunt[2], Michelle Rudden[3], Holly N. Wilkinson[3], Matthew J. Hardman[3], Anthony J. Wilkinson[4], Barry Murphy[2], Gavin H. Thomas[1]

[1] Department of Biology, University of York, Wentworth Way, York, YO10 5DD, UK
[2] Unilever Research & Development, Port Sunlight Laboratory, Quarry Road East, Bebington, Wirral, Merseyside, CH63 3JW, UK
[3] Centre for Biomedicine, Hull York Medical School, University of Hull, Hull, HU6 7RX, UK
[4] York Structural Biology Laboratory, Department of Chemistry, University of York, Wentworth Way, York, YO10 5DD, UK

Manuscript under review with Microbial Genomics.
Reviewer's comments have been addressed and changes have been incorporated.

Preprint DOI: 10.1101/2025.04.01.646391

Post viva status: Manuscript was published with Microbial Genomics on the 8th of October 2025.

DOI: 10.1099/mgen.0.001531

Co-authors have given permission for this manuscript to be used in this thesis.

3.2 Context within thesis (Abstract)

The overarching objective of the work presented in this thesis is to assess the roles of communities within the axillary microbiome in human malodour production. To facilitate this, bacterial communities of the same axillae were collected and genetically characterised (Chapters 3 and 4) in order to assemble a genetically diverse library of commensals for the purposes of community-based phenotypic malodour characterisation (Chapter 5).

In Chapter 3, the diversity of the genus *Staphylococcus* from a single high-malodour individual was assessed through a genetic characterisation workflow specifically developed for the large-scale genomic characterisation of isolates. The work reported in Chapter 3 identified 7 genetically diverse strains spanning three species of *Staphylococcus* and the potential inter-*Staphylococcus* networks that could have existed in the source axilla.

3.3 Abstract

The skin microbiome is dominated by a few key genera, among which *Staphylococcus* is one of the most well-characterised. Recent studies have examined the roles of various *Staphylococcus* species such as *S. epidermidis* and *S. hominis* within broader skin microbial communities. However, these investigations often rely on isolates from multiple individuals and hence limit their ability to capture intra-community interactions. In this study, we focused on the axillary microbiome of a single healthy individual to characterise the genetic and functional diversity of resident *Staphylococcus* isolates. Using a low-cost, high-throughput DNA extraction and long-read whole genome sequencing pipeline, we generated complete genomes for 93 isolates spanning seven genetically distinct lineages across three major skin species. These comprised of one dominant and three additional lineages of *S. epidermidis*, two of *S. hominis* and one of *S. capitis*. Functional and metabolic analyses revealed species- and strain-specific features, suggesting potential metabolic cross-feeding and specialisation within this community, including within strains of *S. epidermidis*. These findings highlight the metabolic complexity and potential interdependence of staphylococci inhabiting a single skin site and the need for strain-level resolution of the community. The strains form part of the York Skin Microbiome (YSM) collection, a growing open biobank of genetically diverse skin isolates from matched individuals.

3.4 Impact statement

Understanding how closely related microbial strains co-exist and interact within the same human body site is critical for deciphering the ecological and functional dynamics of the skin microbiome. This research pushes the boundaries of microbiome studies by employing a high-throughput long-read genome sequencing strategy on a single skin niche, revealing that strain-level genomic and functional diversity among coagulase-negative *Staphylococcus* is not only present but likely ecologically relevant. The discovery of discrete biosynthetic capacities, nutrient dependencies, and antimicrobial resistance traits, even among isolates of the same species, suggests a division of labour and potential metabolic cross-feeding within the community. These findings shift our understanding of skin microbiota from being merely species-rich to being structured by fine-scale functional variation.

Beyond its biological insights, this work provides a scalable, low-cost sequencing and bioinformatics framework for microbiome studies. This work forms part of the York Skin Microbiome (YSM) collection: a curated, genomically resolved resource for experimental and translational research. By anchoring microbiome investigations at the strain level, this approach enables more precise interrogation of host–microbe and microbe–microbe interactions, informing therapeutic development. As public interest grows in microbiome-targeted interventions, these results underscore the importance of moving beyond species averages to capture the true complexity within our microbial ecosystems.

3.5 Introduction

Human skin is home to a wide range of microorganisms that have evolved with their host over many millions of years (1–3). Naturally, these organisms interact with each other and the host in a symbiotic relationship to establish a favourable environment which protects the host while allowing for these microorganisms to thrive in generally nutrient limiting conditions(4,5). The multiple skin sites of the human body have some characteristic differences, such as being oily/sebaceous (e.g. face), dry/sebaceous poor (e.g. volar forearm) or moist (e.g. underarm/axilla) which then dictates the abundance and the diversity of microbes in the microbiome. Multiple synergistic and antagonistic relationships exist in these niches which helps maintain a healthy environment for the host (4). The genus *Staphylococcus* is one of the most abundant genera on human skin and can be found in multiple skin environments. While some species within this genus are well-known human pathogens (e.g. *Staphylococcus aureus*), cutaneous species of *Staphylococcus* play multiple roles in maintaining skin health (6–11). Species within this genus can also participate in ancient processes like malodour production. We previously identified a monophyletic group of coagulase-negative staphylococci (CoNS) including *S. hominis* and *S. lugdunensis* which are able to metabolise the human apocrine gland derived metabolite, Cys-Gly-3M3SH, and convert it to an odorous molecule we associate to human malodour (3,12–14). Interestingly, different isolates of the same species of *Staphylococcus* can have varying levels of activity against Cys-Gly-3M3SH suggesting as yet undiscovered genetic features that could control malodour formation (12,15).

Studies understanding the diversity within microbiomes have traditionally involved amplicon (e.g. 16S rRNA) sequencing of mixed communities or cultured isolates (16–18). Next generation sequencing (e.g. Illumina) then unlocked the ability to perform metagenomics; deep sequencing of these microbiome communities without the bias introduced by amplicon sequencing (19). The development of high-accuracy long read sequencing technologies (e.g. Oxford Nanopore) further equipped metagenomics studies with the ability to generate longer sequences and hence, could provide more complete genes, operons and even more complete genomes (20). In this study, we combine the strengths of long read sequencing and large-scale isolate culturing methods to characterise a microbiome community. We isolated the axillary staphylococci sampled from a single individual to understand the different genetic components that form the *Staphylococcus* community in a distinct body site. We adapted a low-cost high throughput DNA extraction and whole genome sequencing pipeline to enable the genetic assessment of these isolates. Through this, we reveal the genetic differences among these *Staphylococcus* isolates which could contribute to maintaining the skin microbiome community and ultimately skin health. This collection of staphylococci contributes to the establishment of the York Skin Microbiome (YSM) isolate resource of genetically diverse isolates from a variety of genera from matched individuals.

3.6 Results

3.6.1 Isolation and genetic characterisation of axillary *Staphylococcus*

The axilla of the volunteer was swabbed and then the captured material was plated onto a selective agar to enrich for cutaneous staphylococci. 96 isolates were harvested and stocked in the York Skin Microbiome (YSM) collection. The genomic DNA from these isolates was then extracted for long-read whole genome sequencing using long-read sequencing on the PromethION platform. We generated complete genome sequences together with accompanying plasmid sequences for 93 of the 96 isolates. Using Multi-locus Sequence Typing (MLST) (21), 90 of these isolates were determined to be *S. epidermidis* and 3 were *S. hominis* (Supplementary Tables 1 and 2). To determine the diversity of these isolates, we generated a Hadamard matrix (Fig 1A) using the pairwise average nucleotide identities and coverage of each alignment

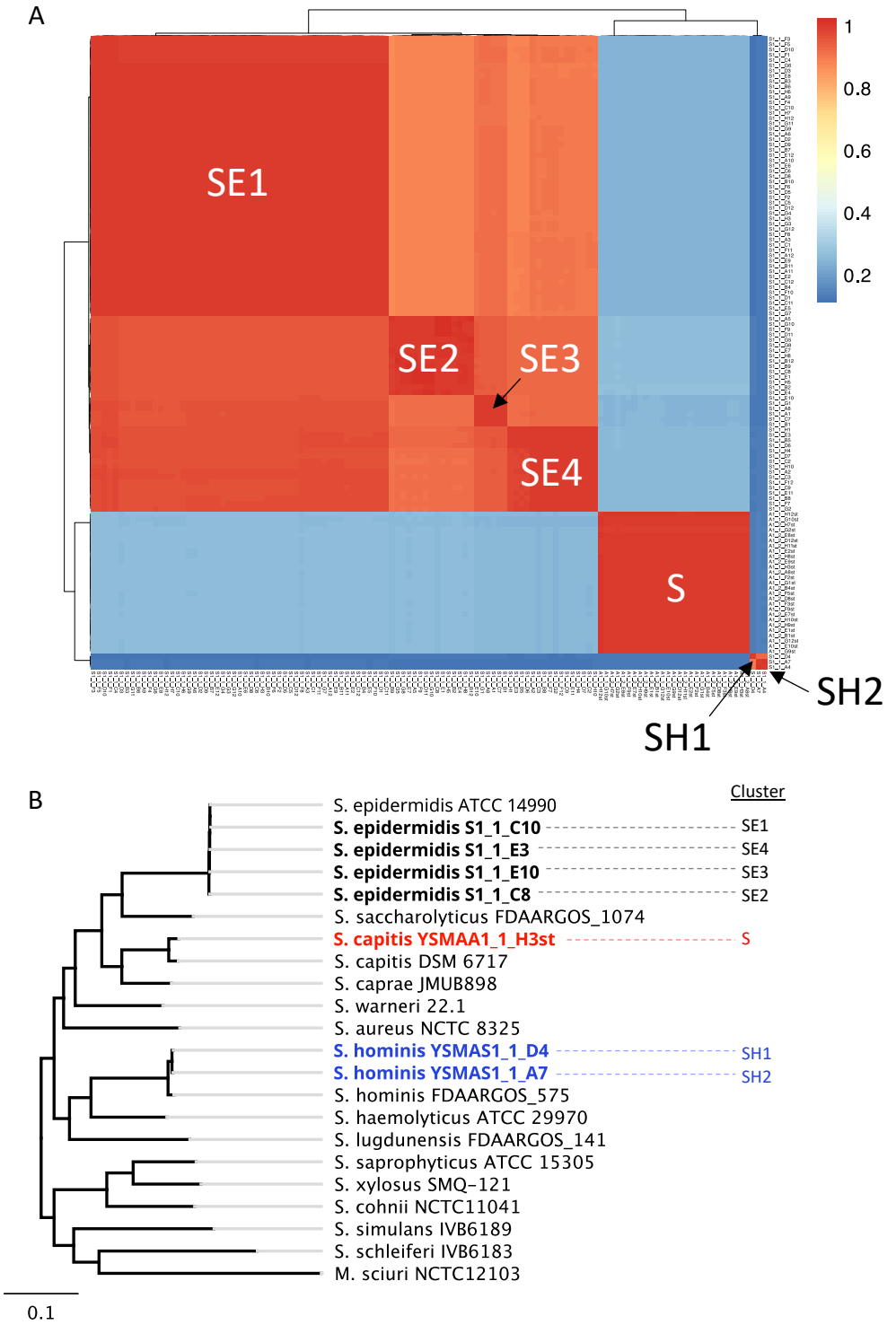


Figure 1. Diversity of 120 axillary staphylococci isolates. **(A)** A Hadamard matrix was constructed using the product of the pairwise-average nucleotide identities and the pairwise-alignment coverage and plotted using the pheatmap package on R. Isolate clusters were identified as SE1-4, SH1-2 or S. **(B)** Core genome maximum likelihood tree of all 7 dereplicated isolates with type strains of known cutaneous staphylococci. Clusters from (A) were identified next to the respective isolates.

calculated on pyani (20). The 90 *S. epidermidis* isolates are grouped into 4 distinct clusters (SE1, SE2, SE3 and SE4) with isolates of SE1 dominating this population of *S. epidermidis* isolates. The 3 *S. hominis* isolates were found to be from two separate clusters (SH1 and SH2). We then performed species specific phylogenetic analyses of the sequence types identified within our *Staphylococcus* isolates against all known MLST allelic profiles of the respective species. For *S. epidermidis*, clusters SE1 and SE4 were most similar, followed by SE2. However, cluster SE3 was found to be considerably more distant (Supplementary Fig 1). The MLST phylogenetic analysis of the *S. hominis* isolates suggested that SH1 and SH2, while clearly genetically distant, share a more recent common ancestor than most other sequence types within this species (Supplementary Fig 2).

We also unexpectedly identified 27 additional isolates of staphylococci using corynebacteria enriching media containing Fosfomycin which we will include in our downstream analyses (23). Analysis using pyani suggested these 27 isolates are genetically highly similar and hence are grouped in the same cluster (S) (Fig 1A). Unsurprisingly, data from Fig 1A suggest we have highly similar genomes in the library of *Staphylococcus* isolates from the volunteer due to indiscriminate large-scale isolation. Hence, we performed dereplication of the isolates using drep (24) with a secondary cluster threshold of 99.5% to define a set of 7 genomes to represent isolates from clusters SE1, SE2, SE3, SE4, SH1, SH2 and S for use in downstream analysis. The core genome of the 7 dereplicated isolates and type strains of other skin associated staphylococci were derived and aligned using Roary (25) (Supplementary Table 3-6). The resulting alignment was then used to derive a maximum-likelihood tree (Fig 1B). The core genome of isolate YSMAA1_1_H3st, representing isolate cluster S (Fig 1A), clustered with *S. capitis* and hence was assigned the species. The full 16S sequences of YSMAA1_1_H3st and *S. capitis* DSM 6717 were found to be identical, further confirming the species assignment (Supplementary Fig 3). This approach provided us with complete genomes of multiple genetically different staphylococci (4 *S. epidermidis*, 2 *S. hominis* and 1 *S. capitis*) from the same individual which enables us to understand how these bacteria may have interacted in the microbiome they were isolated from.

3.6.2 Species- and isolate-specific genetic traits reveal potential roles and functional networks within a single environment

We then performed a pangenome analysis using anvi'o (26) which allowed us to identify genes specific to each species as well as genes only found in individual isolates (singletons) with the aim of revealing the roles of each genetically distinct isolate in this environment (Fig 2).

The pangenome analysis identified 1609 core genes shared by all tested isolates of this genus with an additional 116 core *S. epidermidis* genes and 281 core *S. hominis* genes. These genes were assigned putative functions (COG) using eggNOG-mapper (27,28) to help elucidate the roles of the various species in the underarm of this single individual. As we were only able to isolate one subtype of *S. capitis*, we included the singletons identified for isolate *S. capitis* YSMAA1_1_H3st to represent the species. Major species-level differences were in the number of genes predicted to be involved in nucleotide related processes like transcription (K) and replication, recombination and repair (L) and also cell surface biogenesis (M) (Fig 2B). We also observed significant differences in the number of genes predicted to be involved in transport processes and metabolism, suggesting there may be species specific pathways. To further resolve the diversity of these isolates, we performed the same functional analysis on the singletons of all *S. hominis* and *S. epidermidis* isolates and compared them to other isolates of the same species (Fig 3). Even within species level comparisons, we observe the largest differences in transcription (K) and replication, recombination and repair (L). This pangenome analysis workflow was also used to demonstrate the suitability of the isolates that were chosen to represent the isolate clusters reported in Fig 1. Isolates within each cluster were compared between each other to identify non-core genes (i.e. not found in every isolate) and the representative isolates were assessed for the presence of these non-core genes (Supplementary Table 3-7). We were not able to identify any non-core genes in the *S. hominis* isolates of this study. Non-core genes which were found in the majority of the S. epidermidis and *S. capitis* isolates of each cluster were also found in each of the representative isolates. An exception to this was in cluster SE2 where there were plasmid located genes found within 8 out of 15 isolates, not including the representative isolate YSMAS1_1_C8 (Supplementary Table 4). However, functional annotation was only

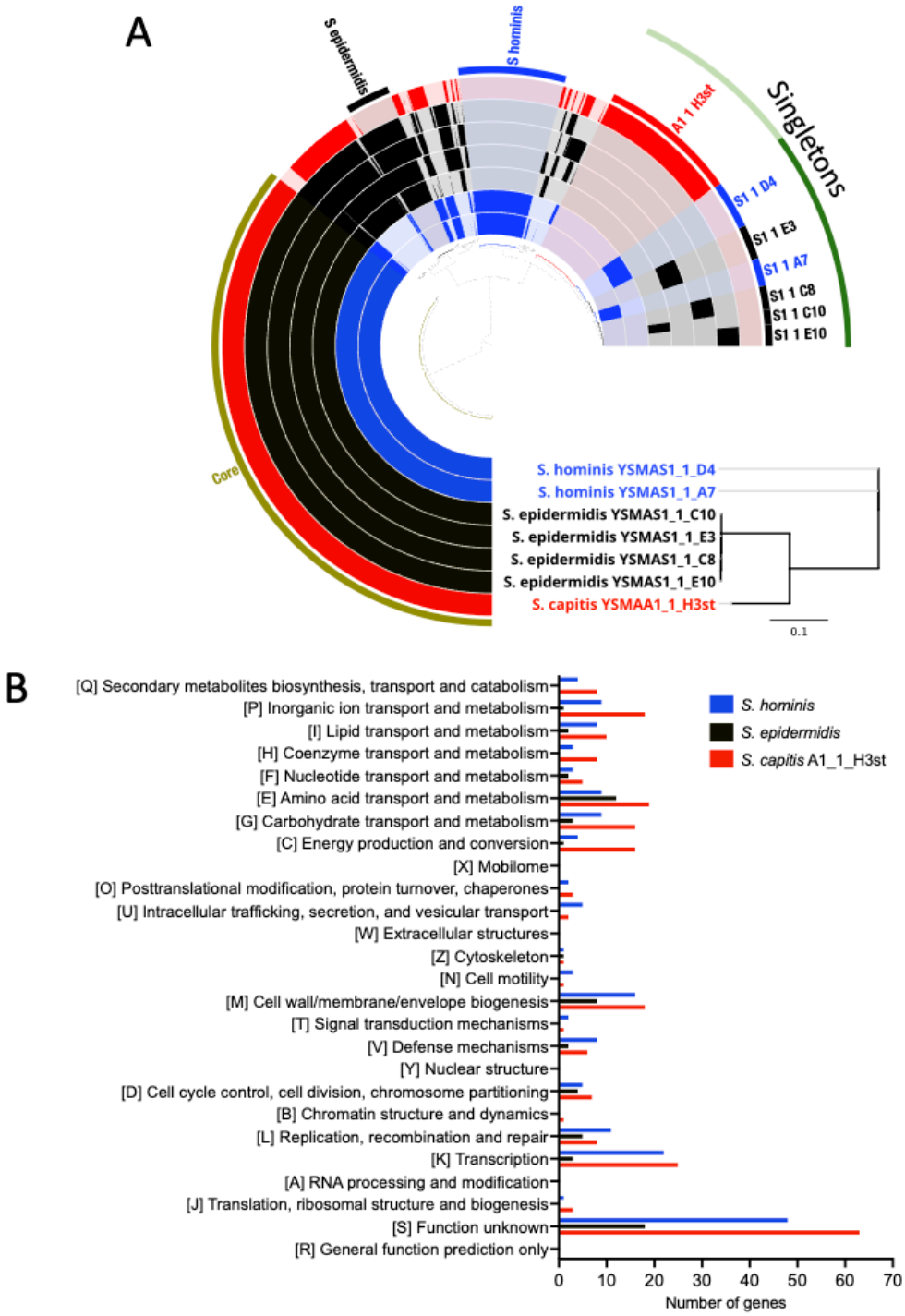


Figure 2. (A) Pangenome analysis of the 7 dereplicated axillary staphylococci isolates identified species-specific clusters (*S. epidermidis* and *S. hominis*) and gene clusters belonging to only one isolate (singletons). The A1 1 H3st gene cluster identified in *S. capitis* YSMAA1_1_H3st was used in species-specific comparisons. The pangenome was calculated and visualised on anvio-8. (B) COG categories were assigned to species-specific gene clusters using eggNOG-mapper and visualised on Graphpad Prism 10.

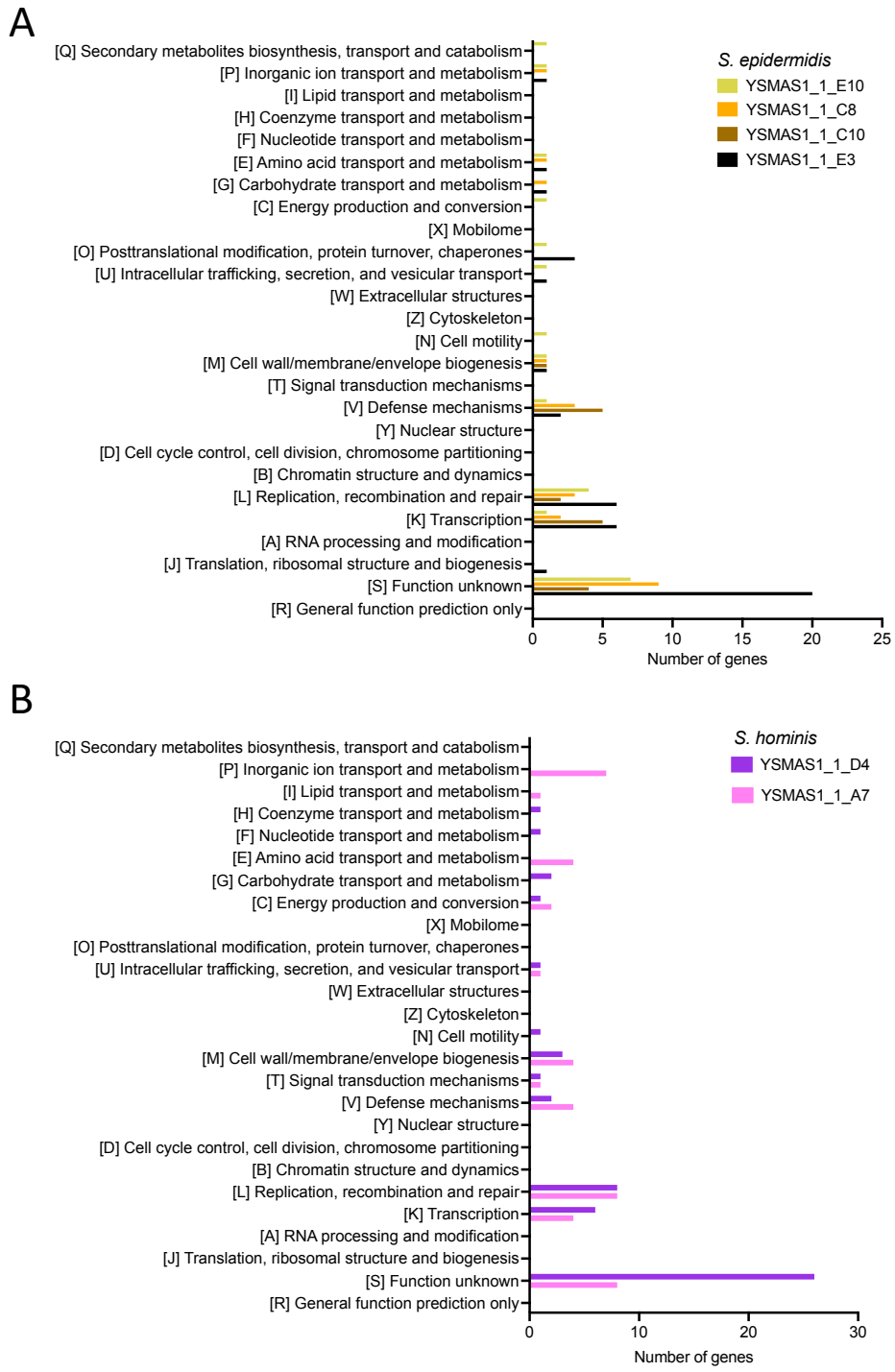


Figure 3. COG categories were assigned to singletons from (A) *S. epidermidis* and (B) *S. hominis* isolates using eggNOG-mapper and visualised on Graphpad Prism 10.

successful for one of these genes which was likely to be a plasmid replication protein and hence, deemed to not be critical in the downstream analysis.

We then performed KEGG analysis on all 7 isolates to map full pathways to complement the COG analysis, in particular to assess the presence of known metabolic pathways (Fig 4, Supplementary Fig 4 & 5). As observed in the COG analysis, there seem to be some correlation between the presence of some complete pathways and specific species. Here, we discuss some examples of the species- and subspecies/strain-level differences which could allude to potential metabolic functions in the microbiome. Within the amino acid metabolism pathways, the ornithine biosynthesis module (M00028) seems to only be complete in *S. epidermidis* isolates and the entire module was found to be missing from *S. hominis* and *S. capitis*. Ornithine can be used in the biosynthesis of arginine as described in the KEGG module M00844 which is found to be present in all 7 isolates (Fig 4). For this purpose, ornithine could be imported from the environment using the ornithine transporter ArcD (29). While the gene encoding for ArcD was identified in *S. capitis* YSMAA1_1_H3st, this transporter was not identified in the genomes of either of the two *S. hominis* isolates, suggesting the presence of a different ornithine transport pathway.

Similarly, the biotin biosynthesis pathways (M00123 and M00577) seem to only be present in *S. epidermidis* strains, but is missing in *S. epidermidis* YSMAS1_1_C10, with the loss of the *bioDKFW* genes in a large strain specific deletion on the chromosome (Supplementary Fig 6). The strain now resembles the *S. hominis* and *S. capitis* isolates in requiring exogenous biotin, which suggests some subtypes of cutaneous *S. epidermidi*s could be the source of biotin for other auxotrophic species of this genus in the same community. The transporter BioY has been identified as a probable biotin uptake system in *S. aureus* (30) with the gene encoding BioY identified in all 4 isolates deficient in biotin biosynthetic pathways. Using the complete genomes of these isolates, we were able to perform an in-depth metabolic analysis to suggest the presence of metabolic cross-talk between members of the same skin microbiome.

We also observed differences within isolates of individual species. For example, the staphylopine biosynthesis pathway (M00949), deriving the zinc-binding siderophore from histidine (31), is found in three *S. epidermidis* isolates. The genes involved in

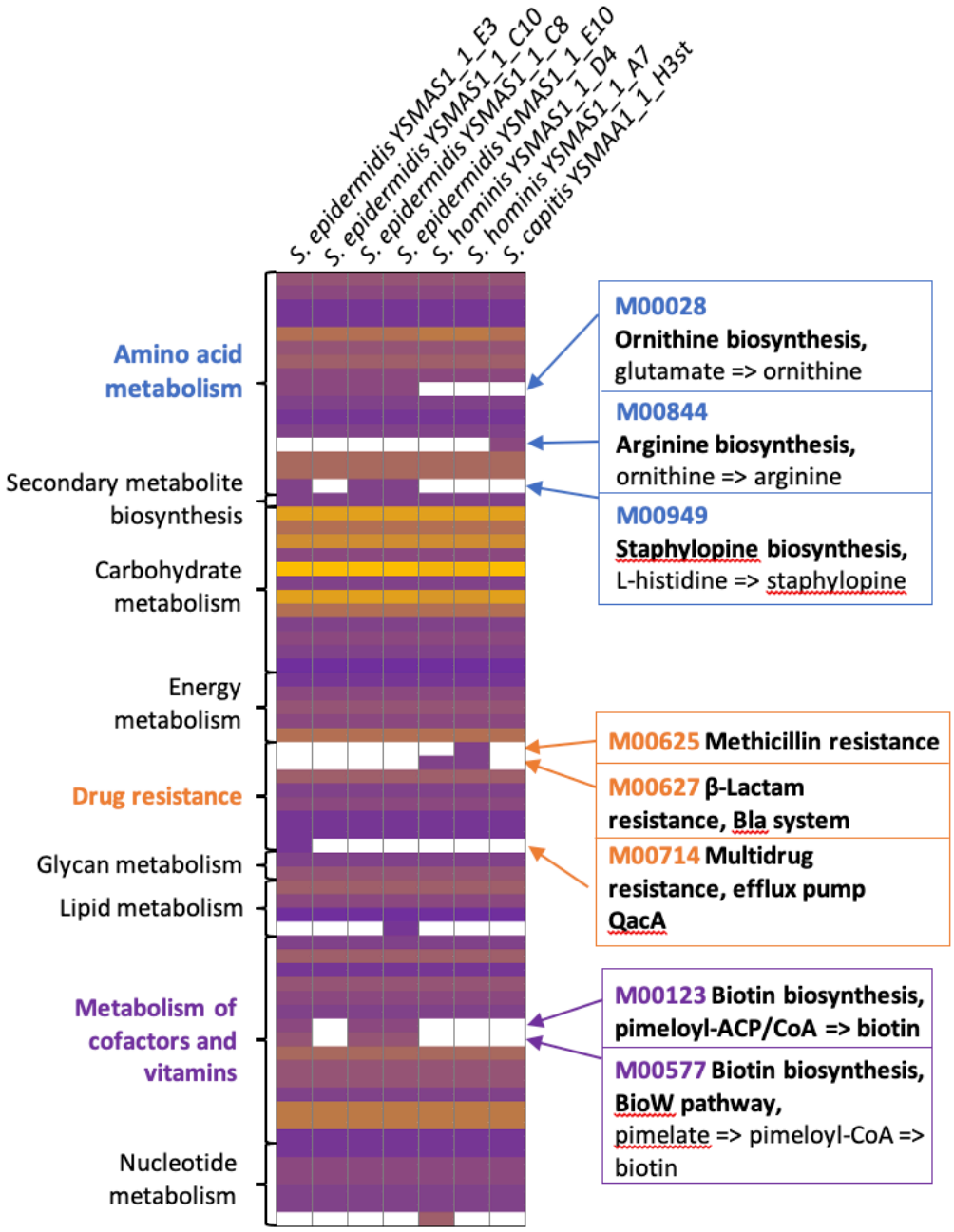


Figure 4. Metabolic pathways in genetically distinct axillary staphylococci reveals species and isolate level variation within the isolates. Each complete KEGG module for each isolate were placed on a heatmap where the modules with the largest number of reactions are in yellow and lowest in purple with absent modules in white. Modules of interest were described on the right of the heatmap. KEGG modules were grouped according to the categories on the left of the heatmap.

staphylopine biosynthesis (*cntKLM*) are found in the same region of the chromosome as the biotin biosynthetic genes (*bioDKFW*) and hence, the large chromosomal deletion also removed *cntKLM* from *S. epidermidis* YSMAS1_1_C10 (Supplementary Fig 6). Alongside *S. epidermidis* YSMAS1_1_C10, both *S. hominis* and the *S. capitis* isolates were also missing *cntKLM*. Rather than synthesising staphylopine, these isolates could instead utilise and import this siderophore, however, they lack the direct orthologue of the *S. aureus* staphylopine CntABCDF ATP-binding cassette (ABC) permease (32) encoded by genes found downstream from *cntKLM* in all three *S. epidermidis* isolates with the complete staphylopine biosynthetic pathway.

In addition to metabolism, we also assessed the genomes for resistance cassettes and defence systems to understand the inherent protective measures which could maintain the community of staphylococci from the same site. In the KEGG analysis (Fig 4) we observed the presence of modules providing resistance to antimicrobials including β-lactams via the bla system (M00627) and the mec systems (M00625) (33,34) in specific *S. hominis* isolates and to some antiseptics via the efflux pump QacA (35) in *S. epidermidis* YSMAS1_1_E3. These findings were confirmed using ResFinder (Supplementary Table 8).

The COG analysis (Fig 3) also identified species- and isolate-specific anti-phage defence systems (COG Category: X) which suggest that we can expect an array of different systems in these isolates. We then analysed the genomes of these isolates on padloc (36) to reveal putative phage defence systems (Fig 5). As expected, the padloc analysis predicted the presence of a wide range of defence systems. Some systems are found in all isolates of the same species like the Type I and IV restriction modification (RM) systems in S*. epidermidis* isolates and the Type II RM system in *S. hominis* isolates. However, as suggested by the species-specific COG analyses (Fig 3), we also see isolate specific systems. Notably, in *S. epidermidis*, we observed the presence of the Type II RM system, the Type I defence-associated reverse transcriptase (DRT_class_I), the abortive infection related systems (AbiG and AbiD) and SEFIR in

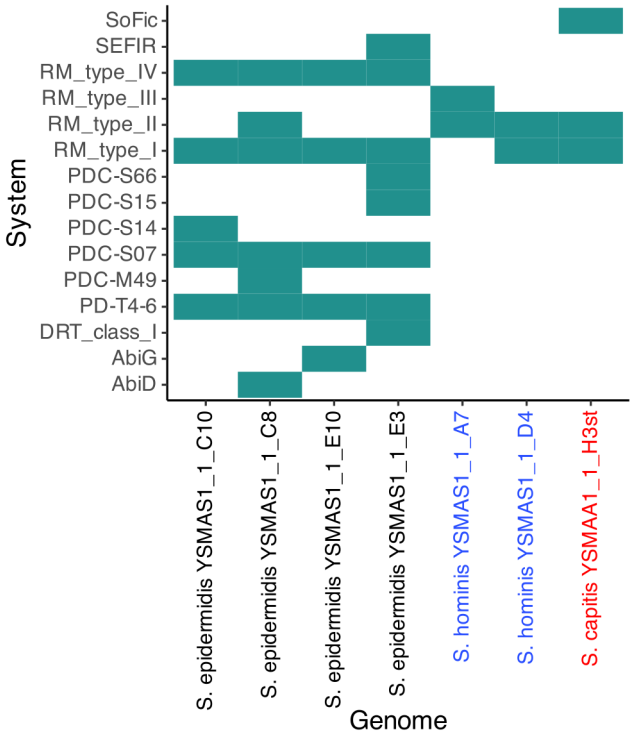


Figure 5. Phage defence systems were predicted in the 7 dereplicated genomes using padloc. Only one of each system was identified in each relevant genome. Names of isolates were coloured according to the species they belong to (*S. epidermidis*: Black, *S. hominis*: Blue, *S. capitis*: Red). Data was visualised using R.

either of the 4 dereplicated isolates. Further, we also identified putative defence systems including some previously reported by Vassallo *et al*. (37) but these were not found in both *S. hominis* isolates and *S. capitis* YSMAA1_1_H3st. In *S. hominis*, predicted defence systems were limited to RM systems with the isolates either possessing a Type III or a Type I RM system alongside the Type II RM system. We could use information from this analysis to aid in the design of genetic tools to target specific species or even specific isolates which will be invaluable in the understanding the communities of the skin microbiome.

3.7 Discussion

Recent advances in sequencing technologies have played a crucial role in revolutionising our understanding of the human skin microbiome. Both short- and long-read sequencing technologies have provided us with the means to unlock the microbiome to study the abundance of different species in a single community through metagenomics and also enable cheaper whole genome sequencing of isolates. In this study, we have proposed the use of a staphylococci-focused isolation approach involving the use of a low-cost total DNA extraction procedure and whole genome sequencing for the isolation and characterisation of cutaneous species. We tested this approach using isolates derived from a single volunteer which allowed us to assemble a library of 120 staphylococcal isolates containing 7 genetically distinct lineages of isolates from three different species.

Through genomic analysis and downstream functional annotation, we suggested possible roles for isolates of three major cutaneous *Staphylococcus* species during colonisation. The subspecies-level resolution of the KEGG analysis allowed us to identify potential metabolic networks involving molecules like ornithine, staphylopine and biotin, indicating possible cross-talk across different species (Fig 4, Supplementary Fig 4 & 5). Considering the completeness of the study, our results may be influenced by limitations in current isolation methods. Our protocols yielded a disproportionate number of *S. epidermidis* isolates, potentially underrepresenting other *Staphylococcus* species. To address this, future studies could explore the design of targeted enrichment media aimed at improving species diversity. Insights from the KEGG analysis (Fig. 4) could inform the development of such media, demonstrating

the practical utility of this genomic dataset beyond descriptive analysis. Further, our findings reported in this study are entirely based on *in silico* analyses and forms the initial step in predict potential interactions within microbiome communities.

Our workflow, powered by Oxford Nanopore's long-read whole genome sequencing technology, provided us with the ability to genetically characterise the library of isolates at a low per sample cost and could potentially be adapted to other laboratory setups. The methods and findings of this study provide a platform to establish larger scale studies of volunteers and isolates of their skin microbiome, not just on healthy individuals but could also be adapted for studies on diseased skin. The strain-level resolution described by our approach provides unparalleled access to fine-scale genetic variation, offering a powerful tool for dissecting microbe–microbe and microbe–host interactions. Ultimately, such insights will be critical for understanding how specific microbial strains contribute to skin health and disease, and for identifying potential targets for therapeutic modulation of the skin microbiome.

3.8 Materials and methods

3.8.1 Sampling procedures

The adult male volunteer was recruited for axillary swabbing at Unilever R&D, Bebington, UK. The volunteer was in good general health, had not used any cosmetic products for 24 hours before the study, was not taking antibiotics, had no active skin conditions, has not suffered from eczema in the last 5 years nor have they ever had psoriasis. Axillary swabs were collected by swabbing in a linear motion 20 times using eSwabs (Copan) on both underarms and then placed into Amies transport medium.

3.8.2 Bacterial culturing

Swabs were plated onto SS solid media (10 g L^{-1} tryptone, 5 g L^{-1} Lemco powder, 3 g L^{-1} yeast extract, 13 g L^{-1} agar no. 1, 10 g L^{-1} sodium pyruvate, 0.5 g L^{-1} glycine, 22.5 g L^{-1} potassium thiocyanate, 1.2 g L^{-1} disodium hydrogen orthophosphate dihydrate, 0.67 g L^{-1} sodium dihydrogen phosphate, 2 g L^{-1} lithium chloride, 10 mL L^{-1} glycerol, pH 7.2) or ACP solid media (39.5 g L^{-1} Blood Agar base no. 2, 3 g L^{-1} yeast

extract, 2 g L^{-1} glucose, 5 mL L^{-1} Tween 80, 50 mL L^{-1} defibrinated horse blood, 100 mg L^{-1} fosfomycin) to enrich for staphylococci and corynebacteria respectively. Colonies were allowed to grow for up to 2 days at 37°C. Colonies were picked and grown overnight in brain-heart infusion broth + 1 % Tween-80 (BHIT) at 37°C, shaking at 200 rpm. Cultures were stocked in BHIT supplemented with 10 % glycerol in 2 mL 96 well plates.

3.8.3 DNA extraction

Total (genomic and plasmid) DNA extraction was performed using a previously reported procedure (23) with the following changes: the enzymatic lysis buffer used also included 125 μg/mL lysostaphin, the enzymatic lysis step was performed over 5 hours and DNA precipitation was performed for 15 minutes at room temperature.

3.8.4 DNA library preparation and whole genome sequencing

The modified Oxford Nanopore Native Barcoding Kit 96 V14 DNA library preparation protocol reported in Herman *et al*. (23) was used. The prepared library was loaded into a PromethION R10.4.1 flow cell and sequencing was performed for 72 hours and base-called using the singleplex high-accuracy model, 400 bps on MinKNOW 23.07.5. The reads were assembled with the EPI2ME Labs wf-bacterial-genomes isolate workflow (Flye (38), Medaka (Oxford Nanopore Technologies), Prokka (39).

3.8.5 Bioinformatics tools

MLST phylogeny analyses were performed on RStudio 2024.01.1 by first calculating the pairwise mismatches (distances) between the allelic profiles of all known sequence types of *S. epidermidis* and *S. hominis* available on PubMLST (www.pubmlst.org, accessed on the 29th of August 2025) using the dplyr package. The distances were then used to construct neighbour-joining trees using the ape package. Pairwise genome comparisons of all assembled genomes were performed using the ANIm (40) analysis on pyANI v0.2.12 (41). The genomes were dereplicated with drep v3.4.2 (24) using Mash (42) and MUMmer 3.0 (43), with a primary clustering ANI cutoff of 95% to

distinguish between different species and a secondary clustering ANI cutoff of 99.5% to capture strain diversity to derive a representative set of genomes. The core genomes were determined using Roary (44) with an 80% BLASTp cutoff. PhyML 3.3 (45) trees of the core genome alignments of the representative genomes were constructed using the HKY85 substitution model with approximate likelihood ratio tests (aLRT). Trees were visualised on Geneious Prime 2024.0.7 (Biomatters). 16S rRNA alignments were performed using the Clustal Omega algorithm and visualised on Geneious Prime 2024.0.7 (Biomatters).

Pangenome analysis was performed using the anvi-pan-genome package and visualised on anvi'o v8 (26). Relevant gene clusters were extracted and functionally analysed with eggNOG-mapper v2 (27,28) to determine the respective COG categories. Metabolic pathways were identified using blastKOALA and KEGG Reconstruct (46,47). Phage defence systems were predicted using PADLOC v4.3 (36).

3.9 Ethics statement

This study was conducted in accordance with the ethical principles of Good Clinical Practice and the Declaration of Helsinki. The local ethics committees of Unilever R&D (Port Sunlight, UK) approved the protocols before commencement of the studies and all subjects gave written informed consent.

3.10 Data availability

Genomes of the 7 dereplicated genomes were deposited to the GenBank repository within the BioProject PRJNA1191350.

3.11 Conflict of interest

The authors declare that B. M., J. H. and R. C. are current employee of Unilever when this study was carried out. All other authors declare that they have no conflicts of interest with the contents of this article.

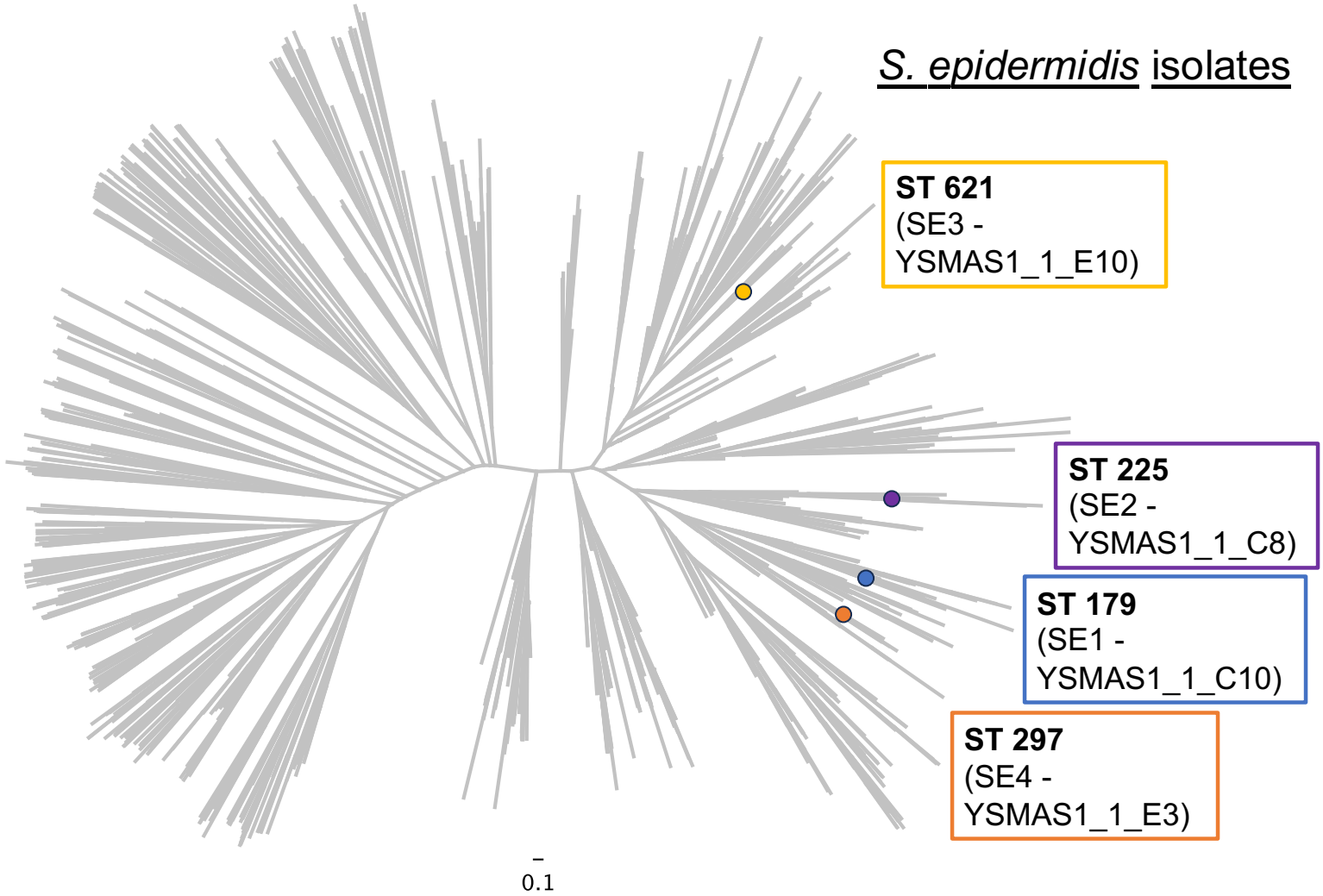
3.12 Funding

This work was supported by the Biotechnology and Biological Sciences Research Council grant BB/W510531/1 supporting R. H.

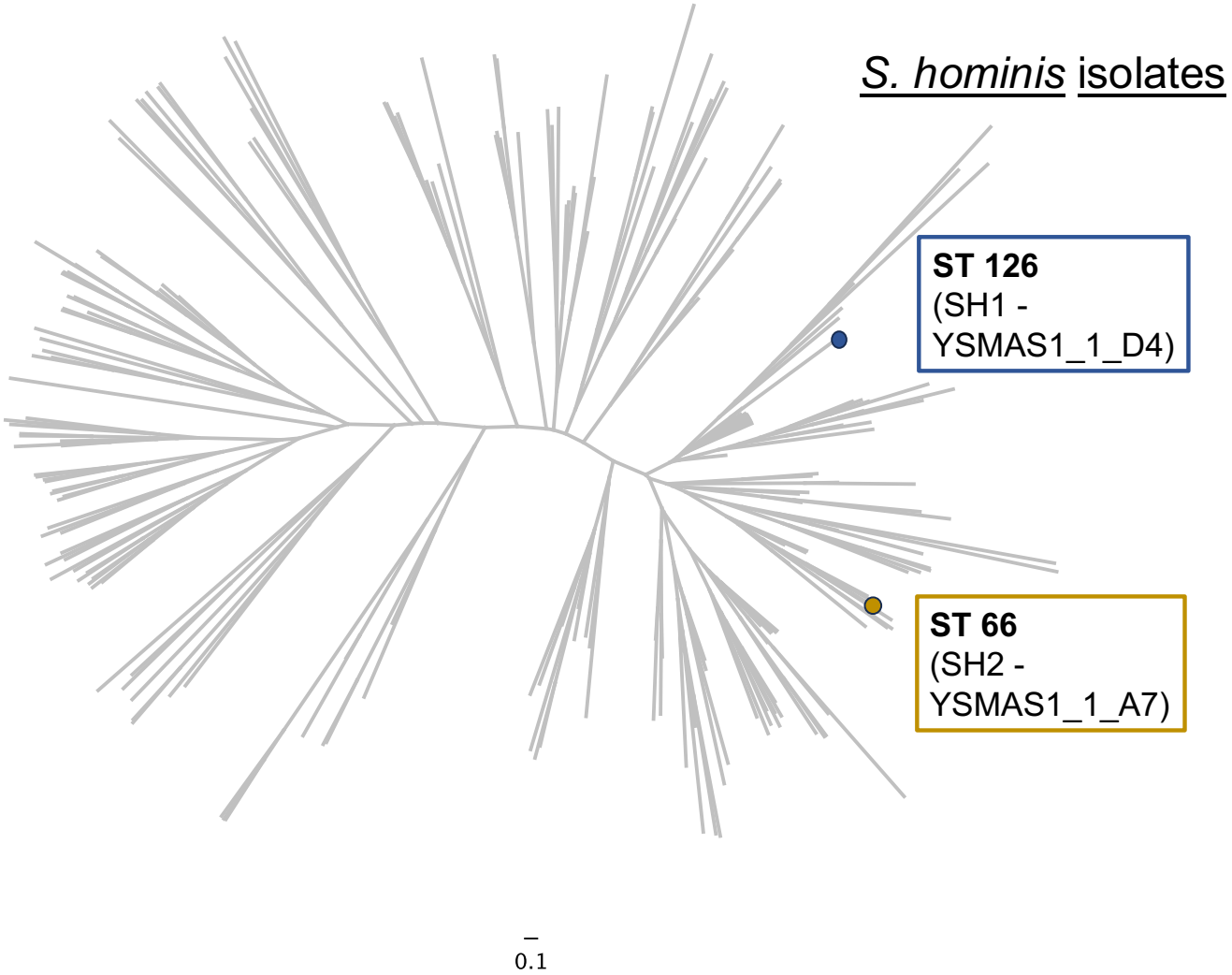
3.13 Author contributions

R. H., R. C, B. M. and G. H. T. conceptualization; R. H., D. T. W, M. R., and B. M. validation; R. H., D. T. W., and G. H. T. formal analysis; R. H., D. T. W., I. K., J. H. and M. R. investigation; R. H. data curation; H. W., M. R. and M. H. resources; R. H. writing–original draft; R. H., D. T. W., S. M., I. K., H. W., A. J. W., B. M., and G. H. T. writing–review & editing; R. H. and S. M. visualization; A. J. W., B. M., and G. H. T. supervision; R. H., B. M., and G. H. T. project administration; B. M. and G. H. T. funding acquisition.

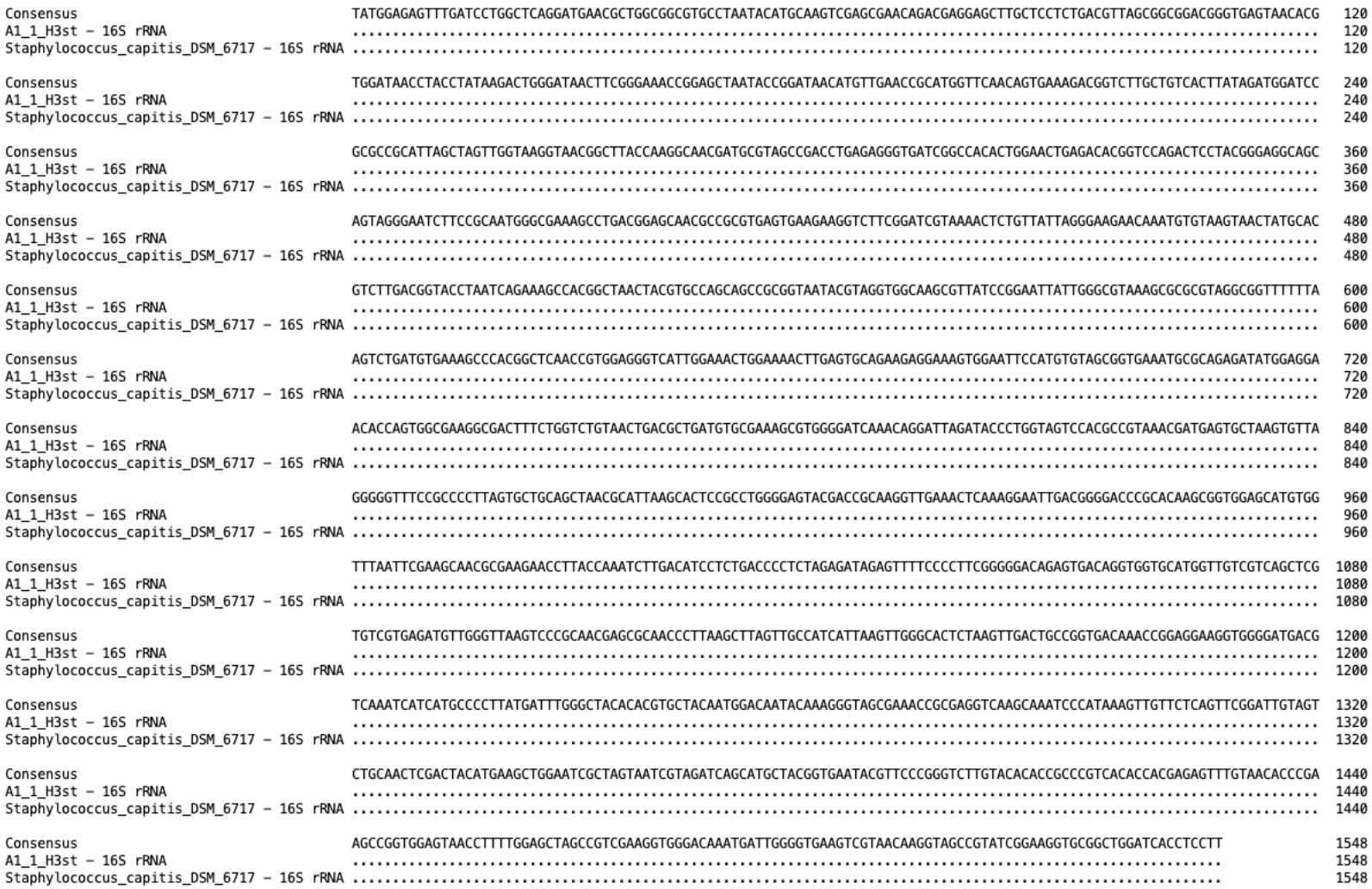
3.14 Supporting data



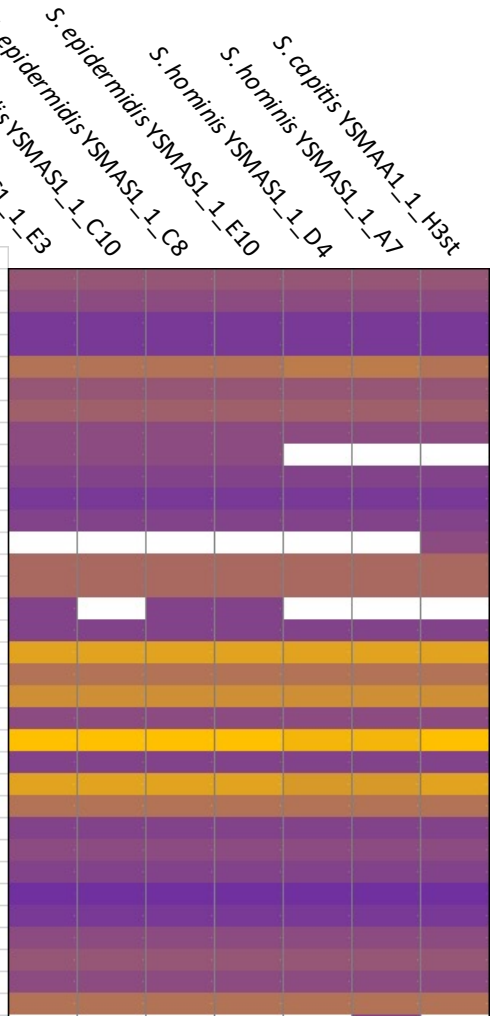
Supplementary Fig 1. Phylogenetic analysis of *S. epidermidis* MLST sequence types using the defined allelic profiles on PubMLST (www.pubmlst.org). The four clusters (SE1-4) were identified by the sequence types and the representative isolate.



Supplementary Fig 2. Phylogenetic analysis of *S. hominis* MLST sequence types using the defined allelic profiles on PubMLST (www.pubmlst.org). The two clusters (SH1-2) were identified by the sequence types and the representative isolate.

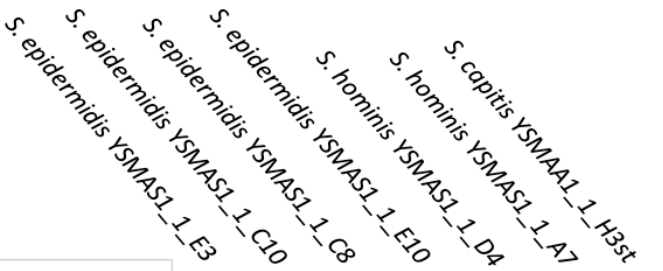


Supplementary Fig 3. Full 16S rRNA alignment of YSMAA1_1_H3st and *S. capitis* DSM 6717 (type strain).



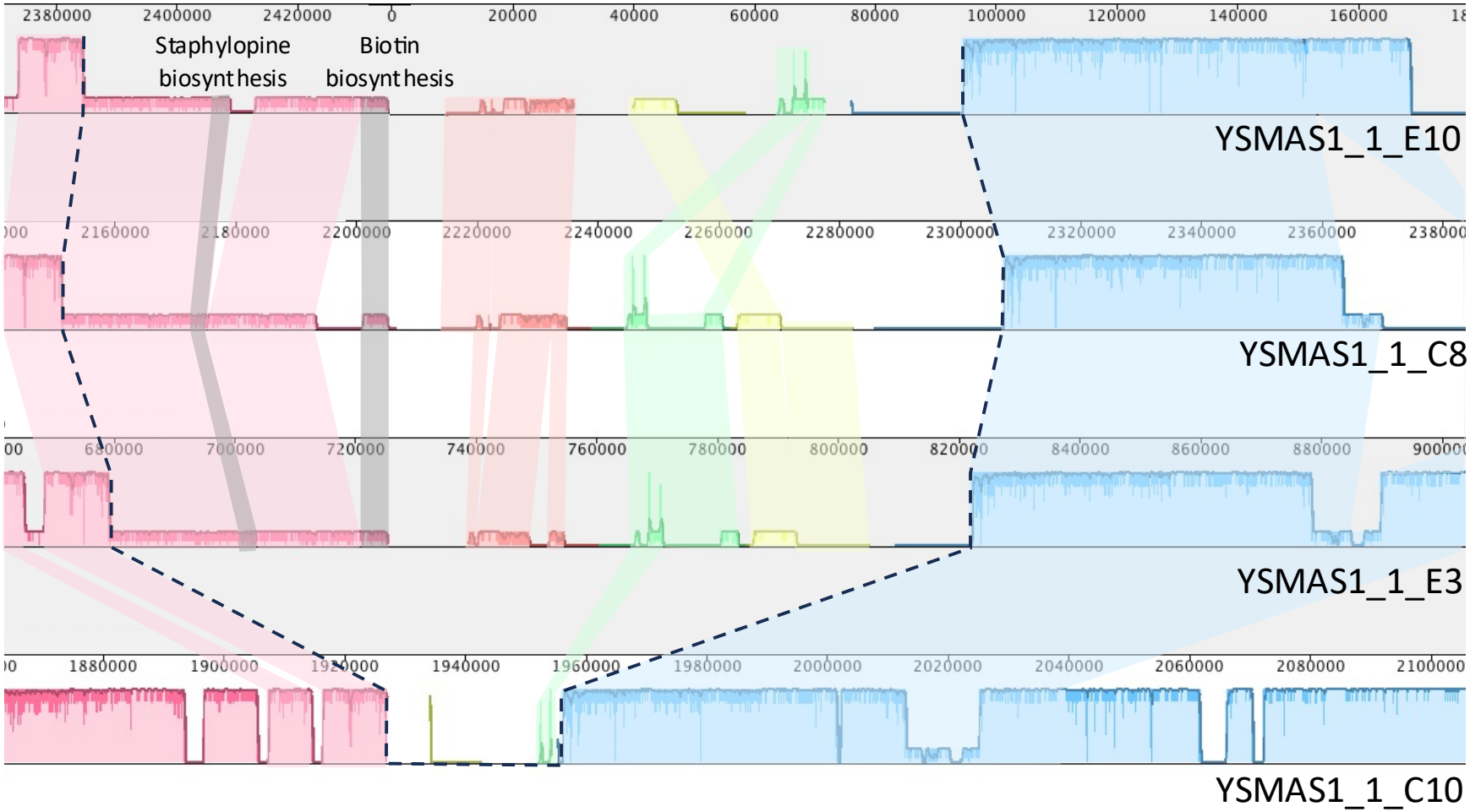
Super Categories	Categories	Module	
Amino acid metabolism	Serine and threonine metabolism	M00018	Threonine biosynthesis, aspartate => homoserine => threonine
Amino acid metabolism	Serine and threonine metabolism	M00621	Glycine cleavage system
Amino acid metabolism	Serine and threonine metabolism	M00555	Betaine biosynthesis, choline => betaine
Amino acid metabolism	Cysteine and methionine metabolism	M00021	Cysteine biosynthesis, serine => cysteine
Amino acid metabolism	Cysteine and methionine metabolism	M00017	Methionine biosynthesis, aspartate => homoserine => methionine
Amino acid metabolism	Branched-chain amino acid metabolism	M00019	Valine/isoleucine biosynthesis, pyruvate => valine / 2-oxobutanoate => isoleucine
Amino acid metabolism	Branched-chain amino acid metabolism	M00570	Isoleucine biosynthesis, threonine => 2-oxobutanoate => isoleucine
Amino acid metabolism	Branched-chain amino acid metabolism	M00432	Leucine biosynthesis, 2-oxoisovalerate => 2-oxoisocaproate
Amino acid metabolism	Arginine and proline metabolism	M00028	Ornithine biosynthesis, glutamate => ornithine
Amino acid metabolism	Arginine and proline metabolism	M00844	Arginine biosynthesis, ornithine => arginine
Amino acid metabolism	Arginine and proline metabolism	M00970	Proline degradation, proline => glutamate
Amino acid metabolism	Arginine and proline metabolism	M00972	Proline metabolism
Amino acid metabolism	Histidine metabolism	M00045	Histidine degradation, histidine => N-formiminoglutamate => glutamate
Amino acid metabolism	Aromatic amino acid metabolism	M00022	Shikimate pathway, phosphoenolpyruvate + erythrose-4P => chorismate
Amino acid metabolism	Aromatic amino acid metabolism	M00023	Tryptophan biosynthesis, chorismate => tryptophan
Amino acid metabolism	Other amino acid metabolism	M00949	Staphylopine biosynthesis, L-histidine => staphylopine
Biosynthesis of terpenoids and polyketides	Terpenoid backbone biosynthesis	M00876	Staphyloferrin A biosynthesis, L-ornithine => staphyloferrin A
Carbohydrate metabolism	Central carbohydrate metabolism	M00001	Glycolysis (Embden-Meyerhof pathway), glucose => pyruvate
Carbohydrate metabolism	Central carbohydrate metabolism	M00002	Glycolysis, core module involving three-carbon compounds
Carbohydrate metabolism	Central carbohydrate metabolism	M00003	Gluconeogenesis, oxaloacetate => fructose-6P
Carbohydrate metabolism	Central carbohydrate metabolism	M00307	Pyruvate oxidation, pyruvate => acetyl-CoA
Carbohydrate metabolism	Central carbohydrate metabolism	M00009	Citrate cycle (TCA cycle, Krebs cycle)
Carbohydrate metabolism	Central carbohydrate metabolism	M00010	Citrate cycle, first carbon oxidation, oxaloacetate => 2-oxoglutarate
Carbohydrate metabolism	Central carbohydrate metabolism	M00011	Citrate cycle, second carbon oxidation, 2-oxoglutarate => oxaloacetate
Carbohydrate metabolism	Central carbohydrate metabolism	M00004	Pentose phosphate pathway (Pentose phosphate cycle)
Carbohydrate metabolism	Central carbohydrate metabolism	M00006	Pentose phosphate pathway, oxidative phase, glucose 6P => ribulose 5P
Carbohydrate metabolism	Central carbohydrate metabolism	M00007	Pentose phosphate pathway, non-oxidative phase, fructose 6P => ribose 5P
Carbohydrate metabolism	Central carbohydrate metabolism	M00580	Pentose phosphate pathway, archaea, fructose 6P => ribose 5P
Carbohydrate metabolism	Central carbohydrate metabolism	M00005	PRPP biosynthesis, ribose 5P => PRPP
Energy metabolism	Carbon fixation	M00579	Phosphate acetyltransferase-acetate kinase pathway, acetyl-CoA => acetate
Energy metabolism	Methane metabolism	M00345	Formaldehyde assimilation, ribulose monophosphate pathway
Energy metabolism	Sulfur metabolism	M00176	Assimilatory sulfate reduction, sulfate => H2S
Energy metabolism	ATP synthesis	M00416	Cytochrome aa3-600 menaquinol oxidase
Energy metabolism	ATP synthesis	M00157	F-type ATPase, prokaryotes and chloroplasts

Supplementary Fig 4. KEGG modules shown in Fig 4 are identified. Included are modules from the Super Categories amino acid metabolism, biosynthesis of terpenoids and polyketides, carbohydrate metabolism and energy metabolism.



Super Categories	Categories	Module	
Gene set	Drug resistance	M00625	Methicillin resistance
Gene set	Drug resistance	M00627	beta-Lactam resistance, Bla system
Gene set	Drug resistance	M00725	Cationic antimicrobial peptide (CAMP) resistance, dltABCD operon
Gene set	Drug resistance	M00726	Cationic antimicrobial peptide (CAMP) resistance, lysyl-phosphatidylglycerol (L-PG) synthase MprF
Gene set	Drug resistance	M00730	Cationic antimicrobial peptide (CAMP) resistance, VraFG transporter
Gene set	Drug resistance	M00700	Multidrug resistance, efflux pump AbcA
Gene set	Drug resistance	M00702	Multidrug resistance, efflux pump NorB
Gene set	Drug resistance	M00714	Multidrug resistance, efflux pump QacA
Glycan metabolism	Nucleotide sugar biosynthesis	M00549	Nucleotide sugar biosynthesis, glucose => UDP-glucose
Glycan metabolism	Nucleotide sugar biosynthesis	M00909	UDP-N-acetyl-D-glucosamine biosynthesis, prokaryotes, glucose => UDP-GlcNAc
Lipid metabolism	Fatty acid metabolism	M00082	Fatty acid biosynthesis, initiation
Lipid metabolism	Fatty acid metabolism	M00083	Fatty acid biosynthesis, elongation
Lipid metabolism	Fatty acid metabolism	M00086	beta-Oxidation, acyl-CoA synthesis
Lipid metabolism	Lipid metabolism	M00098	Acylglycerol degradation
Metabolism of cofactors and vitamins	Cofactor and vitamin metabolism	M00899	Thiamine salvage pathway, HMP/HET => TMP
Metabolism of cofactors and vitamins	Cofactor and vitamin metabolism	M00125	Riboflavin biosynthesis, plants and bacteria, GTP => riboflavin/FMN/FAD
Metabolism of cofactors and vitamins	Cofactor and vitamin metabolism	M00916	Pyridoxal-P biosynthesis, R5P + glyceraldehyde-3P + glutamine => pyridoxal-P
Metabolism of cofactors and vitamins	Cofactor and vitamin metabolism	M00119	Pantothenate biosynthesis, valine/L-aspartate => pantothenate
Metabolism of cofactors and vitamins	Cofactor and vitamin metabolism	M00120	Coenzyme A biosynthesis, pantothenate => CoA
Metabolism of cofactors and vitamins	Cofactor and vitamin metabolism	M00883	Lipoic acid biosynthesis, animals and bacteria, octanoyl-ACP => dihydrolipoyl-H => dihydrolipoyl-E2
Metabolism of cofactors and vitamins	Cofactor and vitamin metabolism	M00123	Biotin biosynthesis, pimeloyl-ACP/CoA => biotin
Metabolism of cofactors and vitamins	Cofactor and vitamin metabolism	M00577	Biotin biosynthesis, BioW pathway, pimelate => pimeloyl-CoA => biotin
Metabolism of cofactors and vitamins	Cofactor and vitamin metabolism	M00126	Tetrahydrofolate biosynthesis, GTP => THF
Metabolism of cofactors and vitamins	Cofactor and vitamin metabolism	M00880	Molybdenum cofactor biosynthesis, GTP => molybdenum cofactor
Metabolism of cofactors and vitamins	Cofactor and vitamin metabolism	M00988	PreQ1 biosynthesis, GTP => 7-Aminomethyl-7-deazaguanine
Metabolism of cofactors and vitamins	Cofactor and vitamin metabolism	M00140	C1-unit interconversion, prokaryotes
Metabolism of cofactors and vitamins	Cofactor and vitamin metabolism	M00926	Heme biosynthesis, bacteria, glutamyl-tRNA => coproporphyrin III => heme
Metabolism of cofactors and vitamins	Cofactor and vitamin metabolism	M00116	Menaquinone biosynthesis, chorismate (+ polyprenyl-PP) => menaquinol
Metabolism of cofactors and vitamins	Cofactor and vitamin metabolism	M00364	C10-C20 isoprenoid biosynthesis, bacteria
Nucleotide metabolism	Purine metabolism	M00048	De novo purine biosynthesis, PRPP + glutamine => IMP
Nucleotide metabolism	Purine metabolism	M00049	Adenine ribonucleotide biosynthesis, IMP => ADP,ATP
Nucleotide metabolism	Purine metabolism	M00050	Guanine ribonucleotide biosynthesis, IMP => GDP,GTP
Nucleotide metabolism	Purine metabolism	M00053	Deoxyribonucleotide biosynthesis, ADP/GDP/CDP/UDP => dATP/dGTP/dCTP/dUTP
Nucleotide metabolism	Pyrimidine metabolism	M00052	Pyrimidine ribonucleotide biosynthesis, UMP => UDP/UTP,CDP/CTP
Nucleotide metabolism	Pyrimidine metabolism	M00938	Pyrimidine deoxyribonucleotide biosynthesis, UDP => dTTP

Supplementary Fig 5. KEGG modules shown in Fig 4 are identified. Included are modules from the Super Categories glycan metabolism, lipid metabolism, metabolism of cofactor and vitamins, nucleotide metabolism and the Category drug resistance.



Supplementary Fig 6. *S. epidermidis* YSMAS1_1_C10 possesses a large chromosomal deletion which contains the staphylopine and biotin biosynthetic genes. The four representative genomes of S. epidermidis were aligned and visualised using Mauve (2015-02-25). The deletion was identified by the dashed lines. Similar genome sections were identified by coloured shadings.

Supplementary Table 1. Multilocus sequence typing (MLST) allelic profiles of representative *S. epidermidis* isolates. Listed for each isolate are the detected sequence type (ST), allele types of each locus used in the relevant MLST schemes (www.pubmlst.org).

S. epidermidis (scheme: sepidermidis)

Isolate	Cluster	ST	ArcC	AroE	Gtr	MutS	PyrR	TpiA	YqiL
YSMAS1_1_E3	SE4	297	1	2	2	2	2	1	3
YSMAS1_1_E10	SE3	621	7	1	1	6	2	1	1
YSMAS1_1_C8	SE2	225	1	13	7	2	2	1	29
YSMAS1_1_C10	SE1	179	1	2	2	2	1	1	1

Supplementary Table 2. Multilocus sequence typing (MLST) allelic profiles of representative *S. hominis* isolates. Listed for each isolate are the detected sequence type (ST), allele types of each locus used in the relevant MLST schemes (www.pubmlst.org).

S. hominis (scheme: shominis)

Isolate	Cluster	ST	ArcC	GlpK	Gtr	Pta	TpiA	Tuf
YSMAS1_1_A7	SH2	66	6	5	3	6	6	3
YSMAS1_1_D4	SH1	126	18	15	7	1	6	3

Supplementary Table 3: Non-core gene clusters from isolate cluster SE1 identified by the gene cluster number, location, COG category and description. The number of genomes within SE1 and whether or not the gene cluster is present in the representative genome S1_1_C10 were indicated.

Gene Cluster	Number of Genomes (Total = 53)	Present in S1_1_C10	Location	COG Category	Description
GC_00002060	52	Yes	Chromosome	S	Bifunctional serine threonine kinase and phosphorylase involved in the regulation of the pyruvate, phosphate dikinase (PPDK) by catalyzing its phosphorylation dephosphorylation
GC_00002061	52	Yes	Chromosome	O	Protein of unknown function (DUF2847)
GC_00002066	48	Yes	Chromosome	L	Belongs to the 'phage' integrase family
GC_00002069	48	Yes	Chromosome	S	D5 N terminal like
GC_00002070	48	Yes	Chromosome	K	Helix-turn-helix XRE-family like proteins
GC_00002072	48	Yes	Chromosome	-	-
GC_00002079	1	No	Plasmid	-	-
GC_00002082	1	No	Plasmid	L	Initiator Replication protein
GC_00002083	1	No	Plasmid	S	Possible lysine decarboxylase
GC_00002084	1	No	Plasmid	-	-
GC_00002086	1	No	Plasmid	-	-
GC_00002087	1	No	Plasmid	-	-
GC_00002088	1	No	Plasmid	-	-
GC_00002092	1	No	Plasmid	-	-
GC_00002093	1	No	Plasmid	-	-

Supplementary Table 4: Non-core gene clusters from isolate cluster SE2 identified by the gene cluster number, location, COG category and description. The number of genomes within SE2 and whether or not the gene cluster is present in the representative genome S1_1_C8 were indicated.

Gene Cluster	Number of Genomes (Total = 15)	Present in S1_1_C8	Location	COG Category	Description
GC_00002225	14	Yes	Chromosome	S	Phage portal protein, SPP1 Gp6-like
GC_00002227	8	No	Plasmid	-	-
GC_00002228	8	No	Plasmid	-	-
GC_00002229	8	No	Plasmid	L	Initiator Replication protein
GC_00002230	8	No	Plasmid	-	-
GC_00002231	8	No	Plasmid	-	-
GC_00002234	6	No	Plasmid	D	Plasmid recombination enzyme
GC_00002235	6	No	Plasmid	S	Protein of unknown function (DUF3139)
GC_00002237	6	No	Plasmid	K	Cro/C1-type HTH DNA-binding domain
GC_00002240	5	No	Plasmid	S	Possible lysine decarboxylase
GC_00002241	5	No	Plasmid	-	-
GC_00002243	5	No	Plasmid	-	-
GC_00002247	5	No	Plasmid	-	-
GC_00002248	3	No	Plasmid	L	Replication initiation factor
GC_00002251	1	No	Plasmid	N	Hep Hag repeat protein

Supplementary Table 5: Non-core gene clusters from isolate cluster SE3 identified by the gene cluster number, location, COG category and description. The number of genomes within SE3 and whether or not the gene cluster is present in the representative genome S1_1_E10 were indicated.

Gene Cluster	Number of Genomes (Total = 6)	Present in S1_1_E10	Location	COG Category	Description
GC_00002156	2	No	Plasmid	L	Initiator Replication protein
GC_00002160	2	No	Plasmid	-	-
GC_00002161	2	No	Plasmid	-	-
GC_00002162	2	No	Plasmid	-	-
GC_00002163	2	No	Plasmid	D	Plasmid recombination enzyme
GC_00002164	1	No	Plasmid	L	Replication initiation factor

Supplementary Table 6: Non-core gene clusters from isolate cluster SE4 identified by the gene cluster number, location, COG category and description. The number of genomes within SE4 and whether or not the gene cluster is present in the representative genome S1_1_E3 were indicated.

Gene Cluster	Number of Genomes (Total = 16)	Present in S1_1_E3	Location	COG Category	Description
GC_00002218	4	Yes	Chromosome	-	-
GC_00002219	4	Yes	Chromosome	T	Domain of unknown function (DUF4118)
GC_00002220	4	Yes	Chromosome	S	SEFIR domain
GC_00002222	4	Yes	Chromosome	-	-
GC_00002223	4	Yes	Chromosome	-	-
GC_00002224	4	Yes	Chromosome	T	catabolite gene activator and regulatory subunit of cAMP-dependent protein kinases
GC_00002225	4	Yes	Chromosome	E	Pyridoxal-dependent decarboxylase conserved domain
GC_00002227	4	Yes	Chromosome	P	ATP-driven potassium transport (or Kdp) system. This subunit is responsible for energy coupling to the transport system
GC_00002228	4	Yes	Chromosome	-	-
GC_00002229	4	Yes	Chromosome	-	-
GC_00002231	4	Yes	Chromosome	P	ATP-driven potassium transport (or Kdp) system. This subunit acts as a catalytic chaperone that increases the ATP- binding affinity of the ATP-hydrolyzing subunit KdpB by the formation of a transient KdpB KdpC ATP ternary complex
GC_00002232	4	Yes	Chromosome	L	RePlication protein
GC_00002233	4	Yes	Chromosome	KT	Transcriptional regulatory protein, C terminal
GC_00002234	4	Yes	Chromosome	V	ATPases associated with a variety of cellular activities
GC_00002236	4	Yes	Chromosome	P	ATP-driven potassium transport (or Kdp) system. This subunit binds and transports the potassium across the cytoplasmic membrane
GC_00002237	4	Yes	Chromosome	T	Histidine kinase-, DNA gyrase B-, and HSP90-like ATPase
GC_00002238	4	Yes	Chromosome	P	MFS/sugar transport protein
GC_00002241	4	Yes	Chromosome	P	P-type ATPase
GC_00002243	4	Yes	Chromosome	S	4Fe-4S single cluster domain
GC_00002244	4	Yes	Chromosome	K	Transcriptional regulatory protein, C terminal
GC_00002245	4	Yes	Chromosome	M	Protein of unknown function (DUF1541)
GC_00002247	1	No	Plasmid	-	-
GC_00002248	1	No	Plasmid	-	-
GC_00002249	1	No	Plasmid	-	-
GC_00002250	1	No	Plasmid	-	-
GC_00002252	1	No	Plasmid	D	Plasmid recombination enzyme

Supplementary Table 7: Non-core gene clusters from isolate cluster S identified by the gene cluster number, location, COG category and description. The number of genomes within S and whether or not the gene cluster is present in the representative genome A1_1_H3st were indicated.

Gene Cluster	Number of Genomes (Total = 27)	Present in A1_1_H3st	Location	COG Category	Description
GC_00000047	26	Yes	Chromosome	L	Replication initiation factor
GC_00000055	26	Yes	Plasmid	-	-
GC_00000074	26	Yes	Chromosome	M	C-terminus of bacterial fibrinogen-binding adhesin
GC_00002340	26	Yes	Chromosome	S	YoeB-like toxin of bacterial type II toxin-antitoxin system
GC_00002342	26	Yes	Chromosome	S	Domain of unknown function (DUF4870)
GC_00002344	26	Yes	Chromosome	S	Pfam:DUF867
GC_00002345	26	Yes	Chromosome	P	ABC transporter
GC_00002346	26	Yes	Chromosome	C	Catalyzes the reductive cleavage of azo bond in aromatic azo compounds to the corresponding amines. Requires NADH, but not NADPH, as an electron donor for its activity
GC_00002347	26	Yes	Chromosome	M	Bacterial capsule synthesis protein PGA_cap
GC_00002348	26	Yes	Chromosome	D	Antitoxin component of a toxin-antitoxin (TA) module
GC_00002351	26	Yes	Chromosome	K	Transcriptional regulator
GC_00002352	26	Yes	Chromosome	S	Poly-gamma-glutamate hydrolase
GC_00002353	26	Yes	Chromosome	M	Bacterial capsule synthesis protein PGA_cap
GC_00002354	26	Yes	Chromosome	GH	Belongs to the TPP enzyme family
GC_00002355	26	Yes	Chromosome	K	regulator
GC_00000066	19	Yes	Plasmid	U	Relaxase/Mobilisation nuclease domain
GC_00002360	17	Yes	Plasmid	S	Bacterial mobilisation protein (MobC)
GC_00002363	17	Yes	Plasmid	K	Cupin domain
GC_00002364	17	Yes	Plasmid	-	-
GC_00002365	17	Yes	Plasmid	-	-
GC_00002366	17	Yes	Plasmid	M	Acetyltransferase (GNAT) domain
GC_00002367	17	Yes	Plasmid	L	Replication protein
GC_00002368	17	Yes	Plasmid	L	Resolvase, N terminal domain
GC_00002370	17	Yes	Plasmid	S	Protein of unknown function, DUF536
GC_00002356	9	No	Plasmid	L	RePlication protein
GC_00002371	9	No	Plasmid	S	Staphylococcus haemolytic protein
GC_00002372	8	Yes	Plasmid	S	Bacterial mobilisation protein (MobC)

Supplementary Table 8. Antimicrobial resistance cassettes identified in all representative isolates using ResFinder.

Isolate	1	2	3	4
S. capitis YSMAA1_1_H3st	vga(A)LC, DQ823382, [lincomycin, clindamycin, dalfopristin, pristinamycin iia, virginiamycin m, tiamulin]			
S. hominis YSMAS1_1_A7	mecA, BX571856, [amoxicillin, ampicillin, cefepime, cefixime, cefotaxime, cefoxitin, ceftazidime, ertapenem, imipenem, meropenem, piperacillin]	blaZ, NZ_JVAT01000021, [amoxicillin, ampicillin, piperacillin, penicillin]	fusC, KF527883, [fusidic acid]	
S. hominis YSMAS1_1_D4	blaZ, CP003979, [amoxicillin, ampicillin, piperacillin, penicillin]			
S. epidermidis YSMAS1_1_C8	fosB, CP000029, [fosfomycin]			
S. epidermidis YSMAS1_1_C10	fosB, CP000029, [fosfomycin]	fusB, AY373761, [fusidic acid]		
S. epidermidis YSMAS1_1_E3	fosB, ACHE01000077, [fosfomycin]	fusC, KF527883, [fusidic acid]	msr(A), X52085, [erythromycin, azithromycin, telithromycin, quinupristin, pristinamycin ia, virginiamycin s]	mph(C), AF167161, [erythromycin, spiramycin, telithromycin]
S. epidermidis YSMAS1_1_E10	fosB, CP000029, [fosfomycin]	fusB, AY373761, [fusidic acid]		

3.15 References

1. Grice EA, Segre JA. The skin microbiome. Nature Reviews Microbiology 2011 9:244–53.
2. Findley K, Oh J, Yang J, Conlan S, Deming C, Meyer JA, et al. Topographic diversity of fungal and bacterial communities in human skin. Nature 2013 498:367–70.
3. Rudden M, Herman R, Rose M, Bawdon D, Cox DS, Dodson E, et al. The molecular basis of thioalcohol production in human body odour. Sci Rep 2020 10:12500.
4. Byrd AL, Belkaid Y, Segre JA. The human skin microbiome. Nature Reviews Microbiology 2018 16:143–55
5. Rosenstein R, Torres Salazar BO, Sauer C, Heilbronner S, Krismer B, Peschel A. The Staphylococcus aureus-antagonizing human nasal commensal Staphylococcus lugdunensis depends on siderophore piracy. Microbiome 2024 12:213
6. Zheng Y, Hunt RL, Villaruz AE, Fisher EL, Liu R, Liu Q, et al. Commensal Staphylococcus epidermidis contributes to skin barrier homeostasis by generating protective ceramides. Cell Host Microbe 2022 30:301-313.e9.
7. Naik S, Bouladoux N, Wilhelm C, Molloy MJ, Salcedo R, Kastenmuller W, et al. Compartmentalized control of skin immunity by resident commensals. Science 2012 337:1115–1119.
8. Naik S, Bouladoux N, Linehan JL, Han SJ, Harrison OJ, Wilhelm C, et al. Commensal–dendritic-cell interaction specifies a unique protective skin immune signature. Nature 2015 520: 104–108.
9. Kim P Il, Sohng JK, Sung C, Joo HS, Kim EM, Yamaguchi T, et al. Characterization and structure identification of an antimicrobial peptide, hominicin, produced by Staphylococcus hominis MBBL 2-9. Biochem Biophys Res Commun 2010 399:133–138.
10. Lynch D, O'Connor PM, Cotter PD, Hill C, Field D, Begley M. Identification and characterisation of capidermicin, a novel bacteriocin produced by Staphylococcus capitis. PLoS One 2019 14:e0223541
11. Zipperer A, Konnerth MC, Laux C, Berscheid A, Janek D, Weidenmaier C, et al. Human commensals producing a novel antibiotic impair pathogen colonization. Nature 2016 535: 511–516.

12. Bawdon D, Cox DS, Ashford D, James AG, Thomas GH. Identification of axillary Staphylococcus sp. involved in the production of the malodorous thioalcohol 3-methyl-3-sufanylhexan-1-ol. FEMS Microbiol Lett. 2015 362:1–10.
13. Minhas GS, Bawdon D, Herman R, Rudden M, Stone AP, Gordon James A, et al. Structural basis of malodour precursor transport in the human axilla. Elife 2018 7:e34995
14. Herman R, Kinniment-Williams B, Rudden M, James AG, Wilkinson AJ, Murphy B, et al. Identification of a Staphylococcal dipeptidase involved in the production of human body odour. Journal of Biological Chemistry 2024 300:107928
15. Grimaldi C, Gilardeau S, Verzeaux L, Guedj O, Richer S, Laspoussas A, et al. Identification of a higher C–S lyase activity of Staphylococcus hominis in volunteers with unpleasant axillary odour. Int J Cosmet Sci 2024 46:1088–1096
16. Taylor D, Daulby A, Grimshaw S, James G, Mercer J, Vaziri S. Characterization of the microflora of the human axilla. Int J Cosmet Sci 2003 25:137–145
17. Leyden JJ, McGinley KJ, Holzle E, Labows JN, Kligman AM. The Microbiology of the Human Axilla and Its Relationship to Axillary Odor. Journal of Investigative Dermatology 1981 77:413–416.
18. Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, et al. Topographical and temporal diversity of the human skin microbiome. Science 2009 324:1190–1192.
19. Oh J, Byrd AL, Deming C, Conlan S, Kong HH, Segre JA, et al. Biogeography and individuality shape function in the human skin metagenome. Nature 2014 514:59–64.
20. Agustinho DP, Fu Y, Menon VK, Metcalf GA, Treangen TJ, Sedlazeck FJ. Unveiling microbial diversity: harnessing long-read sequencing technology. Nature Methods 2024 21:954–966.
21. Maiden MCJ, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, et al. Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci U S A 1998 95:3140–3145.
22. Pritchard L, Cock P, Esen Ö. pyani v0.2.8: average nucleotide identity (ANI) and related measures for whole genome comparisons. Anal Methods 2019 8:12-24.
23. Herman R, Meaden S, Rudden M, Cornmell R, Wilkinson AJ, Murphy B, et al. Revealing the diversity of commensal corynebacteria from a single human skin site. biorxiv. 2024 Nov 28

24. Olm MR, Brown CT, Brooks B, Banfield JF. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. ISME J 2017 11:2864–2868.
25. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 2015 31:3691–3693.
26. Eren AM, Kiefl E, Shaiber A, Veseli I, Miller SE, Schechter MS, et al. Community-led, integrated, reproducible multi-omics with anvi'o. Nature Microbiology 2020 6:3–6.
27. Cantalapiedra CP, Herņandez-Plaza A, Letunic I, Bork P, Huerta-Cepas J. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. Mol Biol Evol 2021 38:5825–5829.
28. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, et al. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res 2019 47:D309–D314.
29. Reslane I, Halsey CR, Stastny A, Cabrera BJ, Ahn J, Shinde D, et al. Catabolic Ornithine Carbamoyltransferase Activity Facilitates Growth of Staphylococcus aureus in Defined Medium Lacking Glucose and Arginine. mBio 2022 13:e00395-22.
30. Satiaputra J, Eijkelkamp BA, McDevitt CA, Shearwin KE, Booker GW, Polyak SW. Biotin-mediated growth and gene expression in Staphylococcus aureus is highly responsive to environmental biotin. Appl Microbiol Biotechnol 2018 102:3793–803.
31. Grim KP, San Francisco B, Radin JN, Brazel EB, Kelliher JL, Párraga Solórzano PK, et al. The metallophore staphylopine enables staphylococcus aureus to compete with the host for zinc and overcome nutritional immunity. mBio 2017 8:10-1128.
32. Grim KP, San Francisco B, Radin JN, Brazel EB, Kelliher JL, Párraga Solórzano PK, et al. The Metallophore Staphylopine Enables Staphylococcus aureus To Compete with the Host for Zinc and Overcome Nutritional Immunity. mBio 2017 8:e01281-17.
33. Turner NA, Sharma-Kuinkel BK, Maskarinec SA, Eichenberger EM, Shah PP, Carugati M, et al. Methicillin-resistant Staphylococcus aureus: an overview of basic and clinical research. Nature Reviews Microbiology 2019 17: 203–218.
34. Kreiswirth B, Kornblum J, Arbeit RD, Eisner W, Maslow JN, McGeer A, et al. Evidence for a Clonal Origin of Methicillin Resistance in Staphylococcus aureus. Science 1993 259:227–230.

35. Mitchell BA, Brown MH, Skurray RA. QacA multidrug efflux pump from Staphylococcus aureus: Comparative analysis of resistance to diamidines, biguanidines, and guanylhydrazones. Antimicrob Agents Chemother 1998 42:475–477.
36. Payne LJ, Meaden S, Mestre MR, Palmer C, Toro N, Fineran PC, et al. PADLOC: a web server for the identification of antiviral defence systems in microbial genomes. Nucleic Acids Res 2022 50:W541–50.
37. Vassallo CN, Doering CR, Littlehale ML, Teodoro GIC, Laub MT. A functional selection reveals previously undetected anti-phage defence systems in the E. coli pangenome. Nature Microbiology 2022 7: 1568–1579.
38. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. Assembly of long, error-prone reads using repeat graphs. Nature Biotechnology 2019 37: 540–546.
39. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics 2014 30:2068–2069.
40. Richter M, Rosselló-Móra R. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A 2009 106:19126–19131
41. Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. Analytical Methods 2015 8:12–24.
42. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, et al. Mash: Fast genome and metagenome distance estimation using MinHash. Genome Biol 2016 17:1–14
43. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, et al. Versatile and open software for comparing large genomes. Genome Biol 2004 5:1–9.
44. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 2015 31:3691–3693
45. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Syst Biol 2010 59:307–321
46. Kanehisa M, Sato Y, Kawashima M. KEGG mapping tools for uncovering hidden features in biological data. Protein Sci 2022 31:47–53.
47. Kanehisa M, Sato Y. KEGG Mapper for inferring cellular functions from protein sequences. Protein Sci 2020 29:28–35

Chapter 4

Corynebacterial diversity within the human axilla

4.1 Manuscript information

Revealing the diversity of commensal corynebacteria from a single human skin site

Reyme Herman[1,5], Sean Meaden[1], Michelle Rudden[2], Robert Cornmell[3], Holly N. Wilkinson[2], Matthew J. Hardman[2], Anthony J. Wilkinson[4], Barry Murphy[3], Gavin H. Thomas[1,5*]

[1] Department of Biology, University of York, Wentworth Way, York, YO10 5DD, UK
[2] Centre for Biomedicine, Hull York Medical School, University of Hull, Hull, HU6 7RX, UK
[3] Unilever Research & Development, Port Sunlight Laboratory, Quarry Road East, Bebington, Wirral, Merseyside, CH63 3JW, UK
[4] York Structural Biology Laboratory, Department of Chemistry, University of York, Wentworth Way, York, YO10 5DD, UK
[5] York Biomedical Research Institute, University of York, YO10 5DD, UK

Manuscript under review with mSystems
Reviewer's comments have been addressed and incorporated

Preprint DOI: 10.1101/2024.11.28.625817v3

Post viva status: Manuscript was published with mSystems on the 25th of September 2025

DOI: 10.1128/msystems.00459-25
Co-authors have given permission for this manuscript to be used in this thesis.

Post-submission of publication in mSystems, the following error was found:
Page 146 – "The likely presence of antimicrobials on the human skin may promote the horizontal acquisition of antimicrobial resistance genes…" (Corrected)

4.2 Context within thesis (Abstract)

Alongside the *Staphylococcus* isolates discussed in Chapter 3, *Corynebacterium* isolates were also isolated from the same axilla for the purposes of malodour generation assessments. A similar genetic characterisation and analysis workflow was carried out on isolates of the lesser studied genus *Corynebacterium*. For comparison, this study also assessed a subset of *Corynebacterium* isolates from three other volunteers. A total of 16 genetically diverse isolates spanning 7 species of *Corynebacterium* were derived from the same axilla assessed in Chapter 3. These isolates were then used in the work presented in Chapter 5.

4.3 Abstract

Our understanding of the skin microbiome has significantly advanced with the rise of sequencing technologies. While *Corynebacterium* species are a major component of the human skin microbiome, research has largely centred on other prevalent genera like *Staphylococcus* and *Cutibacterium*. Prior to this study, complete genomes for skin-associated *Corynebacterium* were limited. For instance, only nine were available for the commonly identified species *C. tuberculostearicum*. In this study, we explored *Corynebacterium* diversity from a single skin site, the axilla, using selective media to enrich for these bacteria. Long-read whole genome sequencing and bioinformatic analysis of the enriched isolates provided unprecedented insight into the diversity of this genus at a single body site. The study yielded 215 closed genomes, comprising 30 distinct representative genomes following dereplication. These genomes span seven distinct species, including two new species provisionally named *C. axilliensis* and *C. jamesii*, as well as species not previously linked to the skin. Pangenome analysis of these isolates uncovered potential metabolic differences, antimicrobial resistance genes, novel biosynthetic gene clusters (BGCs), prophages, and phage defence systems. This study represents the most detailed analysis to date of *Corynebacterium* from a single skin niche and highlights extensive variation even within a single host. Our culture-based, Nanopore sequencing approach has expanded the number of publicly available skin *Corynebacterium* genomes, providing a valuable resource for future studies investigating the diversity and function of this important skin genus.

4.4 Importance

This study uncovers the hidden diversity of *Corynebacterium*, an important yet often overlooked group of bacteria inhabiting human skin. Focusing on the underarm and using advanced sequencing techniques, we generated over 200 complete bacterial genomes. This collection includes species not previously known to reside on the skin, as well as two entirely new species, highlighting how little is currently known about this cutaneous genus. Most of these bacteria were isolated from a single individual, underscoring the vast microbial diversity that can exist within one person. By closely examining the corynebacterial community at a single site, we begin to uncover the complex relationships within a true microbial ecosystem. These findings deepen our understanding of the skin microbiome and provide a valuable resource for future research into how these microbes affect skin health, hygiene, and disease.

4.5 Introduction

Exploration of the human skin microbiome allows deeper understanding of how commensal microbes and the host interact to maintain skin health. Human skin, generally a nutrient limited environment, hosts a variety of commensal microbes spanning multiple genera. Skin sites can be characterised as dry, moist or oily (sebaceous), depending on the architecture of the skin and abundance of associated glands, namely eccrine, sebaceous or apocrine glands. These different environmental conditions are known to play a major role in determining microbial diversity (1). Although *Cutibacterium* and *Staphylococcus* are widely reported as major commensal genera, *Corynebacterium* are also highly abundant (2, 3). For example, Grice *et al*., report that corynebacteria comprise 22.8% of 112,000 16S rRNA sequences derived from 20 different skin sites, second only to cutibacteria (23%) (4). Some corynebacterial species have been found to be lipophilic and hence require lipids derived from the sebaceous gland or the stratum corneum (5, 6). The commonly identified commensal corynebacteria, *C. jeikeium* (7), *C. amycolatum* (8) and *C. kroppenstedtii* (9) have been suggested to be opportunistic pathogens. *C. accolens* (10) and *C. tuberculostearicum* (11) have been shown to interact with the human immune system, causing inflammation. *C. accolens* has also been shown to inhibit the growth of *Streptococcus pneumoniae* (12, 13) while *C. striatum* is proposed to be able

to shift *Staphylococcus aureus* from a virulent to commensal state (14), thus demonstrating its potential protective capabilities. Together, these studies highlight the importance of the skin microbiome in protecting against pathogens and maintaining skin homeostasis.

Human malodour is generated by microbial metabolism of a precursor secreted by the apocrine glands. Specific skin-associated bacteria, such as *Staphylococcus* species, convert these odorless precursors into volatile compounds responsible for body odour. Although a monophyletic group of staphylococci was attributed to this phenomenon (15), some studies have proposed a role for axillary *Corynebacterium* in human malodour formation due to the correlation of abundance and odour profiles (16) and the presence of enzymatic components for the biotransformation of the human derived precursor (17–19). Further, a genus level study analysing the diversity of axillary commensals in 9 individuals found that 59.5% of sequences were binned to the genus *Corynebacterium* (20). However, our understanding of the role of cutaneous corynebacteria in human malodour remains unclear.

Despite the abundance of corynebacteria on the human skin, the amount of genetic information on cutaneous corynebacterial species is surprisingly lacking, suggesting an incomplete understanding of true diversity within the skin microbiome. In the case of *C. tuberculostearicum*, a species widely found on various skin sites, only 9 complete genomes were publicly available on GenBank prior to this study. In comparison, there are 306 available complete genomes of another skin commensal, *Staphylococcus epidermidis* (data from 28 March 25). Two recent comparative genomics studies have sought to address the lack of genetic information through a culture-based approach. The first focusses on skin isolates of the *C. tuberculostearicum* species complex which is made up of multiple closely related species including *C. kefirresidentii* (21). In addition, the study reports the identification of a novel species of corynebacteria that was predominantly found on the toe web and toenails of multiple individuals. A second comparative genomics study on human skin corynebacteria (22) revealed the diversity of corynebacteria on human facial skin. This study also identified two novel species of corynebacteria.

Here, we sought to elucidate the diversity of corynebacteria within the human underarm (axilla). This is a moist skin site enriched with metabolites derived from the three skin-associated glands as well as the epidermis, making it a relatively nutrient rich environment compared to most other skin sites. In this study we used a culture-based approach followed by sequencing, consistent with recent comparative genomics studies. However, those studies typically began with 16S rRNA (V1-V3 region) sequencing to identify representative taxa within amplicon sequence variants (ASVs) for subsequent whole genome sequence (WGS). We instead adopted a direct-to long-read whole genome sequencing (WGS) strategy by omitting the 16S rRNA gene sequencing which could introduce sampling bias. In this study, we generated complete genomes of over 200 corynebacteria isolates derived from axillary swabs of four individuals using the Oxford Nanopore PromethION platform. Our sequencing approach revealed 7 distinct species of axillary corynebacteria, two of which are proposed to be novel species. Through comprehensive genomic analyses, we revealed the genetic and functional diversity within our collection of axillary corynebacteria, paving the way to revealing their roles in the skin microbiome.

4.6 Results

4.6.1 High-resolution genomic characterisation reveals corynebacterial diversity in the human axilla

Axillary swabs from 4 human volunteers (A1, A3, A4 and A5) were collected and then enriched for corynebacteria on selective agar. To assess the ability of our approach to capture the diversity of corynebacteria within our isolates, we generated complete genome and plasmid sequences for 154 corynebacteria isolates from a single individual A1 and complemented the study with small subsets of isolates from the three other volunteers; 22 isolates from A3, 26 isolates from A4 and 13 isolates from A5 (**Fig. 1A**).

In order to initially assess the diversity of the corynebacterial isolates from all 4 volunteers, [23] we compared pairwise average nucleotide identities (ANI) between assembled genomes from each individual (**Fig. 1B-E**) and collectively (**Supplementary Fig. 1**). The pairwise percentage identity plots for the 154 isolates

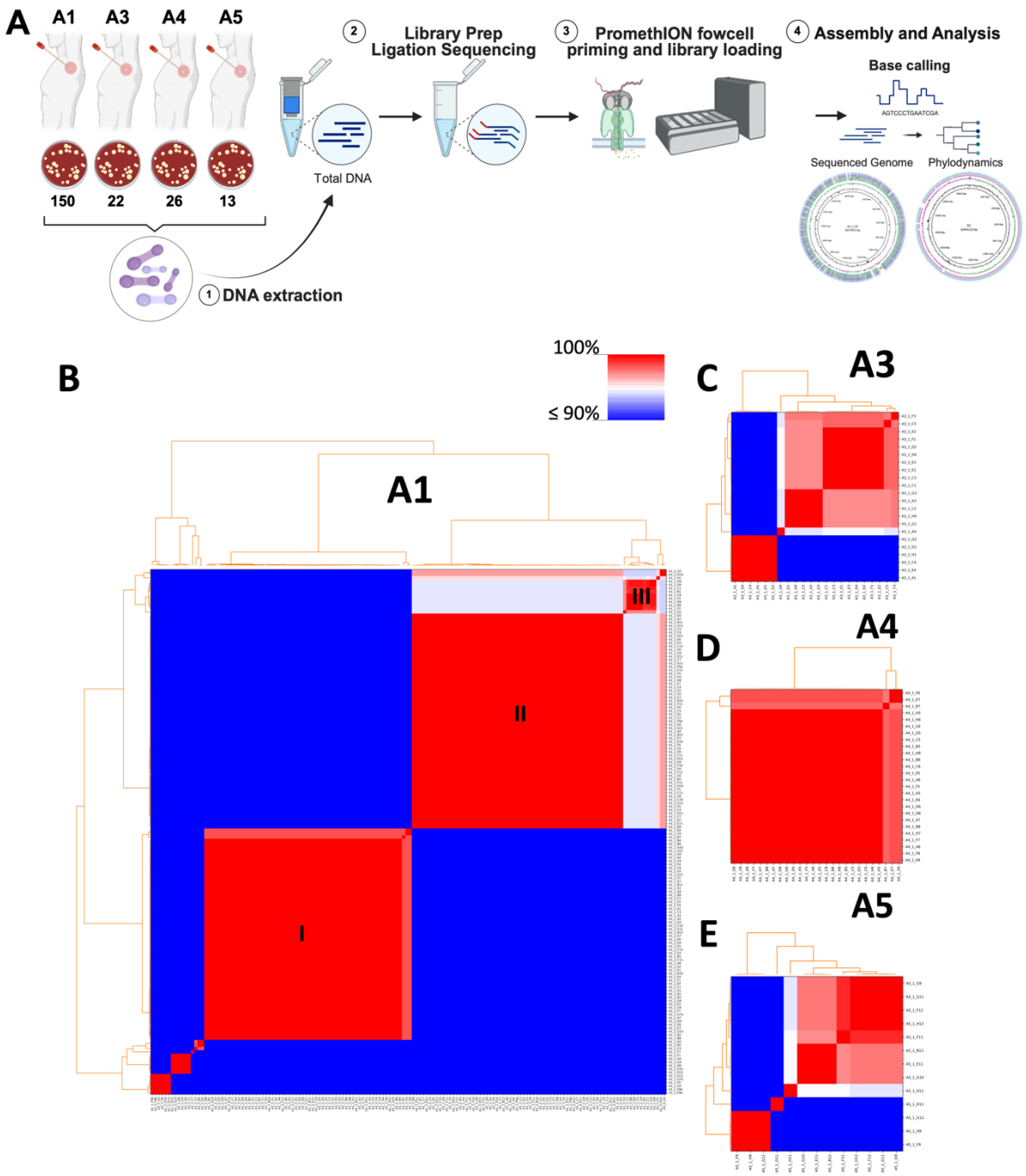


Figure 1. Capturing the diversity of axillary corynebacteria. (A) Axillary *Corynebacterium* isolates were harvested from swabs from 4 volunteers. Long-read WGS and downstream analysis was performed on 154 isolates of a volunteer A1, 22 from volunteer A3, 26 from volunteer A4 and 13 from volunteer A5. (B-E) Pairwise average nucleotide identities (ANI) were calculated with pyANI and used to assess the genetic diversity of the isolates. ANI values were placed onto a coloured pairwise matrix according to the colour scale depicted in (B) for isolates derived from each volunteer. Three major groups of isolates were identified with I, II and III from the isolates from A1. Images in (A) were created with Biorender.com.

from A1 revealed that the majority of isolates recovered by culture belonged to two discrete groups (labelled as I and II) (**Fig. 1B**). However, the depth of sampling also revealed a third cluster (group III) and number of other smaller groups, some only containing one isolate, suggesting the approach had recovered significant diversity from this single body site from a single individual. For individuals A3 and A5 there was a similar pattern of a number of different clusters identified, despite there being a much smaller total number of genomes (22 and 13 isolates respectively) (**Fig. 1C** and **Fig. 1E**). Surprisingly in individual A4 we found that almost all of the 26 isolates were very similar to each other (ANI = >95%), suggesting that one species dominates this community (**Fig. 1D**).

Using the drep workflow (23), we identified representative genomes from groups of genetically similar isolates. Genome-wide ANI values of <95% initially classify isolates into different species, while further refinement using ANI values of 95%–99.5% distinguishes between different strains (**Supplementary Fig. 2**). The chosen ANI thresholds are consistent with recent studies on the diversity of cutaneous corynebacteria (21, 22). The isolates bin to 7 different species with 30 identified strains which we focussed further analysis on. Typically, 16S rRNA (V1-V3 region) sequences are analysed for species-level identification of isolates. Using this method, we observe only 5 distinct clades instead of 7 (Supplementary Fig. 3A). Further, the 16S rRNA (V1-V3 region) sequences of the 24 isolates forming the largest clade were found to be ≥99.8% identical to three closely related but ultimately different species of corynebacteria (Supplementary Fig. 3A, red shading). This highlights the limitations of 16S rRNA analysis in accurately resolving species-level differences. Hereafter, the representative genomes will be referred to with the prefix YSMA (**Y**ork **S**kin **M**icrobiome: Study **A**).

While the 16S information alone could not accurate speciate our isolates, we were able to achieve this with the closed genomes using GTDB-tk (24, 25) (**Supplementary Fig. 3B**, **Supplementary Table 1**). This analysis, further classified the 24 isolates into three separate species: *C. kefirresidentii, C. tuberculostearicum* and *C. aurimucosum*_E. These three species are commonly found on the skin and are thought to be part of the *C. tuberculostearicum* species complex which was previously

reported (21, 26), confirming the broader significance of these organisms across multiple human skin sites. ANI analysis of the GTDB representative for the species *C. aurimucosum*_E suggest that the species is misassigned and matches best with *C. marquesiae* (27). From this point onward, we classify our isolates that matched with *C. aurimucosum*_E on GTDB as *C. marquesiae*. Although two isolates (YSMAA1_1_B1 and YSMAA5_1_F11) matched closest with *C. tuberculostearicum* and *C. tuberculostearicum_C*, they were not classified to a specific species due to the predefined ANI reference radius of the analysis. As they share more than 95% identity to their closest matches, we classified both as *C. tuberculostearicum*.

YSMAA1_2_B5 and YSMAA1_2_C5 were classified as *C. amycolatum* which have been identified not only on the skin but also specifically in the human axilla (28, 29)(**Supplementary Table 1**). Interestingly, YSMAA1_1_F6 was identified to be *C. gottingense* which was first reported in 2017 (30) but until now, not being associated to the human skin. Finally, three isolates (YSMAA1_1_D6, YSMAA5_1_F9 and YSMAA1_1_F7) were matched with metagenome derived genomes of uncultured isolates suggesting that these are two new species. Isolates YSMAA1_1_D6 and YSMAA5_1_F9, tentatively named *C. axilliensis*, are most closely related to *C. singulare* with an average nucleotide identity of ~84% calculated using FastANI (31) (**Supplementary Fig. 4A**). *C. axilliensis* isolates were found on individuals A1 (~4% of corynebacteria genomes) and A5 (~23% of corynebacteria genomes). Isolate YSMAA1_1_F7, tentatively named *C. jamesii*, is most similar to *C. jeikeium* with an average nucleotide identity of ~87.5% calculated using FastANI (31) (**Supplementary Fig. 4B**).

From the 4 volunteers, our culturomics approach identified 30 genetically distinct isolates spanning 7 species of *Corynebacterium* from an initial pool of 215 isolates (Supplementary Tables 2-5). Using our direct to WGS workflow, we revealed the species diversity of our axillary isolates from the four volunteers which would not be captured if a prior 16S rRNA screen was carried out (*C. tuberculostearicum*, *C. kefirresidentii* and *C. marquesiae*). Alongside other known cutaneous species (including *C. amycolatum*), we were also able to identify species not previously associated to healthy human skin (*C. gottingense*) including two novel species (*C. axilliensis* and *C. jamesii*).

4.6.2 Pangenome and metabolic profiling reveal species- and strain-level functional diversity among cutaneous corynebacteria

We performed pangenome analysis of the 30 representative isolates to understand their genetic differences. Firstly, we used Roary (32) to derive a core genome maximum likelihood SNP tree using a BLASTp cutoff of 80% and seeded the tree with species representatives from GenBank (**Supplementary Fig 5**). Consistent with the GTDB-tk analysis, we clearly see 7 different species including distinctions between the isolates from the species of the *C. tuberculostearicum* species complex and obvious divergence between the two novel species and their nearest related species. To further elucidate the genetic differences between our isolates and their closest GenBank representatives, we performed a pangenome analysis using anvi'o (33) (**Fig. 2**). Gene clusters of interest were then functionally predicted with eggNOG-mapper(34, 35).

The two representative *C. axilliensis* isolates share a specific cluster of genes with the largest proportion of genes with assigned functions found to be associated to transcription (K) or replication, recombination and repair (L) (**Supplementary Fig. 6A**). Interestingly, analysis of the gene clusters specific to either *C. axilliensis* YSMAA5_1_F9 or YSMAA1_1_D6 places the largest proportion of genes with annotated functions in the cell wall/membrane/envelope biogenesis (M) category suggesting differences in their cell envelope (**Supplementary Fig. 6B and 6C**). Functionally annotated gene clusters specific to *C. jamesii* YSMAA1_1_F7 are mainly found in the replication, recombination and repair (L) (**Supplementary Fig. 7**).

As expected, *C. kefirresidentii*, *C. marquesiae* and *C. tuberculostearicum* isolates share a larger number of gene clusters than with other species. To clarify the main differences between the three species, we generated the core genomes for each species from isolates within our collection followed by a pangenome analysis (**Supplementary Fig. 8**) to extract the singletons found within each species for functional analysis. Within the functionally predicted singletons, genes classified to the inorganic ion transport and metabolism (P) category make up the largest proportions in each of the three core genomes (**Supplementary Fig. 9A-C**) consistent with a previous study on skin isolates of this species complex (21).

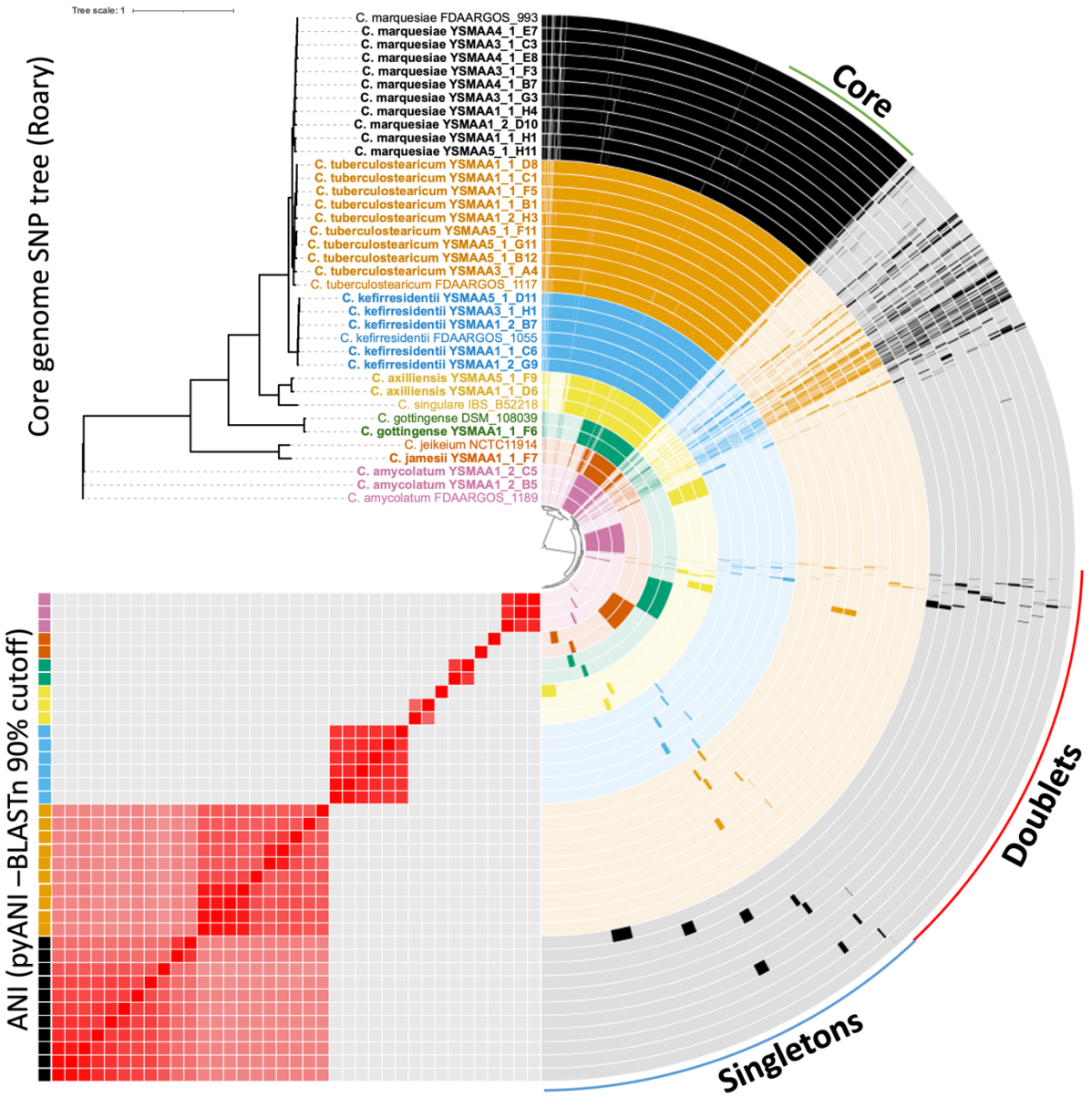


Figure 2. The pangenome of the axillary corynebacteria isolates (bold) with the closest RefSeq representatives for each species calculated and visualised using anvi'o. Gene clusters found in either one (Singletons) or two (Doublets) were identified. A core genome was calculated using Roary with an 80% BLASTp cutoff and visualised on a tree generated using PhyML and visualised on iTOL. An ANI comparison using BLASTn on pyANI with a 90% cutoff was generated and visualised on anvi'o. Isolates of the 7 clusters of species were colour coded (*C. tuberculostearicum* = orange, *C. kefirresidentii* = blue, *C. axilliensis* = gold, *C. gottingense* = green, *C. jamesii* = bronze, *C. marquesiae* = black, *C. amycolatum* = pink).

We further investigated the putative metabolic differences between these isolates by identifying complete metabolic pathways present in each skin isolate using KEGG (**Fig. 3**). As observed in the pangenome analysis, we generally observe KEGG modules clustering with more similar species. Within the *C. tuberculostearicum* species complex isolates, we see conservation of most KEGG modules but the largest differences seem to lie within the amino acid metabolism pathways. Unlike other species within the complex, most *C. maquesiae* isolates do not seem to possess the serine biosynthetic module (M00020) while *C. tuberculostearicum* lack both lysine biosynthetic modules (M00016, M000526) and the ornithine biosynthesis (M00028) modules and *C. kefirresidentii* lacks the tryptophan biosynthetic module (M00023).

The tryptophan biosynthetic module (M00023) seems to also be absent in isolates all of *C. gottingense*, *C. jamesii*, *C. amycolatum* and also *C. axilliensis* YSMAA1_1_D6. Interestingly, the two isolates of *C. axilliensis* seem to possess different metabolic pathway profiles, not limited to tryptophan biosynthesis. Another example is the fatty acid biosynthesis module (M00083), found only in *C. axilliensis* YSMAA5_1_F9 and both *C. amycolatum* isolates.

Another category of metabolic pathways that we observe a large number of differences in is the metabolism of cofactors and vitamins. All but two isolates of *C. kefirresidentii* seem to be biotin auxotrophs while *C. gottingense* and *C. amycolatum* isolates were suggested to be pyridoxal phosphate auxotrophs. Finally, *C. amycolatum* YSMAA1_2_B5 seems to be the only isolate that is able to synthesise pantothenate from valine or aspartate (M00119).

Overall, these differences including intra-species variation are captured in the 16 corynebacteria isolates from a single individual A1 (**Supplementary Fig 10**). Although we have observed differences between isolates of the same species, we identified mainly species-specific metabolic pathways which could be used to understand the potential to modulate the cutaneous corynebacterial community through metabolic means.

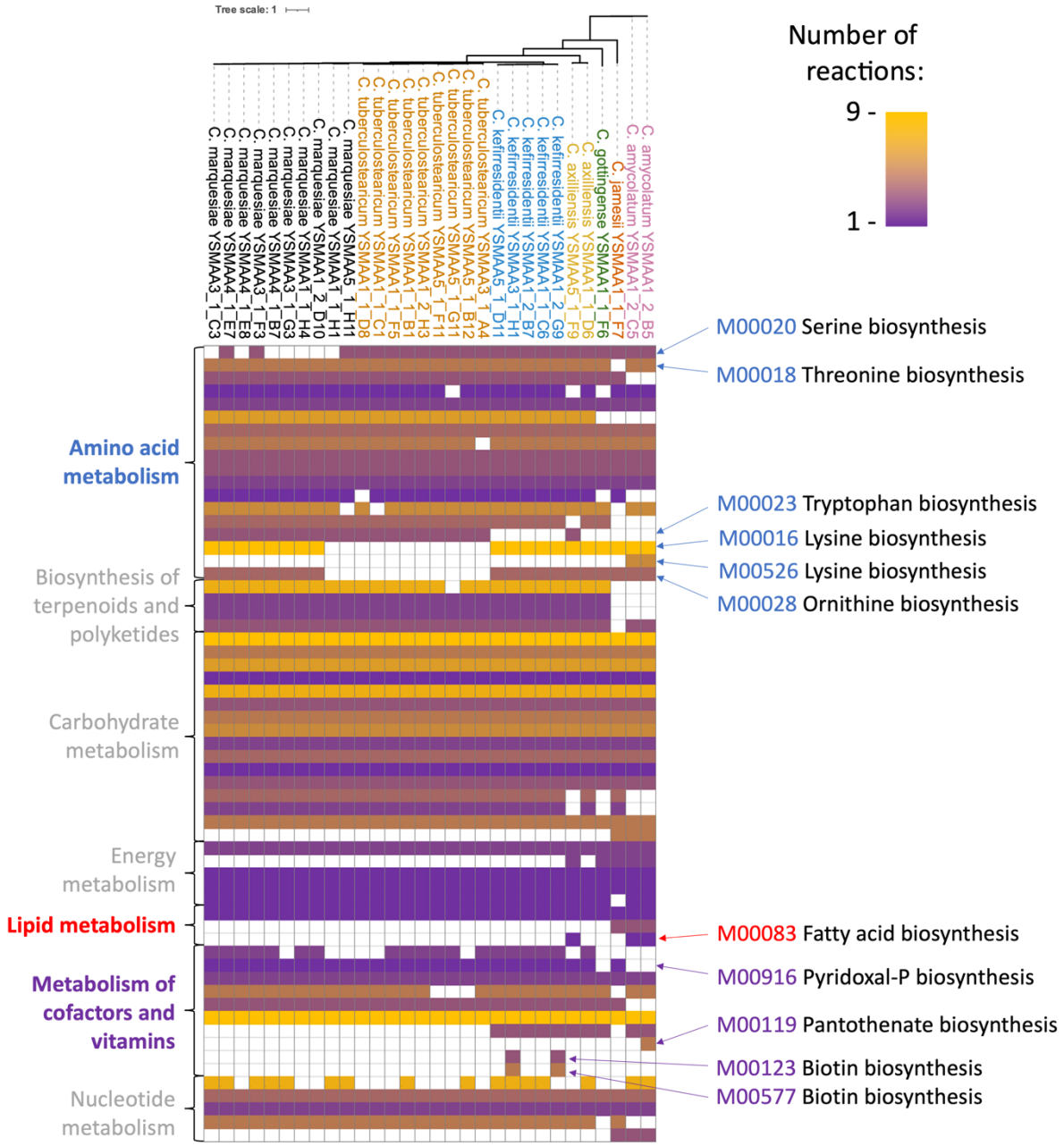


Figure 3. Metabolic pathways in axillary corynebacteria reveals species specific general profiles and some variation within species. Each KEGG module for each isolate were placed on a heatmap where the modules with the largest number of reactions are in yellow and lowest in purple with absent modules in white. Modules of interest were described on the right of the heatmap. KEGG modules were grouped according to the categories on the left of the heatmap. Isolates were grouped by species with a core genome tree depicting the variation. Species names were coloured accordingly (*C. tuberculostearicum* = orange, *C. kefirresidentii* = blue, *C. axilliensis* = gold, *C. gottingense* = green, *C. jamesii* = bronze, *C. marquesiae* = black, *C. amycolatum* = pink).

4.6.3 Diverse biosynthetic capacities and antimicrobial resistance in corynebacterial skin isolates

Bacteria often produce complex small molecules that have various functions relating to microbial competition such as antibiotics and siderophores, that are encoded within recognisable biosynthetic gene clusters (BGCs). Resident bacteria of highly complex microbial communities like the skin are known to produce antibiotics to suppress other species in their niche, including preventing colonisation by pathogenic species (36) and conversely, allow for the proliferation of others (37). We used antiSMASH to identify putative BGCs within the representative set of isolates (38). For the purposes of this global corynebacteria BGC analysis, we group similar predicted clusters together and determined the prevalence of each discrete cluster (**Fig 4**). The majority of the isolates share a predicted non-alpha poly-amino acid (NAPAA), a type 1 polyketide synthase (T1PKS) and a terpene biosynthetic cluster. Other predicted BGCs include non-ribosomal polyketide synthase (NRPS) and multiple terpene clusters with different architectures (**Supplementary Fig. 11 and 12**). The *C. tuberculostearicum* species complex isolates were generally found to possess similar BGCs with *C. tuberculostearicum* isolates having the largest variation (**Supplementary Fig. 13**). As expected, distantly related species like *C. amycolatum, C. gottingense* and *C. jamesii* were predicted to have BGCs that are more distinct or of different classes (**Supplementary Fig. 14**). Despite being outside the *C. tuberculostearicum* species complex, the analysed *C. axilliensis* isolates seemed to also share similar predicted BGCs. With the exception of the cationic peptide ε-Poly-L-lysine (NAPAA), we were not able to confidently predict the products of these BGCs as the most similar known clusters for all of the predicted BGCs shared low percentage similarities with their respective putative clusters. Nevertheless, our analysis suggested different BGCs harboured by the corynebacterial isolates which could contribute to the modulation of the skin microbiome through the production of metabolites like siderophores (e.g. NI siderophore, aminopolycarboxylic-acid, NRPS) and even antimicrobials (e.g. T1PKS, PKS-like + aminoglycoside).

The likely presence of antimicrobials on the human skin may promote the horizontal acquisition of antimicrobial genes or develop beneficial mutants of existing cellular

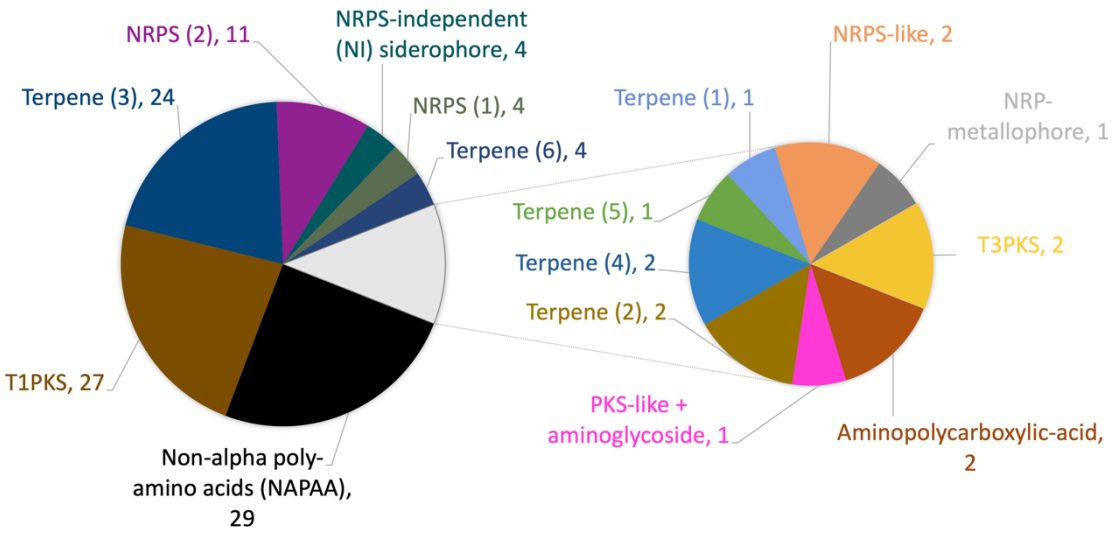


Figure 4. Multiple putative biosynthetic gene cluster (BGC) types were identified in the representative isolate genomes using antiSMASH 7. Names of the different BGC types were included with different cluster architectures in numbered parentheses (i.e. [Name] (X), where X corresponds to a cluster architecture type). The prevalence of each BGC type were also included after the names of each BGC (i.e. …, X).

proteins. ResFinder analysis of the isolates revealed 15 isolates with varying numbers and types of acquired resistance genes but no predicted chromosomal mutations leading to resistance (**Supplementary Table 6**). The most common resistance gene identified was *erm(X)* which is proposed to confer resistance to erythromycin-related antibiotics and is widely found across the genus (39–43). *C. tuberculostearicum* YSMAA5_1_F11 and YSMAA5_1_G11 were also found to have an aminoglycoside resistance gene which could confer resistance to gentamycin and tobramycin. *C. marquesiae* YSMAA1_1_H4, *C. tuberculostearicum* YSMAA3_1_A4 and *C. axilliensis* YSMAA5_1_F9 seem to have a resistance gene against an unknown aminoglycoside initially found in *C. striatum*. Interestingly, *C. axilliensis* YSMAA5_1_F9 has genes conferring resistance to chloramphenicol (cmx), streptomycin (aph(6)-Id, aph(3")-Ib) and kanamycin (aph(3")-Ia). These were found in tandem in the same region of the genome surrounded by multiple insertion sequence (IS) elements and were also not found in *C. axilliensis* YSMAA1_1_D6, suggesting a recent acquisition. Finally, we identified a sole tetracycline resistance gene in *C. tuberculostearicum* YSMAA5_1_B12 which also resides near IS elements, suggesting another example of horizontally acquired resistance mechanisms. We attempted to experimentally verify what we identified using ResFinder using the EUCAST disc diffusion assay (**Supplementary Table 7**). Despite the common identification of erm(X) which should confer resistance to clindamycin, only 10 out of the 14 isolates predicted to be resistant were as such. *C. axilliensis* YSMAA5_1_F9 was experimentally resistant to chloramphenicol, as predicted by ResFinder. However, *C. tuberculostearicum* YSMAA5_1_B12 was thought to be doxycycline resistant but did not display the phenotype in our disc diffusion assay. The presence of certain resistance genes may not confer resistance in lab-based disc diffusion assays as we may not have captured the most appropriate conditions in which these isolates would be resistant. Interestingly, *C. jamesii* YSMAA1_1_F7 seem to be experimentally resistant to Penicillin G even though a β-lactamase was not predicted to be in the genome suggesting this isolate has some other innate mechanism of resistance.

4.6.4 Diverse prophage repertoires and conserved phage defence systems in skin-associated *Corynebacterium* isolates

Another type of commonly acquired mobile genetic element are prophages. Phages are commonly found in the microbiome and could play an important role in skin health by regulating the bacterial community. We identified 17 putative prophage sequences in our 30 representative genomes, with an average of 0.6 (± 1.4) prophages per genome [50] (**Table 1**). Putative prophages were identified using geNomad (44) followed by gene annotation using pharokka and phold(45). Only putative prophages containing genes which were functionally annotated with multiple common features of a phage (e.g. tail proteins, head and packaging) were retained for further analysis. These putative prophages were then searched against the NCBI database using BLASTn to identify the closest related phage and bacteria species. Some putative prophages were either matched with phages but retaining poor coverage (A1_1_F7_phi1, A5_1_F9_phi1) or not matched with any (A1_2_B5_phi1). As expected, various regions of these putative prophages also matched with other corynebacteria and similar to prophage hits, we observe some bacteria hits with poor coverage (A1_1_F7_phi1, A3_1_H1_phi1). As some of these isolates were derived from the same skin site of the same individual, we sought to understand the diversity of these putative prophages as some prophages may be shared between isolates. We performed a multisequence alignment using MAFFT (46) to determine the genetic identities between these putative prophages (**Supplementary Figure 15**). Neither of these putative prophages seem to be shared between isolates, even within the isolates from the same individuals. Remarkably, none of the 9 representative *C. tuberculostearicum* genomes were found to possess any putative prophages.

In addition to the presence of prophages, we analysed the representative genomes with PADLOC (47) to identify potential phage defence systems (**Fig 5**). Type I, II, IIG and IV restriction modification (RM) systems are commonly found in all isolate genomes while the Type III RM system is much less abundant. Other commonly identified systems include the plasmid transformation protection system Wadjet Type I (48), the abortive phage infection system AbiD (49). We also observed the CRISPR-Cas Type I-E system in some isolates of *C. amycolatum*, *C. tuberculostearicum* and

Table 1. The closest phages and bacteria genomes were identified using a BLASTn analysis on the putative prophages identified by geNomad. The percentage sequence identity and percentage coverage of each hit were included. The putative prophage names are coloured consistently with the species of origin (*C. kefirresidentii* = blue, *C. axilliensis* = gold, *C. jamesii* = bronze, *C. marquesiae* = black, *C. amycolatum* = pink)

Putative prophage name	Putative prophage length	Source species	Closest phage hit (% identity, % query cover)	Closest bacteria hit (% identity, % query cover)
A1_1_C6_phi1_2	52250	*C. kefirresidentii*	ctNWy2 (92.03%, 69%)	*C. tuberculostearicum* CT (91.51%, 76%)
A1_1_D6_phi1	42557	*C. axilliensis*	ctIBL2 (83.98%, 29%)	*C. aurimucosum* DSM 44532 (87.61%, 26%)
A1_1_F7_phi1	56384	*C. jamesii*	ctwl93 (77.01%, <1%)	*C. jeikeium* DSM 7171 (91.59%, 7%)
A1_1_H1_phi1	49192	*C. marquesiae*	ctfR81 (90.01%, 67%)	*C. aurimucosum* FDAARGOS_1110 (84.54%, 49%)
A1_2_B5_phi1	38625	*C. amycolatum*	-	*C. jeikeium* FDAARGOS_574 (92.93%, 37%)
A1_2_D10_phi1	42401	*C. marquesiae*	ctKpD5 (81.46%, 43%)	*C. accolens* CA1 (80.24%, 49%)
A1_2_G9_phi1	26374	*C. kefirresidentii*	ctOSk1 (95.99%, 64%)	*C. tuberculostearicum* CT (91.72%, 65%)
A3_1_C3_phi1	38823	*C. marquesiae*	ctSSL1 (93.44%, 45%)	*C. imitans* NCTC13015 (81.58%, 28%)
A3_1_F3_phi1	54302	*C. marquesiae*	ctcbl1 (92.73%, 70%)	*C. accolens* CA9 (75.43%, 38%)
A3_1_G3_phi1	47617	*C. marquesiae*	ct1BD1 (92.15%, 54%)	*C. tuberculostearicum* CTNIH13 (89.64%, 42%)
A3_1_H1_phi1	47347	*C. kefirresidentii*	ctLCM1 (95.25%, 44%)	*C. tuberculostearicum* CTNIH23 (97.62%, 9%)
A4_1_E7_phi1	47738	*C. marquesiae*	ct1BD1 (86.18%, 54%)	*C. tuberculostearicum* CTNIH13 (89.94%, 40%)
A5_1_D11_phi5_2_1	49670	*C. kefirresidentii*	ctNWy2 (95.17%, 85%)	*C. tuberculostearicum* CT (99.84%, 100%)
A5_1_D11_phi3_4	50578	*C. kefirresidentii*	ctNWy2 (96.26%, 64%)	*C. tuberculostearicum* CT (96.35%, 81%)
A5_1_F9_phi1	43162	*C. axilliensis*	ctIBL2 (84.39%, 19%)	*C. aurimucosum* DSM 44532 (88.11%, 56%)
A5_1_H11_phi1	50212	*C. marquesiae*	ctXFF1 (97.39%, 85%)	*C. accolens* CA9 (84.87%, 51%)
A5_1_H11_phi2	49906	*C. marquesiae*	ct1BD1 (92.69%, 46%)	*C. tuberculostearicum* CTNIH13 (91.20%, 75%)

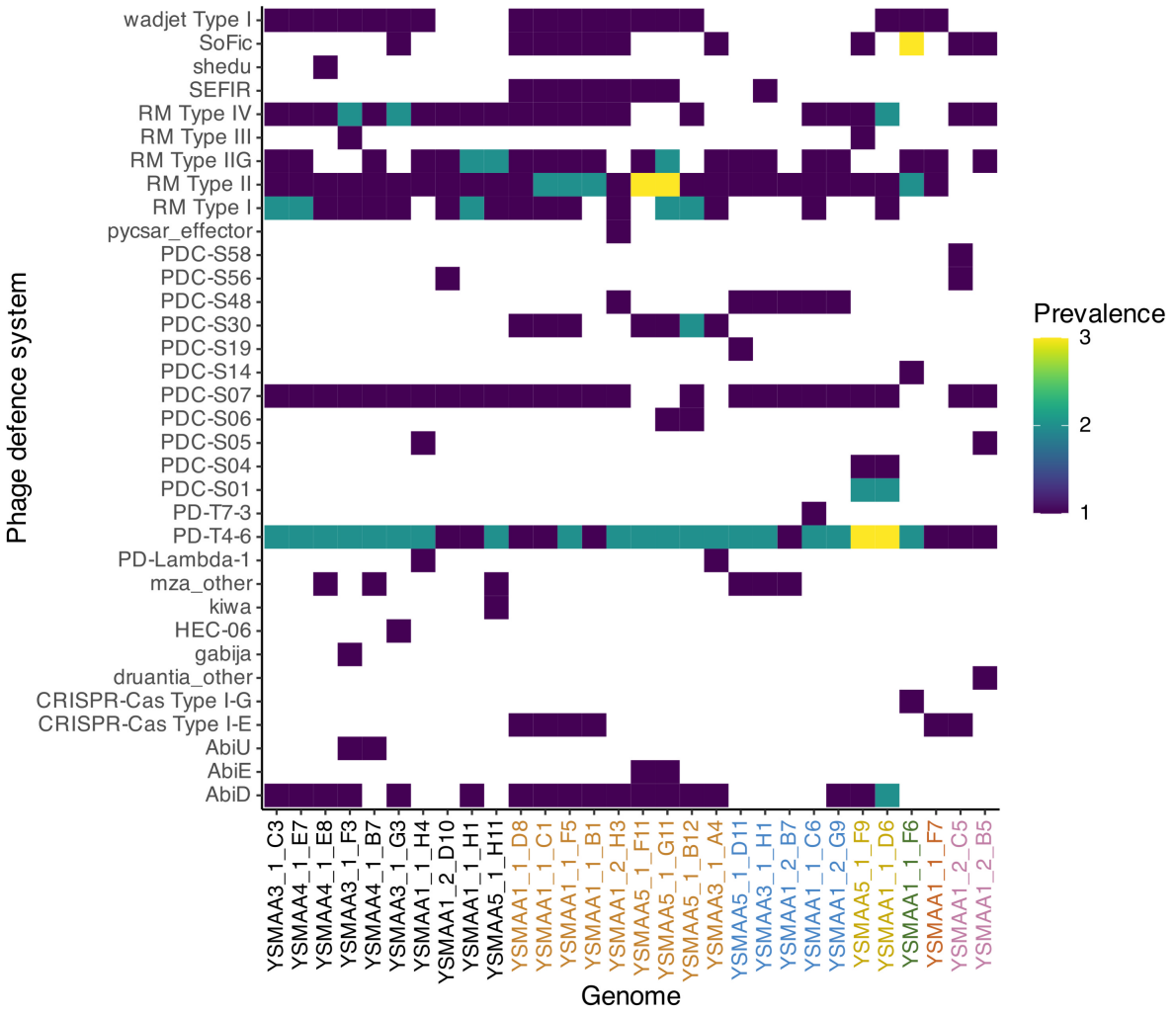


Figure 5. Various phage defence systems were predicted in the representative genomes using PADLOC. The heatmap depicts the prevalence of each phage defence system within each isolate according to the scale on the right with absent systems in white. The genomes were arranged according to the different corynebacteria species and coloured accordingly (*C. tuberculostearicum* = orange, *C. kefirresidentii* = blue, *C. axilliensis* = gold, *C. gottingense* = green, *C. jamesii* = bronze, *C. marquesiae* = black, *C. amycolatum* = pink).

C. jamesii and also the CRISPR-Cas Type I-G system only in *C. gottingense.* We also predicted the presence of multiple low prevalence phage defence systems like the putative nucleotide-sensing Gabija defence (50), the pyrimidine cyclase system for anti-phage resistance (Pycsar) (51) and other defence systems of unknown mechanisms like shedu and kiwa (48). Global PADLOC analysis of all 215 complete genomes of our axillary corynebacteria isolates suggested typical conservation of these predicted phage defence systems within the drep secondary clusters with some exceptions (**Supplementary Fig. 16 - 19**). Some of these differences include select isolates from *C. marquesiae* cluster 4_5 and 4_8 and *C. tuberculostearicum* cluster 3_2 which lack the RM type IV system. A minority of *C. marquesiae* cluster 4_5 isolates were also predicted to have the putative phage RNA interacting Mokosh Type II system (52). Similarly, only two isolates of *C. kefirresidentii* cluster 2_5 lack the RM Type IIG system and two other isolates of the same cluster were predicted to have the AbiD system. Together, we have comprehensively covered the range of phage defence systems predicted to be present in our axillary isolates and we expect these mechanisms to be present in other skin associated corynebacteria.

4.7 Discussion

The growing number of recent studies exploring the diverse roles of skin commensals, whether antagonistic or mutualistic, underscores the need to accurately characterise the true constituents of the skin microbiome. Within the population of skin commensal microbes, staphylococci and cutibacteria have been widely studied (2–4, 53). In contrast, although corynebacteria make up a large proportion of the microbiomes across multiple skin sites, commensal species from this genus remain poorly studied. This highlights the importance of studies like this, not only to uncover the roles of these commensals but also to discover novel skin-associated species.

Others have also sought to address the lack of information on this genus by attempting to clarify the diversity of cutaneous corynebacteria (21, 22, 26) but the genetic information for this genus still lags behind that of cutaneous staphylococci and cutibacteria. The recent advances in long-read sequencing, particularly the R10.4.1 flow cell and Kit 14 chemistry by Oxford Nanopore Technologies, have allowed for

high quality sequencing to be more accessible at lower costs to researchers. With this we developed and optimised a direct-to sequencing method enabling us to contribute another 30-representative cutaneous corynebacteria genomes with another 195 closed genomes which were filtered out during dereplication but are available for analysis.

Unlike our study, previous efforts to understand the skin microbiome usually involved an initial 16S rRNA amplicon sequencing step which serves as a filtering step to select specific isolates for whole-genome sequencing (21, 22). The high throughput nature of 16S rRNA gene amplicon sequencing makes it an appealing first step in determining the diversity of isolates. A common way to compare the 16S rRNA gene sequences is by aligning their respective V1-V3 fragment as it has been shown to be able to distinguish between species more reliably (54, 55). However, this approach does not resolve the differences between isolates of closely related but ultimately different species. For example, the skin commensals *C. kefirresidentii* strain FDAARGOS_1055 and *C. tuberculostearicum* strain FDAARGOS_1198 share a 16S rRNA (V1-V3 fragment) sequence identity of 99.8% but only share a ~89% genomic sequence identity when calculated using FastANI. Clearly, the difference between the two species is only obvious when whole genome sequences are available. The difficulties in resolving the diversity of isolates were also previously reported in multiple other studies as microbial “dark matter”, or isolates that cannot easily speciated using the current tools and information available (2, 56, 57). Further, common mobile elements like prophages which are known contributors to genetic variation would be missed without a WGS approach. We were able to keep our WGS cost low negating the need for an initial 16S rRNA gene filtering approach and allowing us to reveal the multitude of invaluable genetic information from WGS.

Using our method, we were able to confirm the widespread presence of the species *C. tuberculostearicum*, *C. kefirresidentii* and *C. marquesiae* (formerly *C. aurimucosum*_E) (**Supplementary Fig. 20**), which together form the *C. tuberculostearicum* species complex (27). In addition, we found more rarely *C. amycolatum* strains in the axilla and unexpectedly, we also identified isolates of *C. gottingense*, a species previously not associated to the skin. This species has been isolated from blood of patients with bacteraemia, the cerebrospinal fluid (CSF) and

urine (30, 58), suggesting that it possibly originates from the skin. In addition to known species, we also identified two novel species tentatively named *C. jamesii* and *C. axilliensis*. Isolates of *C. axilliensis* were found on individuals A1 and A5 and makes up a significant 23% of sequenced corynebacterial isolates from the latter suggesting this novel species could be a major previously unrecognised components of the axillary microbiome of some individuals.

Although the fundamental roles of cutaneous commensal corynebacteria are still poorly understood, there have been increasing studies demonstrating how bacteria from this genus can contribute to colonisation protection including against other human pathogens. For example, *C. striatum* has been demonstrated to shift *S. aureus* towards commensalism through the suppression of the *agr* quorum sensing system (14). Additionally, Rosenstein *et al*. (37) have proposed cooperation between corynebacteria and *Staphylococcus lugdunensis* against *S. aureus* in the nasal cavity. These studies have highlighted the importance in understanding the true diversity of cutaneous corynebacteria species as we may yet uncover more roles that could be exploited to improve tools and treatments for skin health.

Our isolate characterisation approach provided us with the ability to predict differences in metabolic pathways, the presence of biosynthetic gene clusters, mobile genetic elements and bacterial defence systems. For example, there were general correlations in the presence of complete amino acid pathways between species (e,g, serine biosynthesis) but we also observed intra-species differences, including within the two strains of *C. axilliensis* we identified (**Fig 3**). To demonstrate the power of our culturomics approach, we performed an unparalleled in-depth characterisation of over 150 axillary isolates from the same volunteer which suggested the possibility of metabolic sharing (e.g lysine) between predicted auxotrophs and other members of the microbiome (**Supplementary Fig. 10**). Further, the discovery of novel biosynthetic gene clusters, including antimicrobials and secondary metabolites, could pave the way towards understanding how certain members of the genus *Corynebacterium* could modulate the microbiome to maintain skin health. Together, these findings could form the basis of further studies to explore networks between members of the same human skin site.

We acknowledge that the findings of this study are limited by the number of volunteers tested and the heavy characterisation of only a single volunteer. Further, only one selective media was used at the point of isolation which could exclude certain isolates of this genus. Nevertheless, our study not only significantly increased the number of *Corynebacterium* genomes available to help reveal their roles in the skin microbiome but also suggest a workflow that can be adapted further in larger-scale studies of this genus.

Studies advancing our understanding of the skin microbiome, including those focusing on cutaneous corynebacteria appear to only be scratching the surface of the skin's microbial diversity. Our knowledge of commensal corynebacteria lags behind other equally prevalent genera like *Staphylococcus* and *Cutibacterium*. In this study, we set out to explore the genetic diversity within the genus *Corynebacterium* using a cost-effective, direct-to-WGS approach on isolates collected from a single skin site, the axilla. Through this strategy, we uncovered species previously unreported on human skin, including two novel species, and revealed extensive genomic variation within our collection of volunteer matched axillary corynebacteria. These included putative metabolic distinctions, diverse biosynthetic gene clusters, and integrated prophages with each offering promising avenues for further investigation. The genomes are publicly available through the MORF platform (https://morf-db.org/projects/Public/MORF000083) and represent a rich resource for the wider research community to deepen our understanding of this historically underexplored genus. Together with the broader York Skin Microbiome (YSM) isolate collection, these resources lay essential groundwork for future studies, advancing our efforts to unravel the complexities of the skin microbiome.

4.8 Materials and methods

4.8.1 Sampling procedures

Four adult male volunteers (Age range: 32 to 78) were recruited for axillary swabbing at Unilever R&D, Bebington, UK. The volunteers were in good general health, did not use any body products (e.g. deodorants) for 24 hours before the study, were not taking antibiotics, had no active skin conditions on any part of their bodies, have not suffered

from eczema in the last 5 years nor have they ever had psoriasis. Prior to swabbing, the volunteers were “sniff tested” by a panel of testers and scored according to odour levels (High malodour: A1 and A3, Low malodour: A4 and A5). Axillary swabs were then collected by swabbing in a linear motion 20 times using eSwabs (Copan) on both underarms and then placed into Amies transport medium.

4.8.2 Bacterial culturing

Swabs were plated onto ACP solid media (39.5 g L^{-1} Blood Agar base no. 2, 3 g L^{-1} yeast extract, 2 g L^{-1} glucose, 5 mL L^{-1} Tween 80, 50 mL L^{-1} defibrinated horse blood, 100 mg L^{-1} fosfomycin) to enrich for corynebacteria and grown for 2 days at 37°C. Alongside the 215 *Corynebacterium* isolates, we also isolated 42 *Staphylococcus* isolates from the species *S. capitis* and *S. warneri*. Colonies were picked and grown overnight in brain-heart infusion broth + 1 % Tween-80 (BHIT) at 37°C, shaking at 200 rpm. Cultures were stocked in BHIT supplemented with 10 % glycerol in 2 mL 96 well plates.

4.8.3 DNA extraction

For genomic and plasmid DNA extraction, 0.5 mL of BHIT was inoculated with each isolate in 2 mL 96 well plates and grown for 20 hours at 37°C, shaking at 200 rpm. Cultures were then pelleted by centrifugation before resuspension in 100 μL enzymatic lysis buffer (50 mM Tris-HCl, pH 8, 50 mM EDTA, pH 8, 0.5% Tween 20, 0.5% Triton-X100, 25 mg/mL lysozyme). The mixtures were incubated at 37°C overnight with gentle shaking (<50 rpm). 500 μL of lysis buffer (0.5 M NaCl, 100 mM Tris-HCl, pH 8, 50mM EDTA, pH 8, 1.5% SDS) was then mixed into each well and incubated at 80°C for 30 minutes in a thermomixer. The mixture was allowed to cool before the addition of 4μl of 4mg/mL RNase A Solution (Promega) before incubation at 37°C for a further hour. Cellular proteins were then degraded with 20 μl of 20 mg/mL Proteinase K Solution (Promega) and incubation at 56°C for 15 minutes. Remaining proteins were precipitated with 200 μl of protein precipitation buffer (5 M Potassium acetate). The lysates were then clarified by centrifugation at 16,000 xg for 5 minutes. DNA was precipitated from the supernatant using 600 μl isopropanol and incubated overnight at

4°C. The precipitated DNA were pelleted by centrifugation at 16,000xg for 5 minutes. The DNA pellets were washed with 600 µl 70% ethanol and then air dried for 15 minutes. Finally, the DNA pellets were rehydrated with 50 µl nuclease-free water and incubation overnight at room-temperature. The DNA concentrations and QC measurements were performed using Qubit (Thermo Fisher) and Nanodrop (Thermo Fisher).

4.8.4 DNA library preparation and whole genome sequencing

The Oxford Nanopore Native Barcoding Kit 96 V14 was used to prepare and barcode the isolated genomic and plasmid DNA. The recommended protocol was followed with the following changes: 1. Half the recommended volume of the DNA repair and end-prep reaction was prepared. 2. Half the recommended volume of the native barcode ligation reaction was prepared. The prepared library was loaded into a PromethION R10.4.1 flow cell and sequencing was allowed to take place over 72 hours and base-called using the singleplex high-accuracy model, 400 bps on MinKNOW v23.07.5. The reads were assembled with the EPI2ME Labs wf-bacterial-genomes isolate workflow (Flye v2.92-b1786 (59), Medaka v1.9.1 (Oxford Nanopore Technologies), Prokka v1.14.5 (60), ResFinder v4.3.2 (61, 62)). Genome sequences are available through the York Skin Microbiome project in MORF (https://morf-db.org/projects/Public/MORF000083).

4.8.5 Bioinformatics tools

Pairwise genome comparisons of all 215 assembled genomes were performed using the ANIm (63) analysis on pyANI v0.2.12 (64). The genomes were dereplicated with drep v3.4.2 (23) using Mash v2.3 (65) and MUMmer v3.0 (66), employing a primary clustering ANI cutoff of 95% to distinguish between different species and a secondary clustering ANI cutoff of 99.5% to capture strain diversity to derive a representative set of genomes. The core genomes were determined using Roary v3.13.0 (32) with an 80% BLASTp cutoff. PhyML v3.3 (67) trees 16S rRNA V1-V3 and core genome alignments of the representative genomes were constructed using the HKY85 substitution model with approximate likelihood-ratio tests (aLRT) (68). Trees were

visualised on iTOL v6.9 (69). Speciation using GTDB-tk v2.3.2 (24) was performed on KBase (25). The genome alignments of *C. axilliensis* and *C. jamesii* were prepared using BLASTn and visualised using BRIG v0.95 (70). *C. axilliensis* was named after the axilla, where this species was first discovered. *C. jamesii* was named after Alexander Gordon James who is a pioneer in axillary corynebacteria research.

Pangenome analysis was performed using the anvi-pan-genome package and visualised on anvi'o v8 (33). The representative genomes were also compared using pyANI and visualised with the pangenome on anvi'o. Relevant gene clusters were extracted and functionally analysed with eggNOG-mapper v2 (71, 72) to determine the respective COG categories. Metabolic pathways were identified using BlastKOALA v3.0 and KEGG Reconstruct (73, 74). Biosynthetic gene cluster analysis was performed using the default settings on antiSMASH v7 (38). Putative prophages were predicted using geNomad v1.8 (44) and then annotated using pharokka v1.7.2 and phold v0.1.4 (45). Phage defence systems were predicted using PADLOC v4.3 (75).

4.8.6 Antibiotic susceptibility testing

Antibiotic susceptibility testing was performed using plate-based disc-diffusion assays using the recommended manufacturer's protocol and EUCAST instructions (www.eucast.org). Overnight cultures were prepared BHIT and grown at 37°C. Cultures were then diluted to a 0.5 McFarland standard (Remel, Thermo Scientific) before plating onto BHIT-agar. Six antibiotic susceptibility discs were placed on the agar (Doxycycline-30μg, Chloramphenicol-30μg, Fosfomycin-200μg, Clindamycin-10μg, Vancomycin-30μg, Penicillin G-1μg) (Oxoid). Plates were allowed to grow for 24 hours at 37°C before measuring zones of inhibition.

4.9 Ethics statement

This study was conducted in accordance with the ethical principles of Good Clinical Practice and the Declaration of Helsinki. The local ethics committees of Unilever R&D

(Port Sunlight, UK) approved the protocols before commencement of the studies and all subjects gave written informed consent.

4.10 Data availability

The 24 dereplicated genomes and associated plasmids were deposited to GenBank within BioProject PRJNA1191350.

4.11 Conflict of interest

The authors declare that B. M. and R. C. are current employees of Unilever.

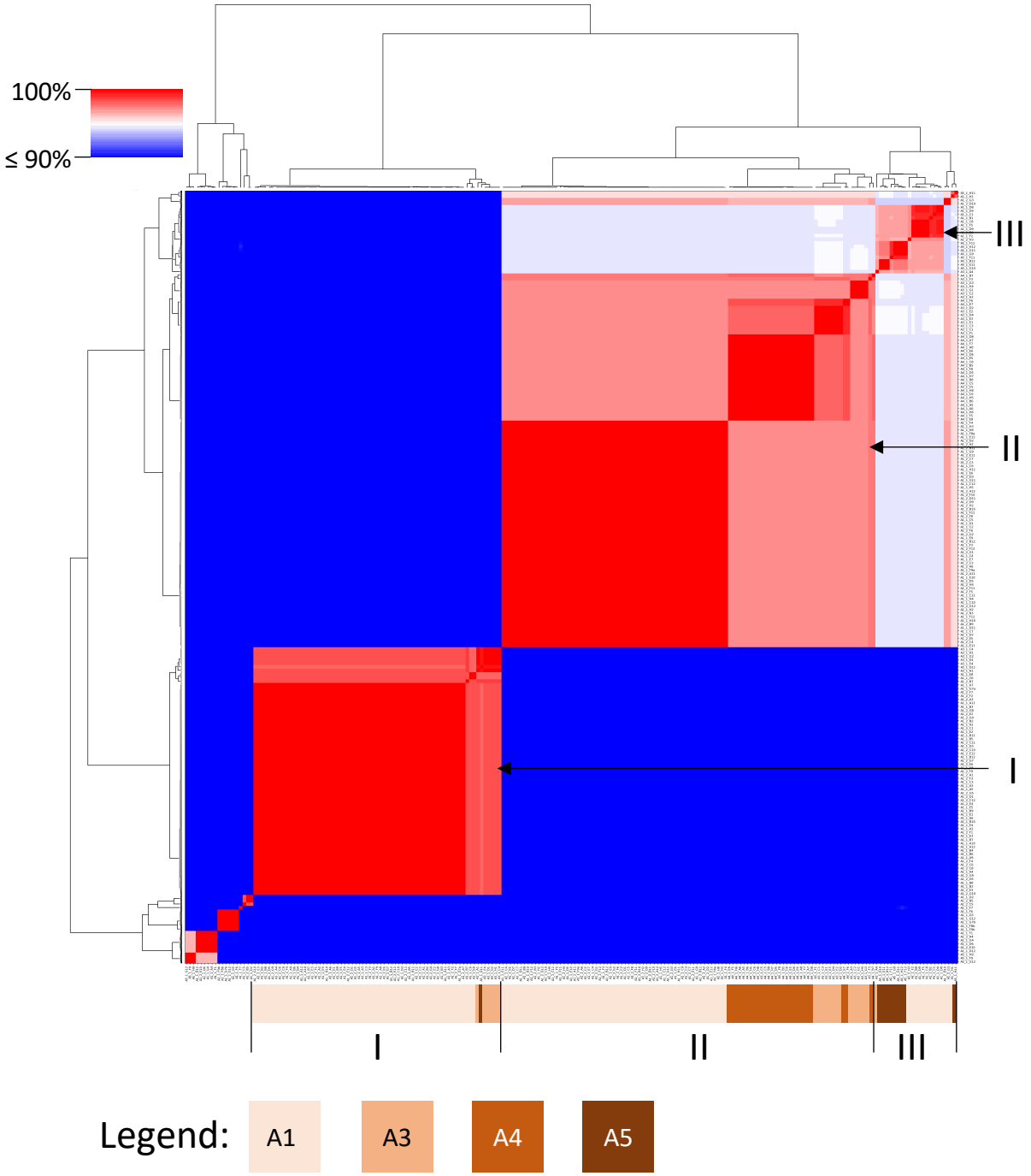
4.12 Funding

This work was supported by the Biotechnology and Biological Sciences Research Council grant BB/W510531/1 supporting R. H.

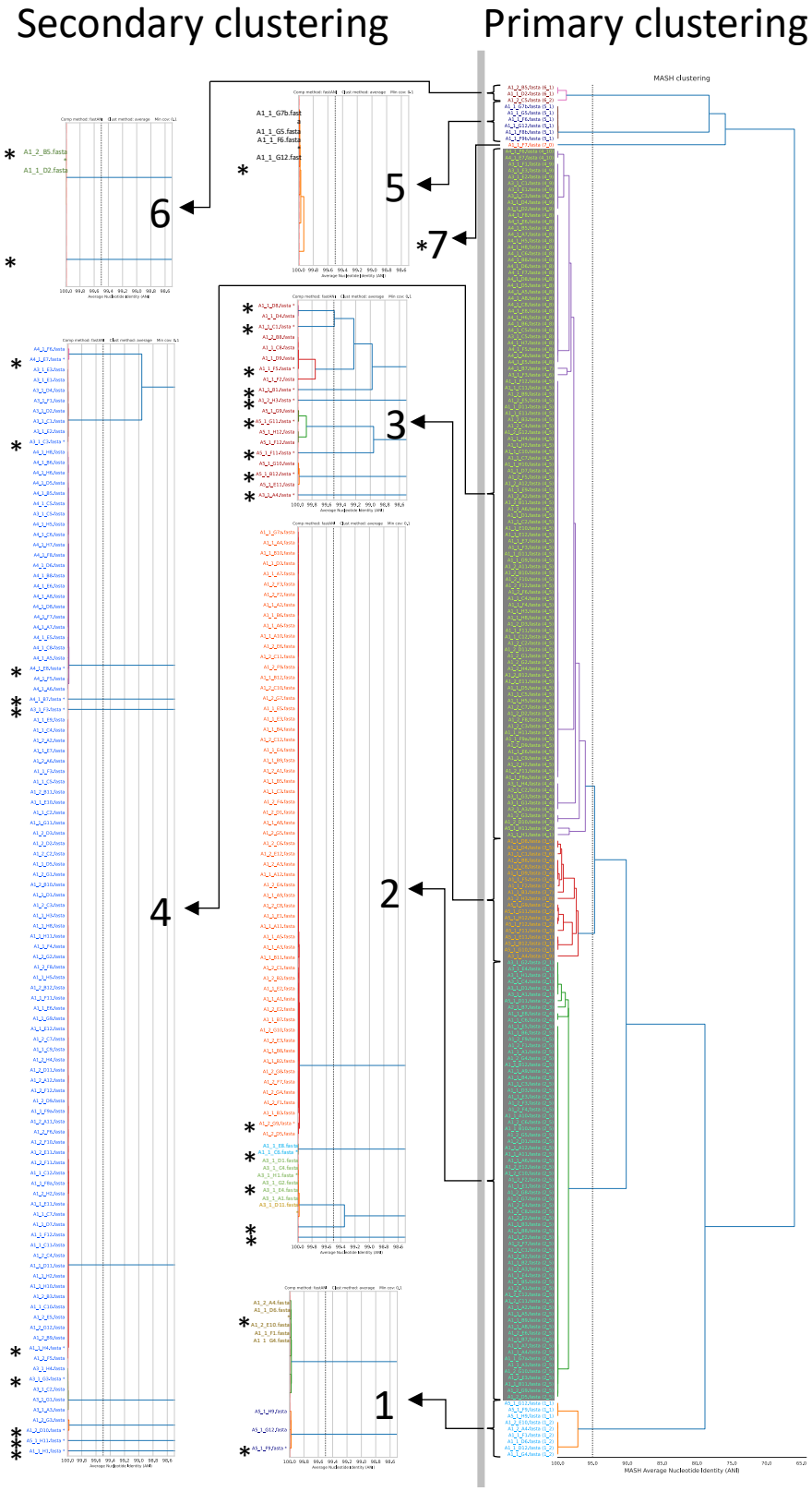
4.13 Author contributions

R. H., R. C, B. M. and G. H. T. conceptualization; R. H., M. R., and B. M. validation; R. H., S. M., and G. H. T. formal analysis; R. H., S. M. and M. R. investigation; R. H. and M. R. data curation; H. W., M. R. and M. J. H. resources; R. H. writing–original draft; R. H., S. M., H. W., M. R., R. C., M. J. H., A. J. W., B. M., and G. H. T. writing–review & editing; R. H., M. R. and S. M. visualization; A. J. W., B. M., and G. H. T. supervision; R. H., B. M., and G. H. T. project administration; B. M. and G. H. T. funding acquisition.

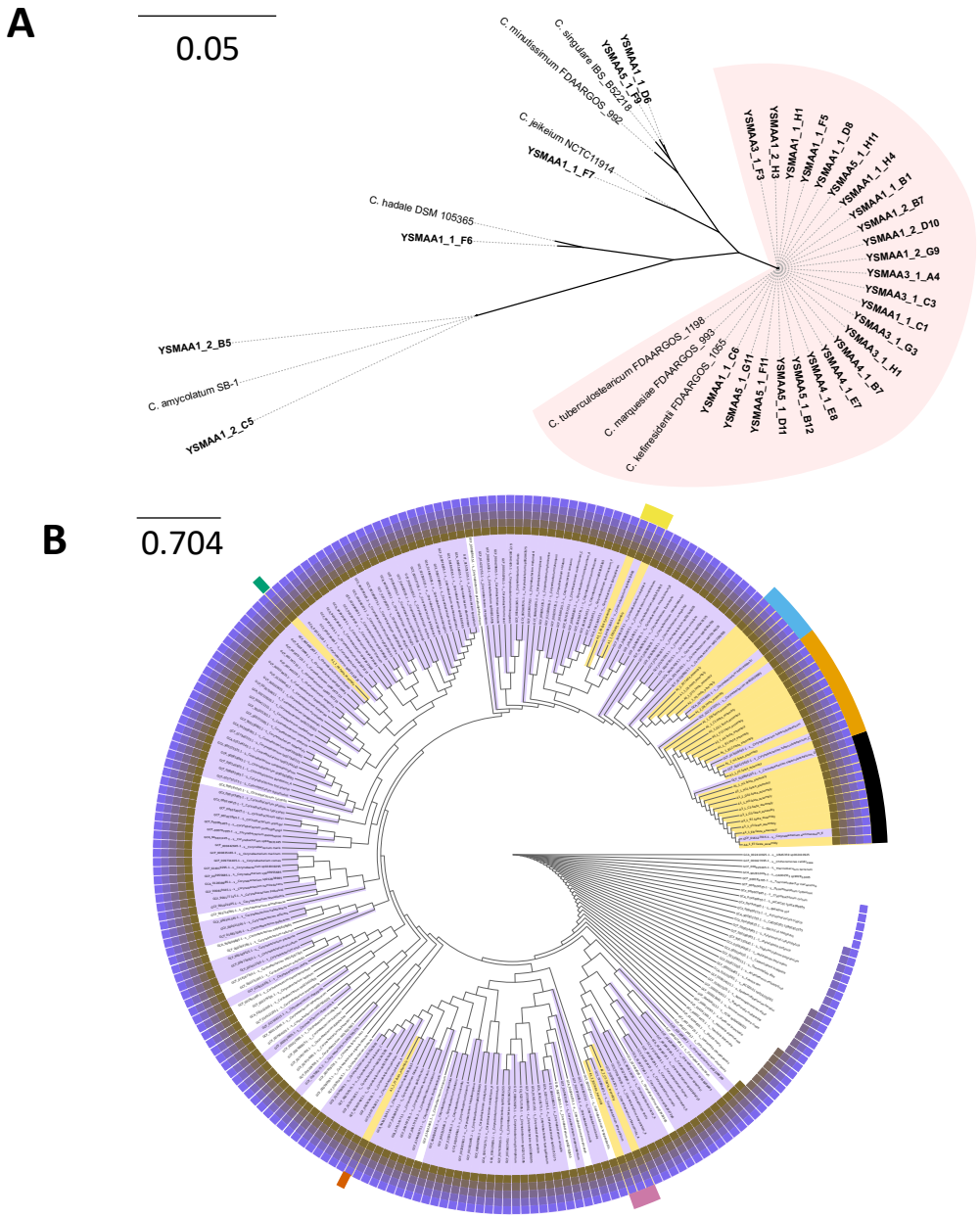
4.14 Supporting data



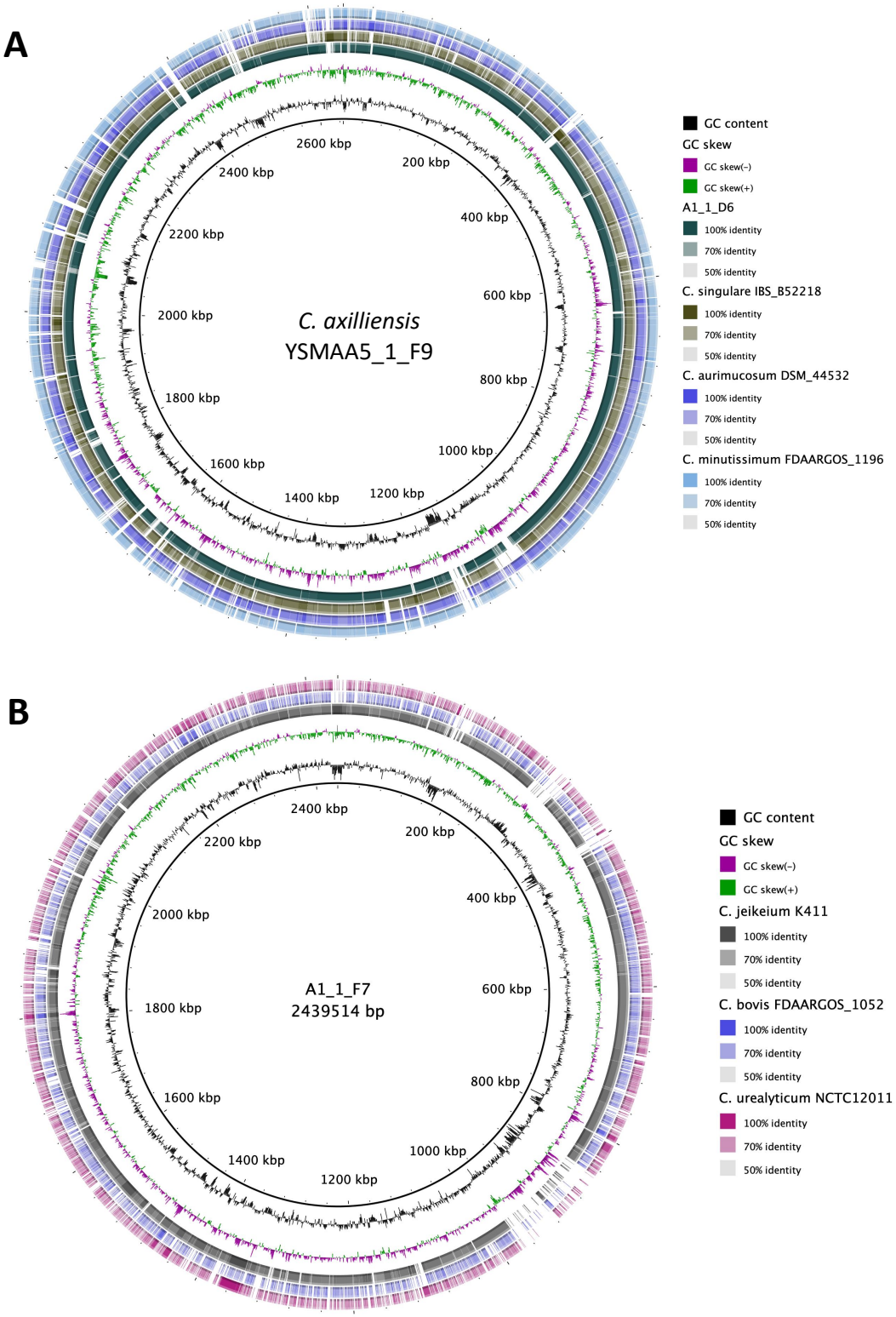
Supplementary Figure 1. Pairwise average nucleotide identities (ANI) of all isolates sequenced in this study. ANI values were placed onto a collective coloured pairwise matrix according to the colour scale depicted Three major groups of isolates were identified with I, II and III. Group I II and III isolates from each of the four individuals were denoted with the coloured bars as indicated in the legend.



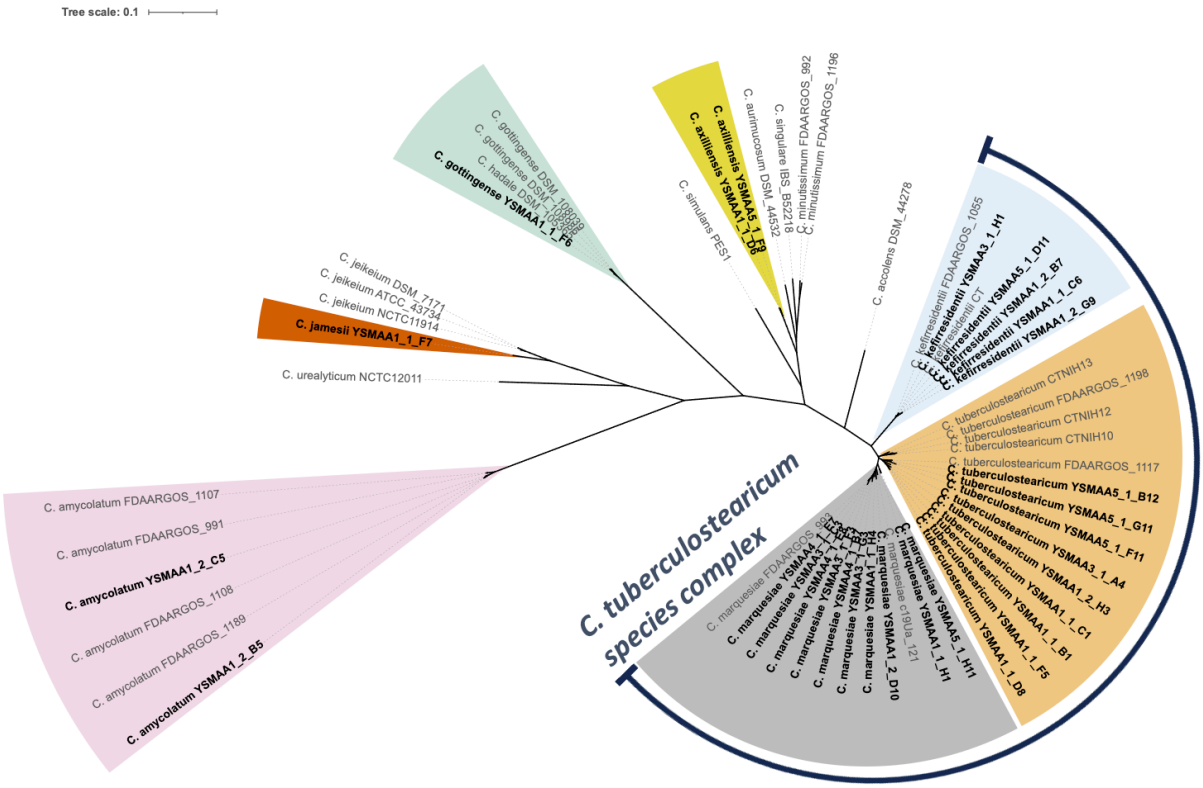
Supplementary Figure 2. Isolates were initially clustered by ANI to 7 primary clusters (species level) using MASH at a 95% threshold. A secondary clustering was performed on the isolates using fastANI at a 99.5% threshold to identify representatives from each cluster (labelled with *).



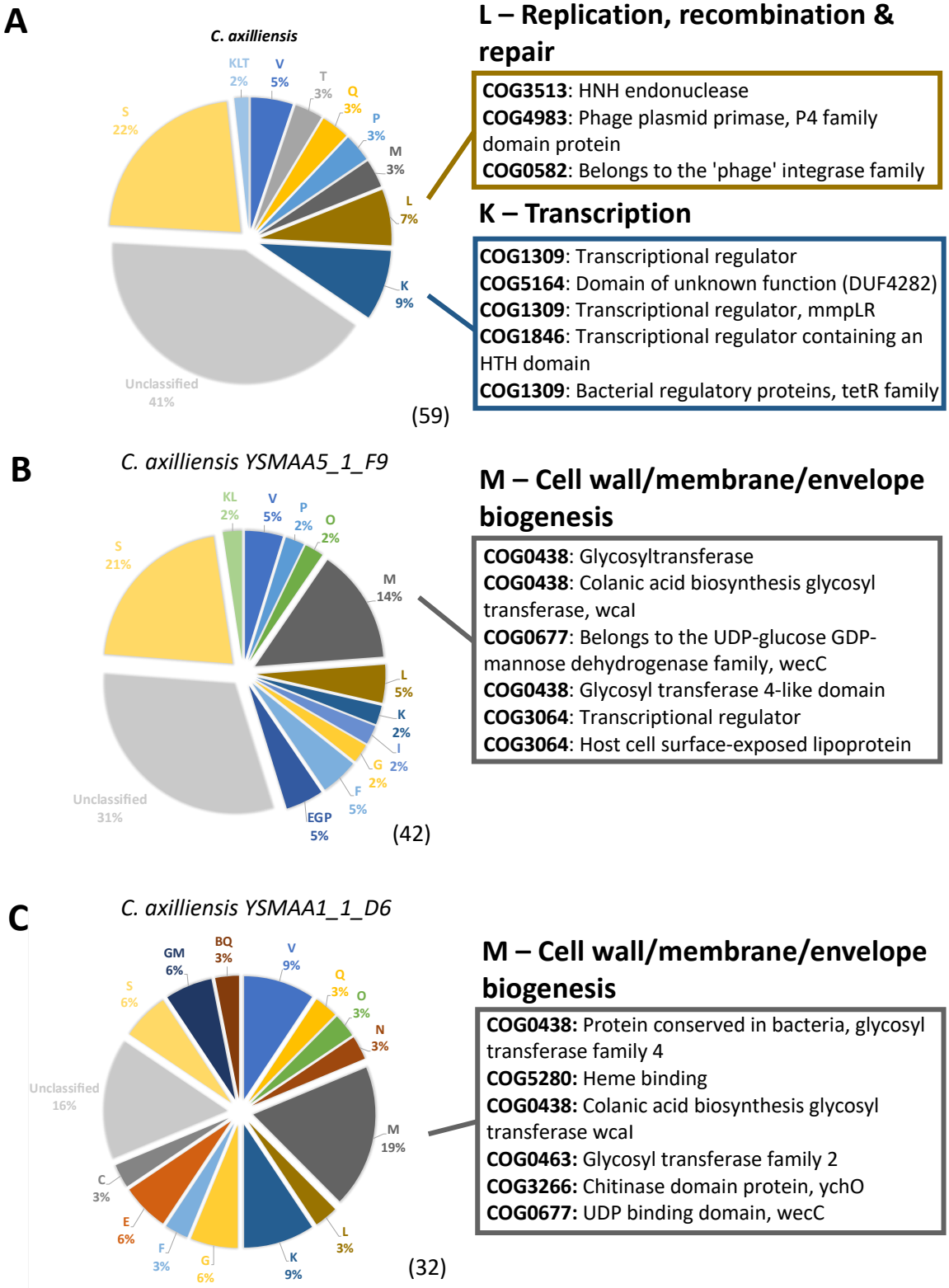
Supplementary Figure 3. Speciating the representative isolates. (A) The V1-V3 regions of the 16S rRNA sequences for each genome were aligned with those of the closest RefSeq genome BLASTn hits and plotted onto an unrooted tree. Branches within the red cloud share ≥ 99.8% identity. The tree was generated using PhyML and visualised on iTOL. (B) Genomes of the representative isolates (yellow) were compared to the genomes of the GTDB. 7 distinct groups of isolates were identified with coloured bars in the outermost ring (*C. kefirresidentii* – blue, *C. tuberculostearicum* – orange and *C. aurimucosum*_E/*C. marquesiae* – black, *C. axilliensis* – yellow, *C. gottingense* – green, *C. jamesii* – dark orange, *C. amycolatum* – magenta).



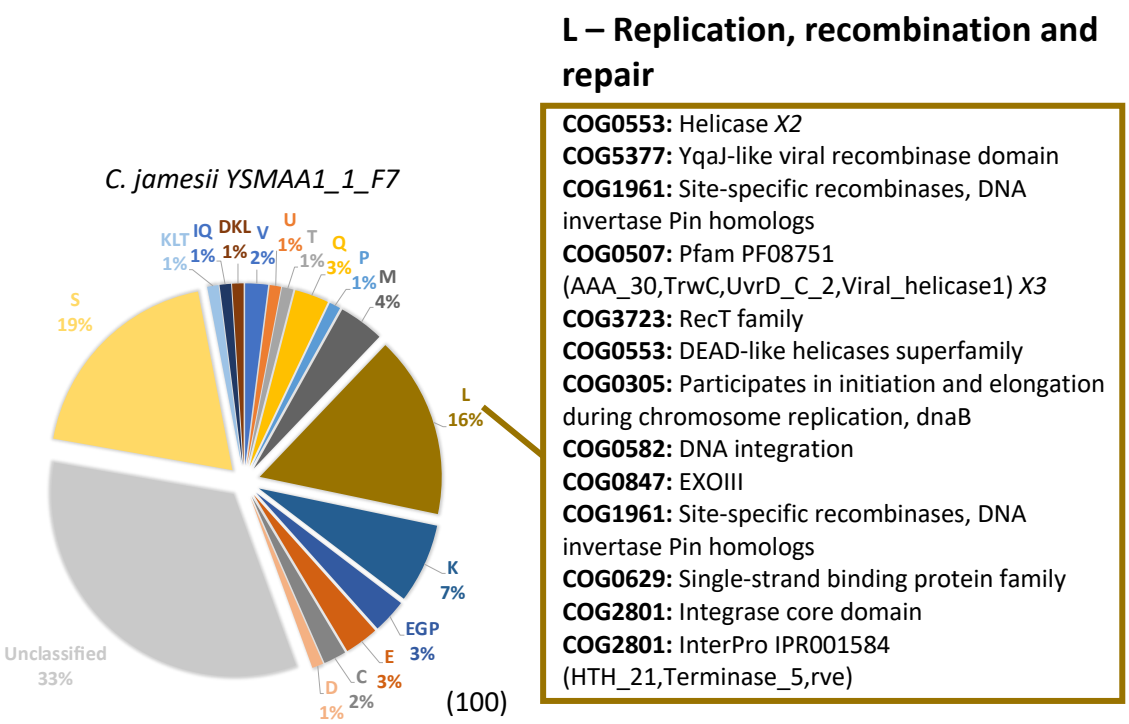
Supplementary Figure 4. Alignments of representatives from the two novel species *C. axilliensis* and *C. jamesii* against closest hits from GenBank. *C. axilliensis* YSMAA5_1_F9 and *C. jamesii* YSMAA1_1_F7 were used as the reference genomes on the plots. *C. axilliensis* YSMAA1_1_D6 was also included in (A) for comparison. Alignments and figures were generated using BRIG.



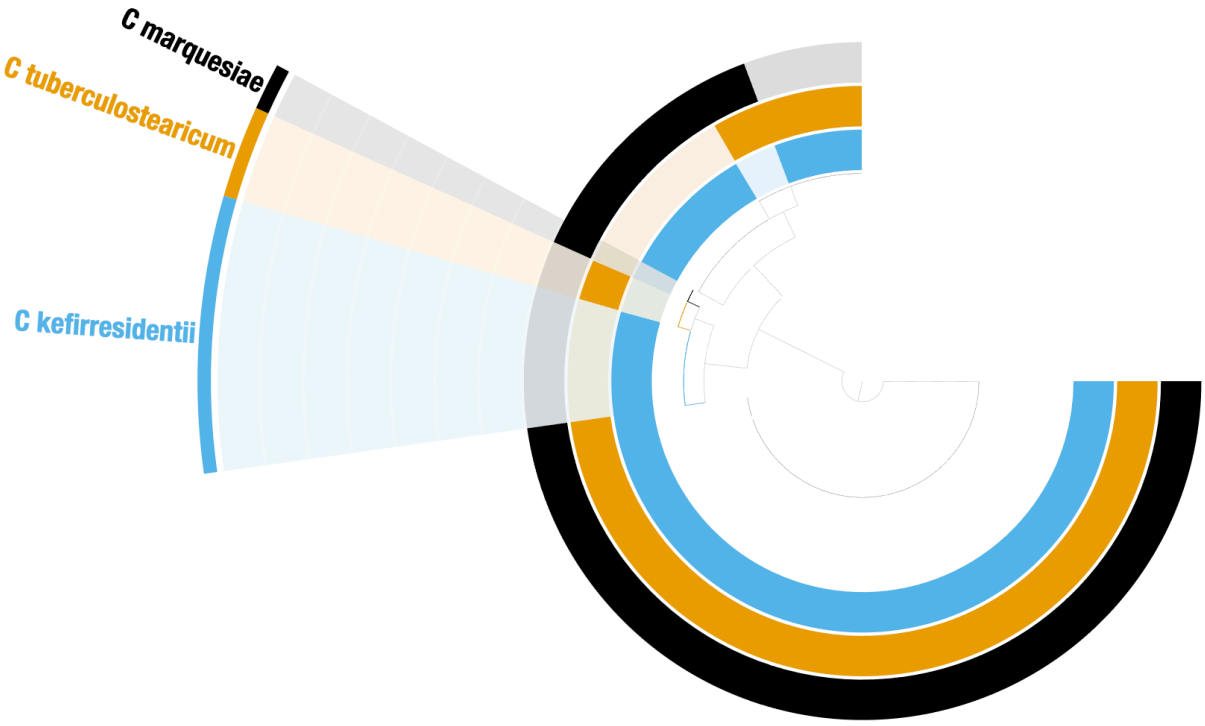
Supplementary Figure 5. A core genome alignment of the representative genomes (bold) and select RefSeq genomes was generated using Roary with a 80% BLASTp cutoff. The 7 species were identified using coloured wedges and the species within the *C. tuberculostearicum* species complex (grey, brown and blue wedges) were also distinguished. The tree was generated using PhyML and visualised on iTOL.



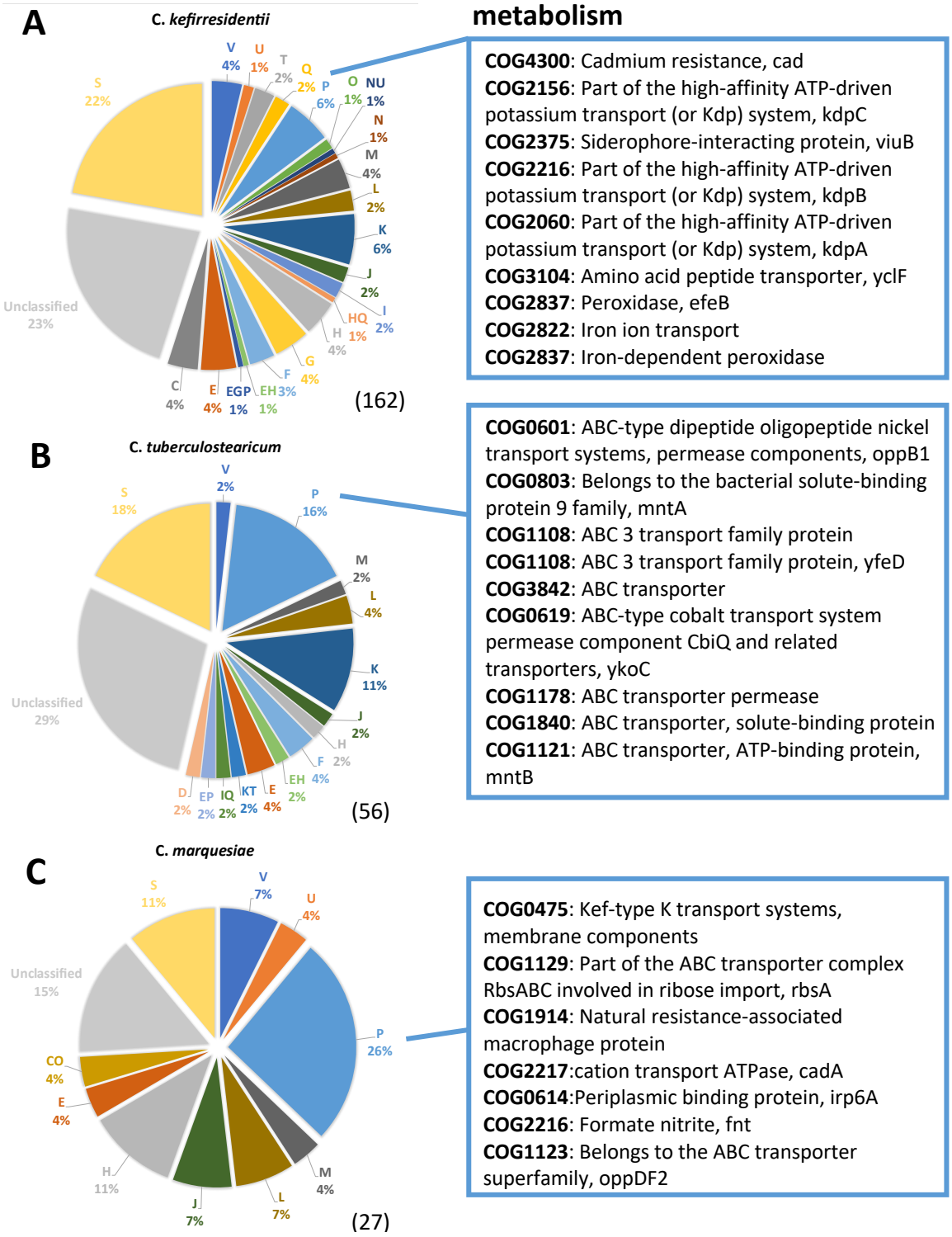
Supplementary Figure 6. COG classification of genes of the *C. axilliensis* (A) core genome and genes specific to (B) YSMAA5_1_F9 and (C) YSMAA1_1_D6. The percentage of genes within each COG classification is calculated against the total number of genes for each analysis. The genes of COG categories with the highest percentage and not unclassified or unknown (S) were shown.



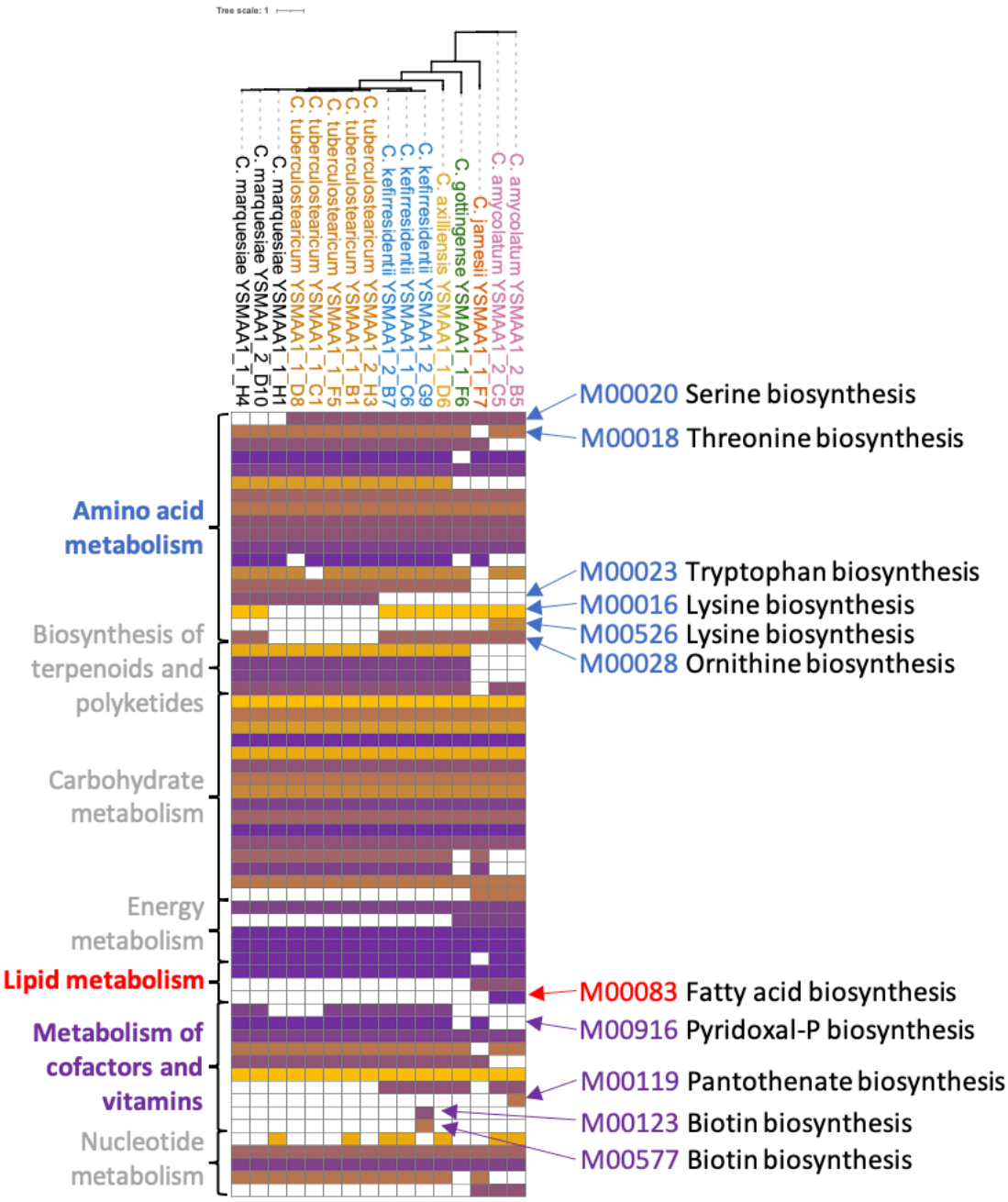
Supplementary Figure 7. COG classification of genes of the *C. jamesii* YSMAA1_1_F7. The percentage of genes within each COG classification is calculated against the total number of genes for each analysis. The genes of COG categories with the highest percentage and not unclassified or unknown (S) were shown.



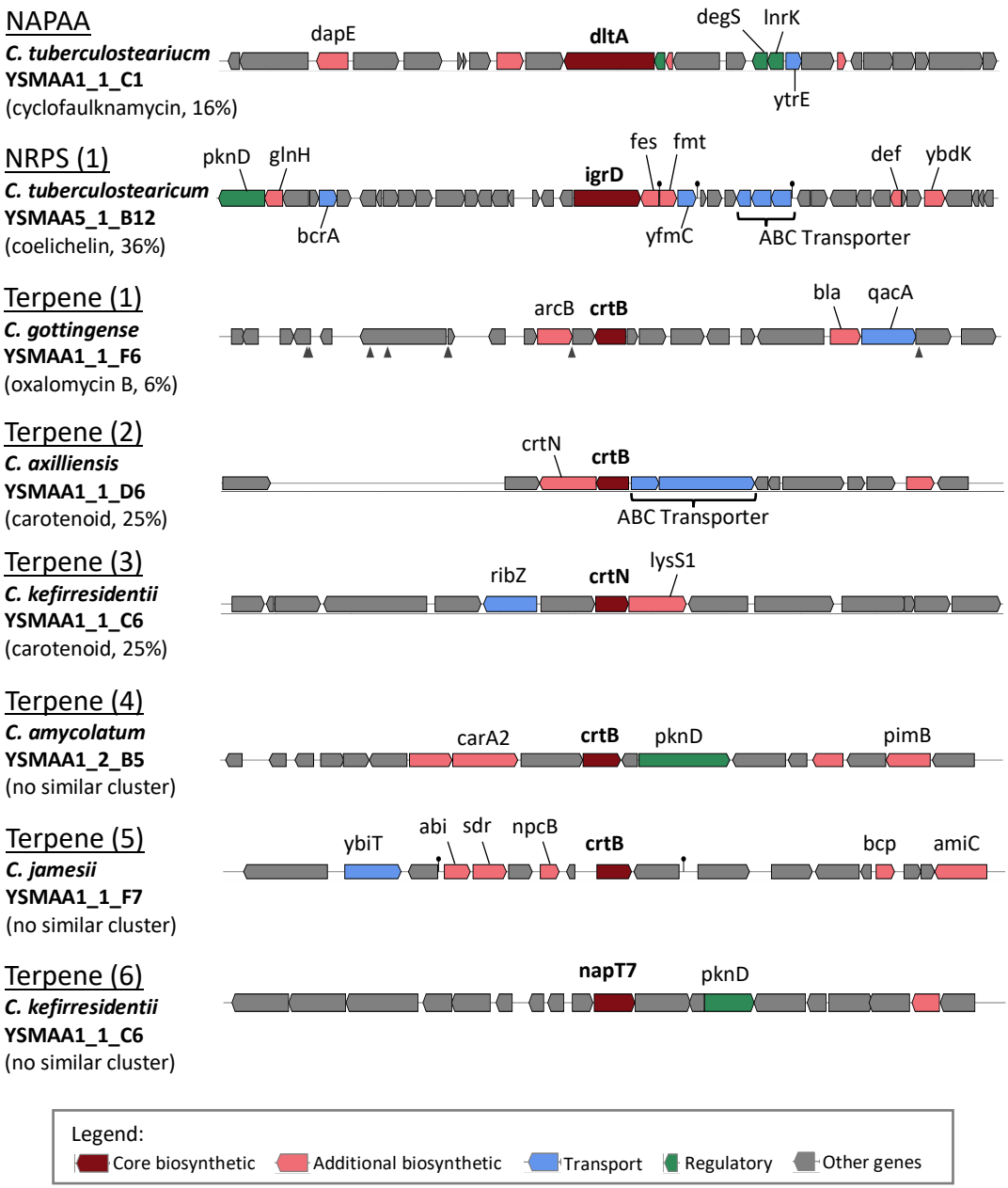
Supplementary Figure 8. Comparisons of the core genomes of species *C. marquesiae*, *C. tuberculostearicum* and *C. kefirresidentii* derived from the respective isolate genomes using anvi'o. Gene clusters specific to each species were highlighted using the coloured wedge.



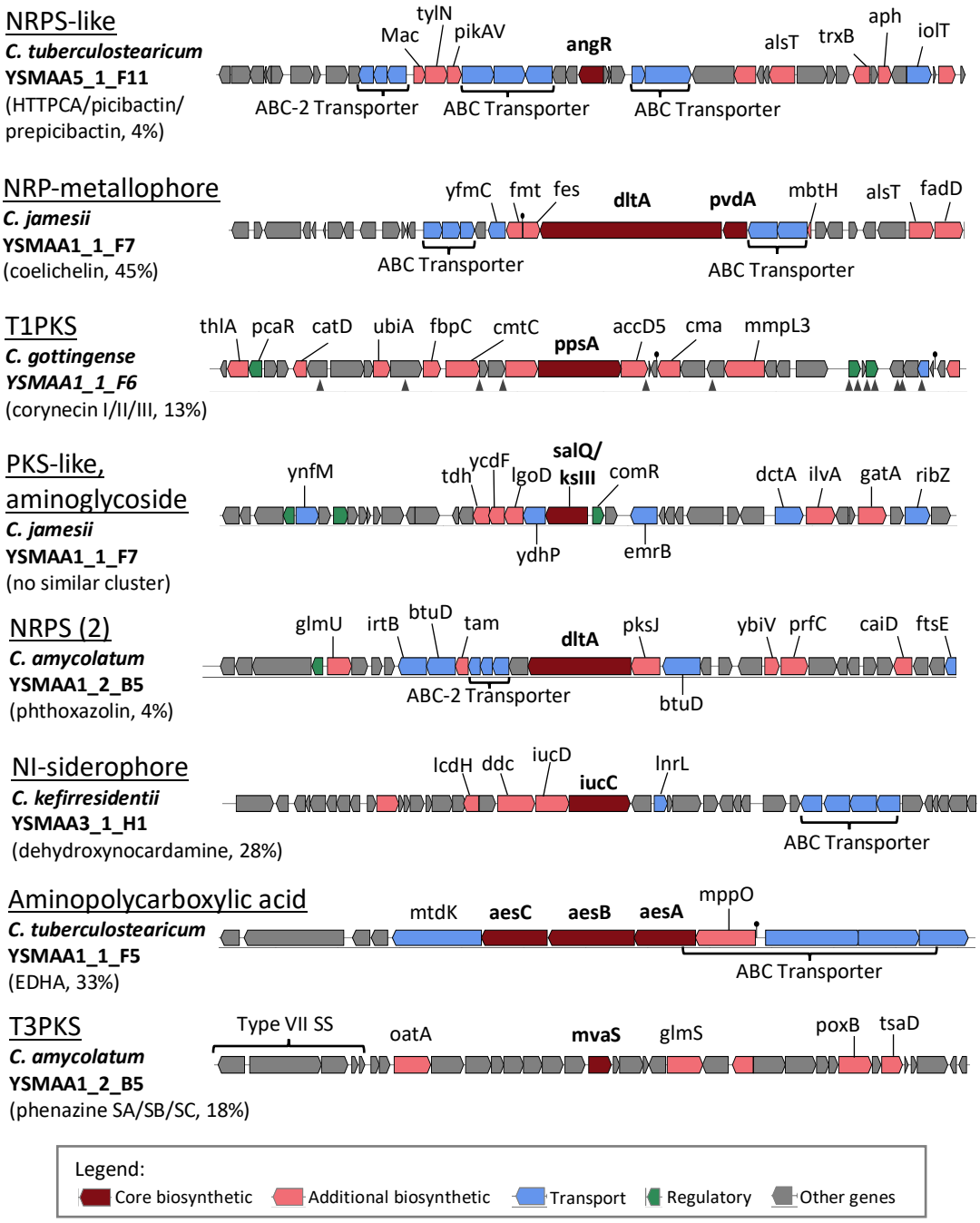
Supplementary Figure 9. COG classification of genes specific to (A) *C. kefirresidentii*, (B) *C. tuberculostearicum* and (C) *C. marquesiae*. The percentage of genes within each COG classification is calculated against the total number of genes for each analysis. The genes of COG categories with the highest percentage and not unclassified or unknown (S) were shown.



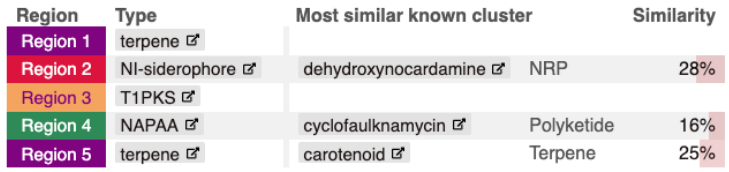
Supplementary Figure 10. Metabolic pathways in axillary corynebacteria of volunteer A1. Each KEGG module for each isolate were placed on a heatmap where the modules with the largest number of reactions are in yellow and lowest in purple with absent modules in white. Modules of interest were described on the right of the heatmap. KEGG modules were grouped according to the categories on the left of the heatmap. Isolates were grouped by species with a core genome tree depicting the variation.



Supplementary Figure 11. Architectures of the 8 distinct putative biosynthetic gene clusters predicted using antiSMASH 7.

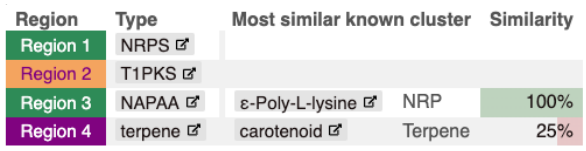


Supplementary Figure 12. Architectures of the 8 distinct putative biosynthetic gene clusters predicted using antiSMASH 7.



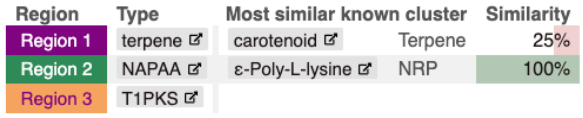
Region	Type	Most similar known cluster		Similarity
Region 1	terpene			
Region 2	NI-siderophore	dehydroxynocardamine	NRP	28%
Region 3	T1PKS			
Region 4	NAPAA	cyclofaulknamycin	Polyketide	16%
Region 5	terpene	carotenoid	Terpene	25%

C. kefirresidentii
A1_1_C6, A1_2_B7, A3_1_H1, A5_1_D11



Region	Type	Most similar known cluster		Similarity
Region 1	NRPS			
Region 2	T1PKS			
Region 3	NAPAA	ε-Poly-L-lysine	NRP	100%
Region 4	terpene	carotenoid	Terpene	25%

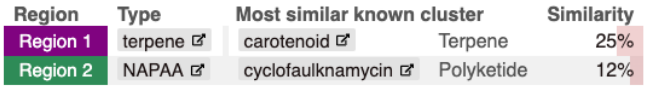
C. kefirresidentii
A1_2_G9



Region	Type	Most similar known cluster		Similarity
Region 1	terpene	carotenoid	Terpene	25%
Region 2	NAPAA	ε-Poly-L-lysine	NRP	100%
Region 3	T1PKS			

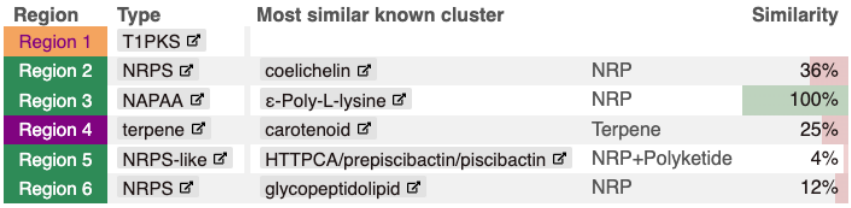
C. marquesiae A1_1_H4 also includes an aminopolycarboxylic acid cluster (most similar known cluster: EDHA, 33%)

C. marquesiae
A1_1_H1, A1_1_H4, A1_2_D10, A3_1_C3, A3_1_F3, A3_1_G3, A4_1_B7, A4_1_E7, A4_1_E8, A5_1_H11



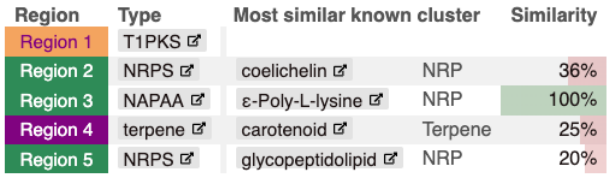
Region	Type	Most similar known cluster		Similarity
Region 1	terpene	carotenoid	Terpene	25%
Region 2	NAPAA	cyclofaulknamycin	Polyketide	12%

C. tuberculostearicum
A3_1_A4



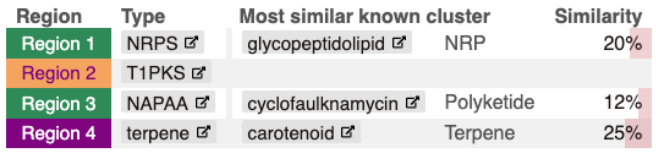
Region	Type	Most similar known cluster		Similarity
Region 1	T1PKS			
Region 2	NRPS	coelichelin	NRP	36%
Region 3	NAPAA	ε-Poly-L-lysine	NRP	100%
Region 4	terpene	carotenoid	Terpene	25%
Region 5	NRPS-like	HTTPCA/prepiscibactin/piscibactin	NRP+Polyketide	4%
Region 6	NRPS	glycopeptidolipid	NRP	12%

C. tuberculostearicum
A5_1_B12, A5_1_F11



Region	Type	Most similar known cluster		Similarity
Region 1	T1PKS			
Region 2	NRPS	coelichelin	NRP	36%
Region 3	NAPAA	ε-Poly-L-lysine	NRP	100%
Region 4	terpene	carotenoid	Terpene	25%
Region 5	NRPS	glycopeptidolipid	NRP	20%

C. tuberculostearicum
A1_2_H3, A5_1_G11

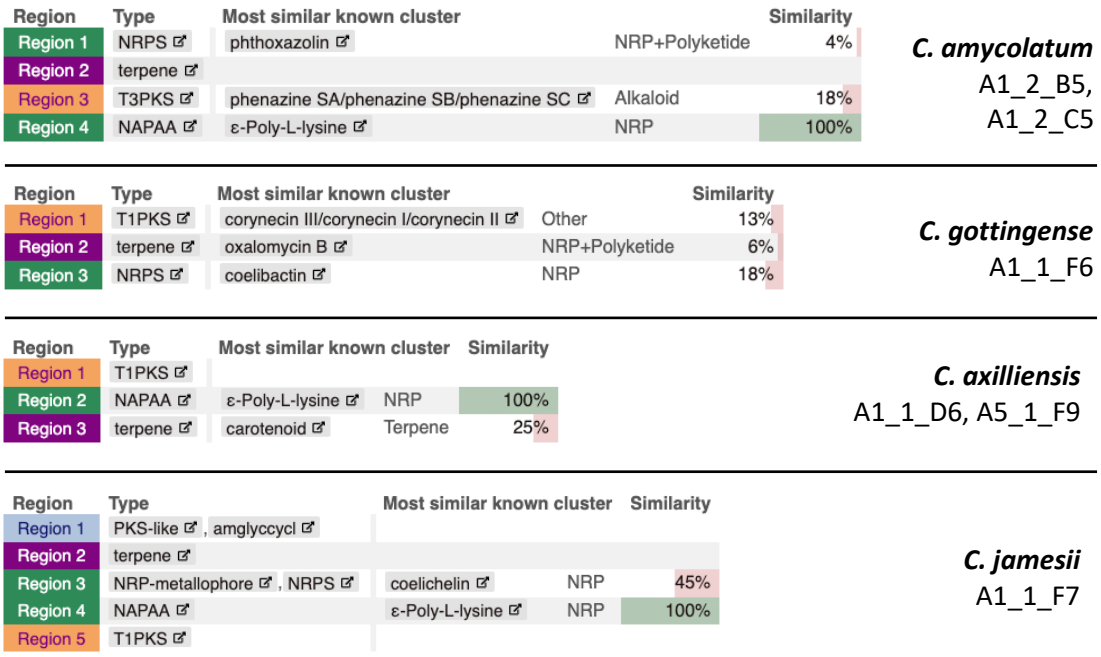


Region	Type	Most similar known cluster		Similarity
Region 1	NRPS	glycopeptidolipid	NRP	20%
Region 2	T1PKS			
Region 3	NAPAA	cyclofaulknamycin	Polyketide	12%
Region 4	terpene	carotenoid	Terpene	25%

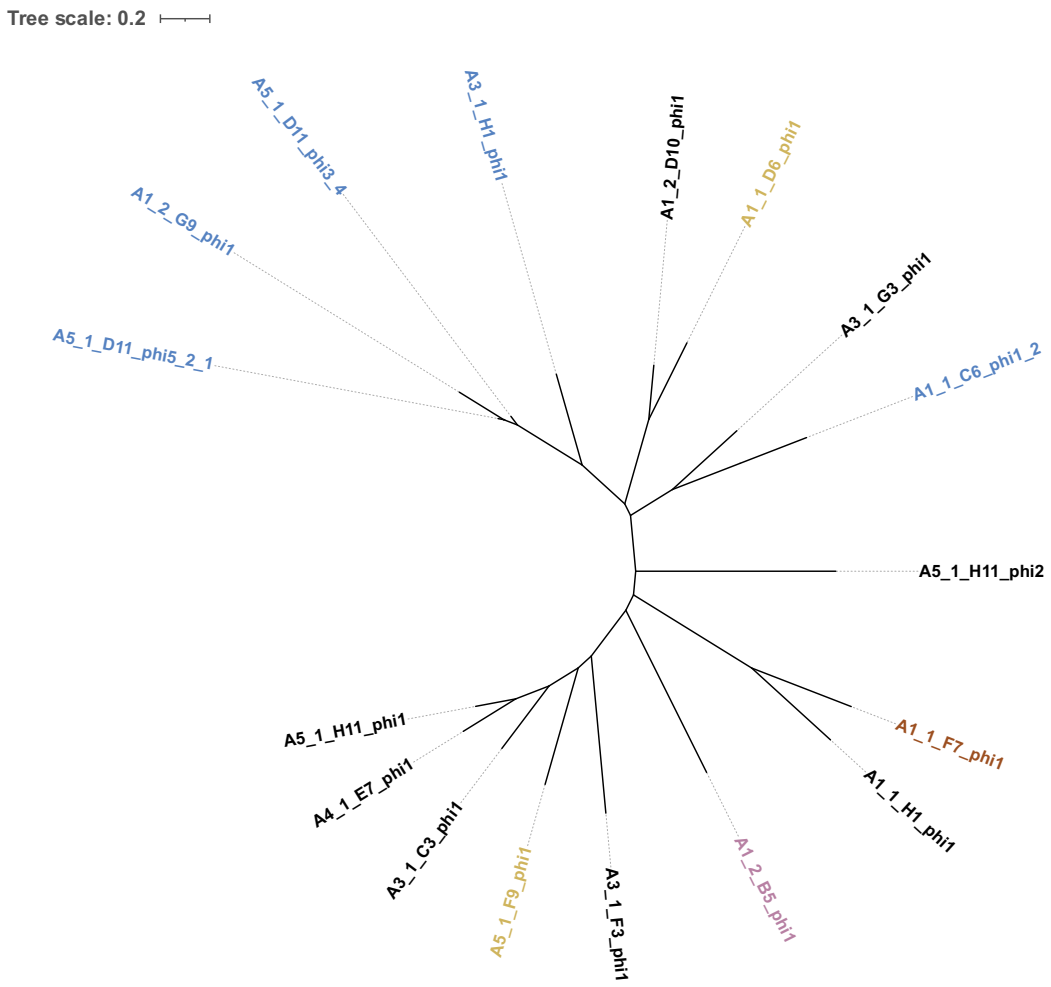
C. tuberculostearicum
A1_1_B1, A1_1_C1, A1_1_D8, A1_1_F5

C. tuberculostearicum A1_1_F5 also includes an aminopolycarboxylic acid cluster (most similar known cluster: EDHA, 33%)

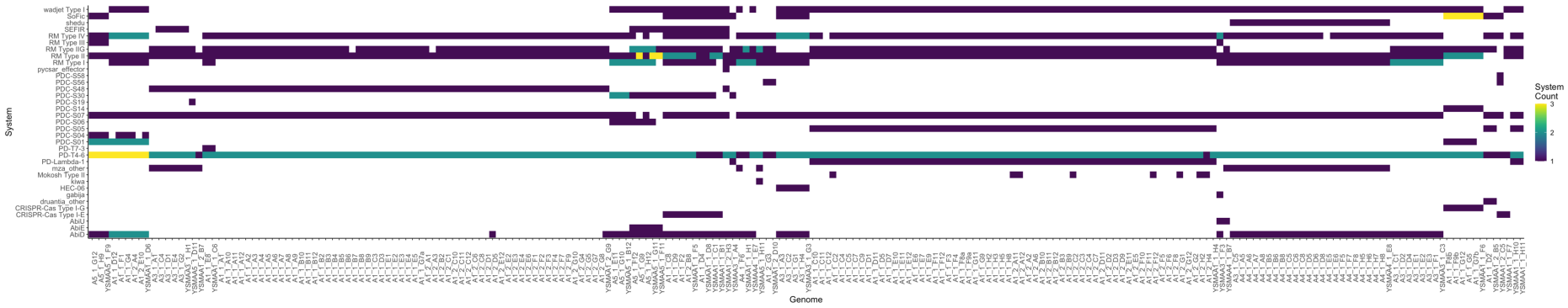
Supplementary Figure 13. The distribution of putative biosynthetic gene clusters in the *C. tuberculostearicum* species complex predicted using antiSMASH 7.



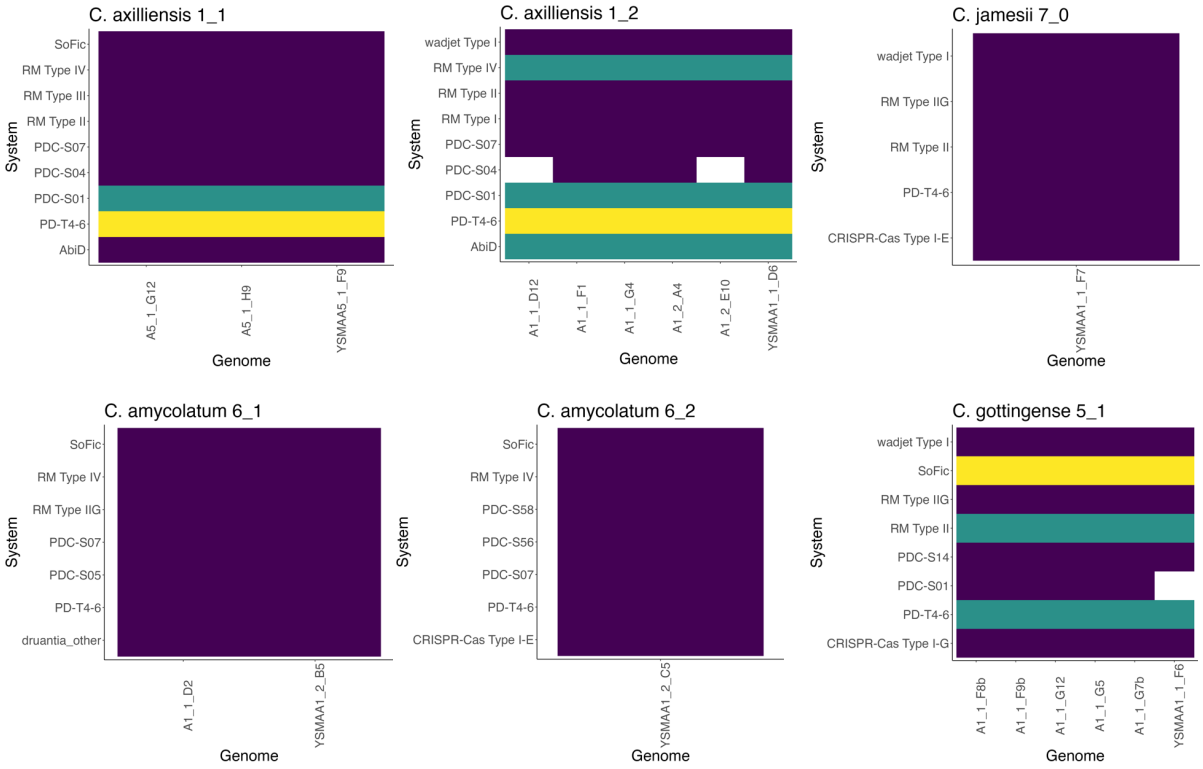
Supplementary Figure 14. The distribution of putative biosynthetic gene clusters in the *C. amycolatum*, *C. gottingense*, *C. axilliensis* and *C. jamesii* isolates predicted using antiSMASH 7.



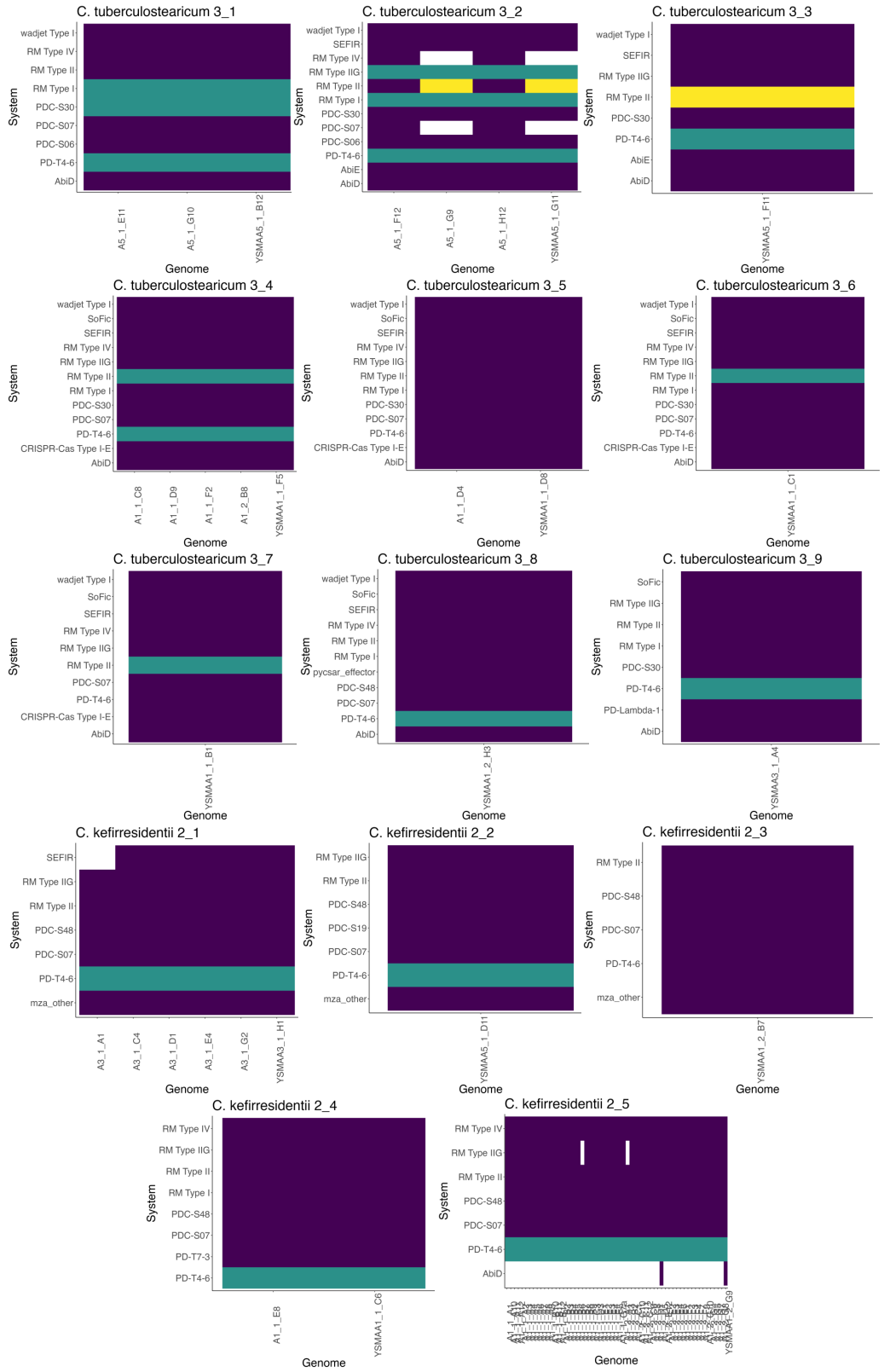
Supplementary Figure 15. None of the putative prophages were completely identical to each other. Putative prophages were aligned and compared using an unrooted tree generated with PhyML and visualised on iTOL. Putative prophages were coloured by the species of the source isolate (Blue: *C. kefirresidentii*, Black: *C. marquesiae*, Gold: *C. axilliensis*, Orange: *C. jamesii* and Pink: *C. amycolatum*).



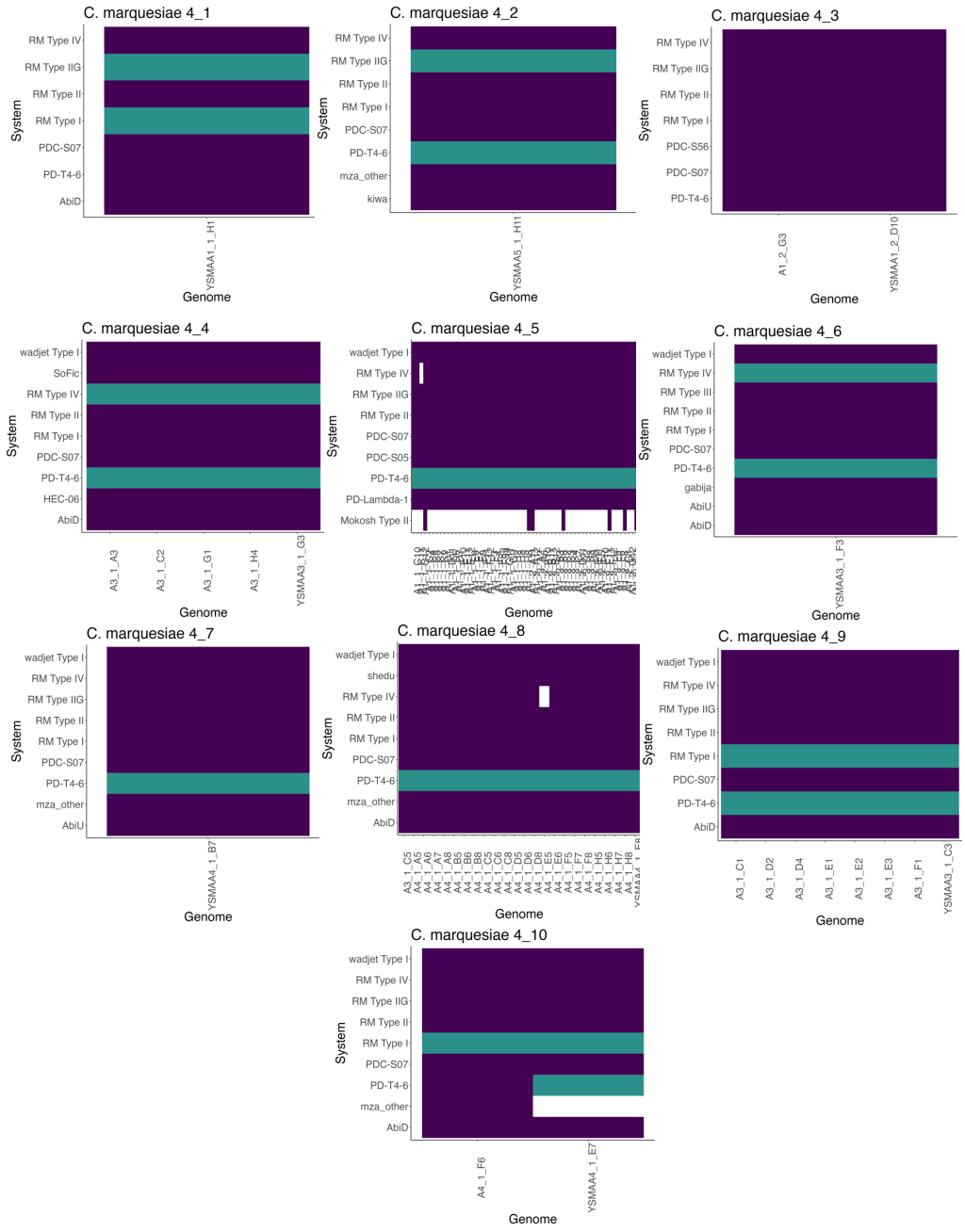
Supplementary Figure 16. Global PADLOC analysis reveals profiles of phage defence systems are generally conserved within drep secondary clusters.



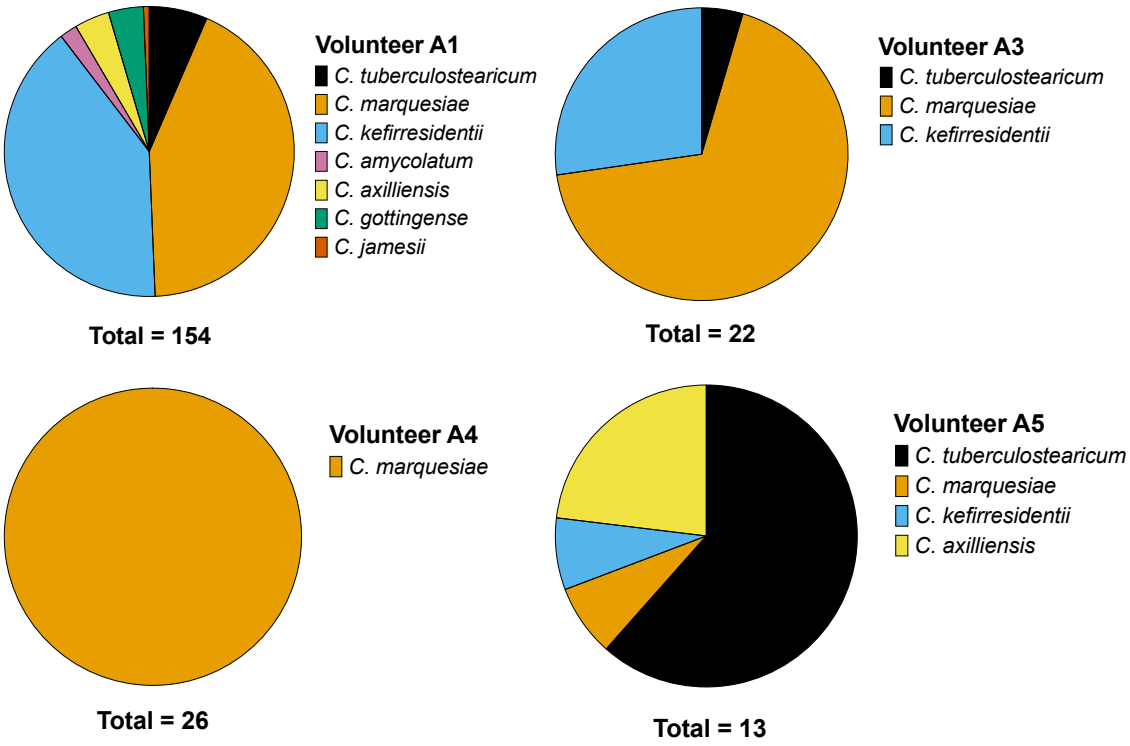
Supplementary Figure 17. Phage defence systems found in drep secondary clusters of (A-B) *C. axilliensis*, (C) *C. jamesii*, (D-E) *C. amycolatum* and (F) *C. gottingense*. The heatmap describes the number of each identified system with (Yellow: 3, Green: 2, Purple: 1 and White: 0).



Supplementary Figure 18. Phage defence systems found in drep secondary clusters of (A-I) *C. tuberculostearicum* and (J-N) *C. kefirresidentii.* The heatmap describes the number of each identified system with (Yellow: 3, Green: 2, Purple: 1 and White: 0).



Supplementary Figure 19. Phage defence systems found in drep secondary clusters of (A-J) *C. marquesiae*. The heatmap describes the number of each identified system with (Yellow: 3, Green: 2, Purple: 1 and White: 0).



Supplementary Figure 20. The proportion of the total number of isolates derived from each of the 4 volunteers which bin to the 7 identified species (*C. tuberculostearicum*, *C. marquesiae*, *C. kefirresidentii*, *C. amycolatum*, *C. axilliensis*, *C. gottingense* and *C. jamesii*)

Isolate	Classification	FastANI Reference	FastANI Reference Radius	FastANI ANI	FastANI Alignment Fraction	Other Related References	MSA AA Percent	RED Value
YSMAA1_2_B5	g__Corynebacterium;s__Corynebacterium amycolatum	GCF_016889425.1	95	99.49	0.978	GCF_000173655.1	96.62	-
YSMAA1_2_C5	g__Corynebacterium;s__Corynebacterium amycolatum	GCF_016889425.1	95	98.55	0.959	GCF_000173655.1	96.6	-
YSMAA1_1_F6	g__Corynebacterium;s__Corynebacterium gottingense	GCF_003693265.1	95	97.62	0.927	GCF_001831515.1	95.65	-
YSMAA1_1_D6^	g__Corynebacterium;s__Corynebacterium sp943912665	GCA_943912665.1	95	96.42	0.928	GCF_000833575.1	96.17	-
YSMAA5_1_F9^	g__Corynebacterium;s__Corynebacterium sp943912665	GCA_943912665.1	95	97.51	0.952	GCF_000833575.1	95.73	-
YSMAA1_1_F7^	g__Corynebacterium;s__Corynebacterium sp943913395	GCA_943913395.1	95	96.82	0.953	GCF_000738175.1	95.39	-
YSMAA1_1_C6	g__Corynebacterium;s__Corynebacterium kefirresidentii	GCA_002154655.1	95	96.37	0.901	GCF_022288805.1	95.89	-
YSMAA1_2_B7	g__Corynebacterium;s__Corynebacterium kefirresidentii	GCA_002154655.1	95	96.43	0.9	GCF_022288805.1	95.85	-
YSMAA1_2_G9	g__Corynebacterium;s__Corynebacterium kefirresidentii	GCA_002154655.1	95	96.42	0.899	GCF_022288805.1	95.15	-
YSMAA3_1_H1	g__Corynebacterium;s__Corynebacterium kefirresidentii	GCA_002154655.1	95	96.5	0.908	GCF_022288805.1	95.65	-
YSMAA5_1_D11	g__Corynebacterium;s__Corynebacterium kefirresidentii	GCA_002154655.1	95	96.31	0.903	GCF_022288805.1	96.17	-
YSMAA1_1_C1	g__Corynebacterium;s__Corynebacterium tuberculostearicum	GCF_013408445.1	96.3332	96.4	0.92	GCF_013408445.1	96.21	-
YSMAA1_1_D8	g__Corynebacterium;s__Corynebacterium tuberculostearicum	GCF_013408445.1	96.3332	96.39	0.928	GCF_000175635.1	96.23	-
YSMAA1_1_F5	g__Corynebacterium;s__Corynebacterium tuberculostearicum	GCF_013408445.1	96.3332	96.42	0.923	GCF_000175635.1	96.21	-
YSMAA1_2_H3	g__Corynebacterium;s__Corynebacterium tuberculostearicum	GCF_013408445.1	96.3332	96.49	0.938	GCF_016894265.1	96.15	-
YSMAA3_1_A4	g__Corynebacterium;s__Corynebacterium tuberculostearicum	GCF_013408445.1	96.3332	96.33	0.924	GCF_013408445.1	92.63	-
YSMAA5_1_B12	g__Corynebacterium;s__Corynebacterium tuberculostearicum	GCF_013408445.1	96.3332	96.61	0.903	GCF_000175635.1	95.83	-
YSMAA5_1_G11	g__Corynebacterium;s__Corynebacterium tuberculostearicum_C	GCF_000175635.1	96.3332	96.47	0.934	GCF_013408445.1	94.2	-
YSMAA1_1_B1*	g__Corynebacterium;s__	-	-	-	-	GCF_013408445.1	96.21	0.99902
YSMAA5_1_F11*	g__Corynebacterium;s__	-	-	-	-	GCF_000175635.1	96.01	0.999
YSMAA1_1_H1	g__Corynebacterium;s__Corynebacterium aurimucosum_E	GCF_016127015.1	95	95.58	0.911	GCF_013408445.1	95.37	-
YSMAA1_1_H4	g__Corynebacterium;s__Corynebacterium aurimucosum_E	GCF_016127015.1	95	97.24	0.94	GCF_022347205.1	94.9	-
YSMAA1_2_D10	g__Corynebacterium;s__Corynebacterium aurimucosum_E	GCF_016127015.1	95	96.54	0.91	GCF_022347205.1	95.75	-
YSMAA3_1_C3	g__Corynebacterium;s__Corynebacterium aurimucosum_E	GCF_016127015.1	95	99.33	0.97	GCF_000175635.1	95.47	-
YSMAA3_1_F3	g__Corynebacterium;s__Corynebacterium aurimucosum_E	GCF_016127015.1	95	98.05	0.935	GCF_000175635.1	95.73	-
YSMAA3_1_G3	g__Corynebacterium;s__Corynebacterium aurimucosum_E	GCF_016127015.1	95	97.35	0.92	GCF_000175635.1	95.45	-
YSMAA4_1_B7	g__Corynebacterium;s__Corynebacterium aurimucosum_E	GCF_016127015.1	95	97.52	0.941	GCF_000175635.1	95.41	-
YSMAA4_1_E7	g__Corynebacterium;s__Corynebacterium aurimucosum_E	GCF_016127015.1	95	99.62	0.98	GCF_000175635.1	96.15	-
YSMAA4_1_E8	g__Corynebacterium;s__Corynebacterium aurimucosum_E	GCF_016127015.1	95	98.31	0.949	GCF_022347205.1	95.83	-
YSMAA5_1_H11	g__Corynebacterium;s__Corynebacterium aurimucosum_E	GCF_016127015.1	95	95.6	0.915	GCF_013408445.1	95.77	-

Supplementary Table 1. Representative genomes were compared on GTDB-tk to identify the closest species on the Genome Taxonomy Database (Classification) using FastANI with the threshold for each classification defined on the “FastANI Reference Radius” tab. Isolates are coloured by classification groups.

Species	Strain (ANI ≥ 99.5%)	Clones	Representative isolate
C. axilliensis	1_2	6	YSMAA1_1_D6
C. kefirresidentii	2_3	1	YSMAA1_2_B7
C. kefirresidentii	2_4	2	YSMAA1_1_C6
C. kefirresidentii	2_5	59	YSMAA1_2_G9
C. tuberculostearicum	3_4	5	YSMAA1_1_F5
C. tuberculostearicum	3_5	2	YSMAA1_1_D8
C. tuberculostearicum	3_6	1	YSMAA1_1_C1
C. tuberculostearicum	3_7	1	YSMAA1_1_B1
C. tuberculostearicum	3_8	1	YSMAA1_2_H3
C. marquesiae	4_1	1	YSMAA1_1_H1
C. marquesiae	4_3	2	YSMAA1_2_D10
C. marquesiae	4_5	63	YSMAA1_1_H4
C. gottingense	5_1	6	YSMAA1_1_F6
C. amycolatum	6_1	2	YSMAA1_2_B5
C. amycolatum	6_2	1	YSMAA1_2_C5
C. jamesii	7_0	1	YSMAA1_1_F7

Supplementary Table 2. Summary of *Corynebacterium* isolates from volunteer A1. The initial pool of 154 isolates were sorted into specific species (ANI ≥ 95%) and strains (ANI ≥ 99.5%) using drep. Species were identified using GTDB-tk. Number of clones and the representative isolates for each strain were included.

Species	Strain (ANI ≥ 99.5%)	Clones	Representative isolate
C. kefirresidentii	2_1	6	YSMAA3_1_H1
C. tuberculostearicum	3_9	1	YSMAA3_1_A4
C. marquesiae	4_4	5	YSMAA3_1_G3
C. marquesiae	4_6	1	YSMAA3_1_F3
C. marquesiae	4_8	1	YSMAA4_1_E8
C. marquesiae	4_9	8	YSMAA3_1_C3

Supplementary Table 3. Summary of *Corynebacterium* isolates from volunteer A3. The initial pool of 22 isolates were sorted into specific species (ANI ≥ 95%) and strains (ANI ≥ 99.5%) using drep. Species were identified using GTDB-tk. Number of clones and the representative isolates for each strain were included.

Species	Strain (ANI ≥ 99.5%)	Clones	Representative isolate
C. marquesiae	4_7	1	YSMAA4_1_B7
C. marquesiae	4_8	23	YSMAA4_1_E8
C. marquesiae	4_10	2	YSMAA4_1_E7

Supplementary Table 4. Summary of *Corynebacterium* isolates from volunteer A4. The initial pool of 26 isolates were sorted into specific species (ANI ≥ 95%) and strains (ANI ≥ 99.5%) using drep. Species were identified using GTDB-tk. Number of clones and the representative isolates for each strain were included.

Species	Strain (ANI ≥ 99.5%)	Clones	Representative isolate
C. axilliensis	1_1	3	YSMAA5_1_F9
C. kefirresidentii	2_2	1	YSMAA5_1_D11
C. tuberculostearicum	3_1	3	YSMAA5_1_B12
C. tuberculostearicum	3_2	4	YSMAA5_1_G11
C. tuberculostearicum	3_3	1	YSMAA5_1_F11
C. marquesiae	4_2	1	YSMAA5_1_H11

Supplementary Table 5. Summary of *Corynebacterium* isolates from volunteer A5. The initial pool of 13 isolates were sorted into specific species (ANI ≥ 95%) and strains (ANI ≥ 99.5%) using drep. Species were identified using GTDB-tk. Number of clones and the representative isolates for each strain were included.

Isolate	1	2	3	4	5	6
YSMAA1_1_C1	**erm(X)**, U21300, [lincomycin, clindamycin, erythromycin, quinupristin, pristinamycin ia, virginiamycin s]					
YSMAA1_1_D6	**erm(X)**, X51472, [lincomycin, clindamycin, erythromycin, quinupristin, pristinamycin ia, virginiamycin s]					
YSMAA1_1_F6	**erm(X)**, X51472, [lincomycin, clindamycin, erythromycin, quinupristin, pristinamycin ia, virginiamycin s]					
YSMAA1_1_H4	**aac(3)-XI**, CTEG01000046, [unknown aminoglycoside]					
YSMAA1_2_B7	**erm(X)**, X51472, [lincomycin, clindamycin, erythromycin, quinupristin, pristinamycin ia, virginiamycin s]					
YSMAA3_1_A4	**erm(X)**, X51472, [lincomycin, clindamycin, erythromycin, quinupristin, pristinamycin ia, virginiamycin s]	**aac(3)-XI**, CTEG01000046, [unknown aminoglycoside]				
YSMAA3_1_C3	**erm(X)**, U21300, [lincomycin, clindamycin, erythromycin, quinupristin, pristinamycin ia, virginiamycin s]					
YSMAA3_1_F3	**erm(X)**, U21300, [lincomycin, clindamycin, erythromycin, quinupristin, pristinamycin ia, virginiamycin s]	**erm(X)**, X51472, [lincomycin, clindamycin, erythromycin, quinupristin, pristinamycin ia, virginiamycin s]				
YSMAA3_1_G3	**erm(X)**, U21300, [lincomycin, clindamycin, erythromycin, quinupristin, pristinamycin ia, virginiamycin s]					
YSMAA4_1_B7	**erm(X)**, U21300, [lincomycin, clindamycin, erythromycin, quinupristin, pristinamycin ia, virginiamycin s]					
YSMAA4_1_E7	**erm(X)**, U21300, [lincomycin, clindamycin, erythromycin, quinupristin, pristinamycin ia, virginiamycin s]					
YSMAA5_1_B12	**erm(X)**, U21300, [lincomycin, clindamycin, erythromycin, quinupristin, pristinamycin ia, virginiamycin s]	**tet(Z)**, AF121000, [tetracycline, doxycycline]				
YSMAA5_1_F9	**erm(X)**, X51472, [lincomycin, clindamycin, erythromycin, quinupristin, pristinamycin ia, virginiamycin s]	**aac(3)-XI**, CTEG01000046, [unknown aminoglycoside]	**cmx**, U85507, [chloramphenicol]	**aph(6)-Id**, M28829, [streptomycin]	**aph(3")-Ib**, AF321551, [streptomycin]	**aph(3')-Ia**, X62115, [kanamycin, neomycin, lividomycin, paromomycin, ribostamycin]
YSMAA5_1_F11	**erm(X)**, U21300, [lincomycin, clindamycin, erythromycin, quinupristin, pristinamycin ia, virginiamycin s]	**ant(2")-Ia**, X04555, [gentamycin, tobramycin]				
YSMAA5_1_G11	**erm(X)**, U21300, [lincomycin, clindamycin, erythromycin, quinupristin, pristinamycin ia, virginiamycin s]	**ant(2")-Ia**, X04555, [gentamycin, tobramycin]				

Supplementary Table 6. Acquired resistance gene cassettes predicted to be present in the representative isolates using ResFinder.

Isolate	Clindamycin	Doxycycline	Penicillin G	Chloramphenicol	Vancomycin	Fosfomycin
YSMAA1_1_B1	-	-	-	-	-	+
YSMAA1_1_C1	-	-	-	-	-	+
YSMAA1_1_C6	-	-	-	-	-	+
YSMAA1_1_D6	-	-	-	-	-	+
YSMAA1_1_D8	-	-	-	-	-	+
YSMAA1_1_F5	-	-	-	-	-	+
YSMAA1_1_F6	+	-	-	-	-	+
YSMAA1_1_F7	-	-	+	-	-	+
YSMAA1_1_H1	-	-	-	-	-	+
YSMAA1_1_H4	-	-	-	-	-	+
YSMAA1_2_B5	-	-	-	-	-	+
YSMAA1_2_B7	+	-	-	-	-	+
YSMAA1_2_C5	-	-	-	-	-	+
YSMAA1_2_D10	-	-	-	-	-	+
YSMAA1_2_G9	-	-	-	-	-	+
YSMAA1_2_H3	-	-	-	-	-	+
YSMAA3_1_A4	+	-	-	-	-	+
YSMAA3_1_C3	+	-	-	-	-	+
YSMAA3_1_F3	+	-	-	-	-	+
YSMAA3_1_G3	-	-	-	-	-	+
YSMAA3_1_H1	-	-	-	-	-	+
YSMAA4_1_B7	+	-	-	-	-	+
YSMAA4_1_E7	+	-	-	-	-	+
YSMAA4_1_E8	-	-	-	-	-	+
YSMAA5_1_B12	+	-	-	-	-	+
YSMAA5_1_D11	-	-	-	-	-	+
YSMAA5_1_F9	-	-	-	+	-	+
YSMAA5_1_F11	+	-	-	-	-	+
YSMAA5_1_G11	+	-	-	-	-	+
YSMAA5_1_H11	-	-	-	-	-	+

Supplementary Table 7. Antibiotic susceptibility disc diffusion assay performed on the representative isolates using the EUCAST protocol. “+” refers to resistance phenotype observed while “-” refers to antibiotic susceptibility.

4.15 References

1. Byrd AL, Belkaid Y, Segre JA. 2018. The human skin microbiome. Nature Reviews Microbiology 2018 16:143–155.

2. Oh J, Byrd AL, Park M, Kong HH, Segre JA. 2016. Temporal Stability of the Human Skin Microbiome. Cell 165:854–866.

3. Oh J, Byrd AL, Deming C, Conlan S, Kong HH, Segre JA, Barnabas B, Blakesley R, Bouffard G, Brooks S, Coleman H, Dekhtyar M, Gregory M, Guan X, Gupta J, Han J, Ho SL, Legaspi R, Maduro Q, Masiello C, Maskeri B, McDowell J, Montemayor C, Mullikin J, Park M, Riebow N, Schandler K, Schmidt B, Sison C, Stantripop M, Thomas J, Thomas P, Vemulapalli M, Young A. 2014. Biogeography and individuality shape function in the human skin metagenome. Nature 2014 514:7520 514:59–64.

4. Grice EA, Kong HH, Conlan S, Deming CB, Davis J, Young AC, Bouffard GG, Blakesley RW, Murray PR, Green ED, Turner ML, Segre JA. 2009. Topographical and temporal diversity of the human skin microbiome. Science 324:1190–1192.

5. POLLOCK MR, WAINWRIGHT SD, MANSON EE. 1949. The presence of oleic acid-requiring diphtheroids on human skin. J Pathol Bacteriol 61:274–276.

6. McGinley KJ, Labows JN, Zechman JM, Nordstrom KM, Webster GF, Leyden JJ. 1985. Analysis of Cellular Components, Biochemical Reactions, and Habitat of Human Cutaneous Lipophilic Diphtheroids. Journal of Investigative Dermatology 85:374–377.

7. Tauch A, Kaiser O, Hain T, Goesmann A, Weisshaar B, Albersmeier A, Bekel T, Bischoff N, Brune I, Chakraborty T, Kalinowski J, Meyer F, Rupp O, Schneiker S, Viehoever P, Pühler A. 2005. Complete genome sequence and analysis of the multiresistant nosocomial pathogen Corynebacterium jeikeium K411, a lipid-requiring bacterium of the human skin flora. J Bacteriol 187:4671–4682.

8. Esteban J, Nieto E, Calvo R, Fernández-Roblas R, Valero-Guillén PL, Soriano F. 1999. Microbiological characterization and clinical significance of Corynebacterium amycolatum strains. European Journal of Clinical Microbiology and Infectious Diseases 18:518–521.

9. Paviour S, Musaad S, Roberts S, Taylor G, Taylor S, Shore K, Lang S, Holland D. 2002. Corynebacterium species isolated from patients with mastitis. Clinical Infectious Diseases 35:1434–1440.

10. Ridaura VK, Bouladoux N, Claesen J, Erin Chen Y, Byrd AL, Constantinides MG, Merrill ED, Tamoutounour S, Fischbach MA, Belkaid Y. 2018. Contextual control of skin immunity and inflammation by Corynebacterium. Journal of Experimental Medicine 215:785–799.

11. Altonsy MO, Kurwa HA, Lauzon GJ, Amrein M, Gerber AN, Almishri W, Mydlarski PR. 2020. Corynebacterium tuberculostearicum, a human skin colonizer, induces the canonical nuclear factor-κB inflammatory signaling pathway in human skin cells. Immun Inflamm Dis 8:62–79.

12. Bomar L, Brugger SD, Yost BH, Davies SS, Lemon KP. 2016. Corynebacterium accolens releases antipneumococcal free fatty acids from human nostril and skin surface triacylglycerols. mBio 7:10-1128.

13. Horn KJ, Jaberi Vivar AC, Arenas V, Andani S, Janoff EN, Clark SE. 2022. Corynebacterium Species Inhibit Streptococcus pneumoniae Colonization and Infection of the Mouse Airway. Front Microbiol 12:804935.

14. Ramsey MM, Freire MO, Gabrilska RA, Rumbaugh KP, Lemon KP. 2016. Staphylococcus aureus Shifts toward commensalism in response to corynebacterium species. Front Microbiol 7:212804.

15. Rudden M, Herman R, Rose M, Bawdon D, Cox DS, Dodson E, Holden MTG, Wilkinson AJ, James AG, Thomas GH. 2020. The molecular basis of thioalcohol production in human body odour. Sci Rep 10:12500.

16. Troccaz M, Gaïa N, Beccucci S, Schrenzel J, Cayeux I, Starkenmann C, Lazarevic V. 2015. Mapping axillary microbiota responsible for body odours using a culture-independent approach. Microbiome 3:1–15.

17. Natsch A, Emter R. 2020. The specific biochemistry of human axilla odour formation viewed in an evolutionary context. Philosophical Transactions of the Royal Society B 375: 20190269.

18. Emter R, Natsch A. 2008. The sequential action of a dipeptidase and a beta-lyase is required for the release of the human body odorant 3-methyl-3-sulfanylhexan-1-ol from a secreted Cys-Gly-(S) conjugate by Corynebacteria. J Biol Chem 283:20645–52.

19. Natsch A, Schmid J, Flachsmann F. 2004. Identification of Odoriferous Sulfanylalkanols in Human Axilla Secretions and Their Formation through Cleavage of Cysteine Precursors by a C□S Lyase Isolated from Axilla bacteria. Chem Biodivers 1:1058–1072.

20. Callewaert C, Kerckhof FM, Granitsiotis MS, van Gele M, van de Wiele T, Boon N. 2013. Characterization of Staphylococcus and Corynebacterium Clusters in the Human Axillary Region. PLoS One 8:e70538.

21. Ahmed N, Joglekar P, Deming C, Lemon KP, Kong HH, Segre JA, Conlan S, Barnabas BB, Black S, Bouffard GG, Brooks SY, Crawford J, Marfani H, Dekhtyar L, Han J, Ho S-L, Legaspi R, Maduro QL, Masiello CA, McDowell JC, Montemayor C, Mullikin JC, Park M, Riebow NL, Schandler K, Schmidt B, Sison C, Stantripop S, Thomas JW, Thomas PJ, Vemulapalli M, Young AC. 2023. Genomic characterization of the C. tuberculostearicum species complex, a prominent member of the human skin microbiome. mSystems 8: e00632-23.

22. Jensen MG, Svraka L, Baez E, Lund M, Poehlein A, Brüggemann H. 2023. Species- and strain-level diversity of Corynebacteria isolated from human facial skin. BMC Microbiol 23:1–14.

23. Olm MR, Brown CT, Brooks B, Banfield JF. 2017. dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. ISME J 11:2864–2868.

24. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. 2022. GTDB-Tk v2: memory friendly classification with the genome taxonomy database. Bioinformatics 38:5315–5316.

25. Chivian D, Jungbluth SP, Dehal PS, Wood-Charlson EM, Canon RS, Allen BH, Clark MM, Gu T, Land ML, Price GA, Riehl WJ, Sneddon MW, Sutormin R, Zhang Q, Cottingham RW, Henry CS, Arkin AP. 2022. Metagenome-assembled genome extraction and analysis from microbiomes using KBase. Nature Protocols 18:208–238.

26. Salamzade R, Swaney MH, Kalan LR. 2023. Comparative Genomic and Metagenomic Investigations of the Corynebacterium tuberculostearicum Species Complex Reveals Potential Mechanisms Underlying Associations To Skin Health and Disease. Microbiol Spectr 11: e03578-22.

27. Cappelli EA, Ksiezarek M, Wolf J, Neumann-Schaal M, Ribeiro TG, Peixe L. 2023. Expanding the Bacterial Diversity of the Female Urinary Microbiome: Description of Eight New Corynebacterium Species. Microorganisms 11: 388.

28. Taylor D, Daulby A, Grimshaw S, James G, Mercer J, Vaziri S. 2003. Characterization of the microflora of the human axilla. Int J Cosmet Sci 25:137–145.

29. Bawdon D, Cox DS, Ashford D, James AG, Thomas GH. 2015. Identification of axillary Staphylococcus sp. involved in the production of the malodorous thioalcohol 3-methyl-3-sufanylhexan-1-ol. FEMS Microbiol Lett 362:1–10.

30. Atasayar E, Zimmermann O, Spröer C, Schumann P, Groß U. 2017. Corynebacterium gottingense sp. Nov., isolated from a clinical patient. Int J Syst Evol Microbiol 67:4494–4499.

31. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. Nature Communications 9:1–8.

32. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, Fookes M, Falush D, Keane JA, Parkhill J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics 31:3691–3693.

33. Eren AM, Kiefl E, Shaiber A, Veseli I, Miller SE, Schechter MS, Fink I, Pan JN, Yousef M, Fogarty EC, Trigodet F, Watson AR, Esen ÖC, Moore RM, Clayssen Q, Lee MD, Kivenson V, Graham ED, Merrill BD, Karkman A, Blankenberg D, Eppley JM, Sjödin A, Scott JJ, Vázquez-Campos X, McKay LJ, McDaniel EA, Stevens SLR, Anderson RE, Fuessel J, Fernandez-Guerra A, Maignien L, Delmont TO, Willis AD. 2020. Community-led, integrated, reproducible multi-omics with anvi'o. Nature Microbiology 6:3–6.

34. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, Von Mering C, Bork P. 2019. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res 47:D309–D314.

35. Cantalapiedra CP, Herɲandez-Plaza A, Letunic I, Bork P, Huerta-Cepas J. 2021. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. Mol Biol Evol 38:5825–5829.

36. Zipperer A, Konnerth MC, Laux C, Berscheid A, Janek D, Weidenmaier C, Burian M, Schilling NA, Slavetinsky C, Marschal M, Willmann M, Kalbacher H, Schittek B, Brötz-Oesterhelt H, Grond S, Peschel A, Krismer B. 2016. Human commensals producing a novel antibiotic impair pathogen colonization. Nature 2016 535:7613 535:511–516.

37. Rosenstein R, Salazar BOT, Sauer C, Heilbronner S, Krismer B, Peschel A. 2024. Siderophore piracy enables the nasal commensal Staphylococcus lugdunensis to antagonize the pathogen Staphylococcus aureus. bioRxiv 2024.02.29.582731.

38. Blin K, Shaw S, Augustijn HE, Reitz ZL, Biermann F, Alanjary M, Fetter A, Terlouw BR, Metcalf WW, Helfrich EJN, Van Wezel GP, Medema MH, Weber T. 2023. antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. Nucleic Acids Res 51:W46–W50.

39. Olender A, Niemcewicz M. 2010. Macrolide, lincosamide, and streptogramin B-constitutive-type resistance in Corynebacterium pseudodiphtheriticum isolated from upper respiratory tract specimens. Microb Drug Resist 16:119–122.

40. Ortiz-Pérez A, Martín-De-Hijas NZ, Esteban J, Fernández-Natal MI, García-Cía JI, Fernández-Roblas R. 2010. High frequency of macrolide resistance mechanisms in clinical isolates of Corynebacterium species. Microb Drug Resist 16:273–277.

41. Tauch A, Kassing F, Kalinowski J, Pühler A. 1995. The Corynebacterium xerosis composite transposon Tn5432 consists of two identical insertion sequences, designated IS1249, flanking the erythromycin resistance gene ermCX. Plasmid 34:119–131.

42. Hodgson ALM, Krywult J, Radford A. 1990. Nucleotide sequence of the erythromycin resistance gene from the Corynebacterium plasmid pNG2. Nucleic Acids Res 18:1891.

43. Rosato AE, Lee BS, Nash KA. 2001. Inducible macrolide resistance in Corynebacterium jeikeium. Antimicrob Agents Chemother 45:1982–1989.

44. Camargo AP, Roux S, Schulz F, Babinski M, Xu Y, Hu B, Chain PSG, Nayfach S, Kyrpides NC. 2024. Identification of mobile genetic elements with geNomad. Nature Biotechnology 42: 1303–1312.

45. Bouras G, Nepal R, Houtak G, Psaltis AJ, Wormald PJ, Vreugde S. 2023. Pharokka: a fast scalable bacteriophage annotation tool. Bioinformatics 39: btac776.

46. Katoh K, Misawa K, Kuma K-I, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30: 3059-3066

47. Payne LJ, Todeschini TC, Wu Y, Perry BJ, Ronson CW, Fineran PC, Nobrega FL, Jackson SA. 2021. Identification and classification of antiviral defence systems in bacteria and archaea with PADLOC reveals new system types. Nucleic Acids Res 49:10868–10878.

48. Doron S, Melamed S, Ofir G, Leavitt A, Lopatina A, Keren M, Amitai G, Sorek R. 2018. Systematic discovery of antiphage defense systems in the microbial pangenome. Science 359: eaar4120.

49. McLandsborough LA, Kolaetis KM, Requena T, McKay LL. 1995. Cloning and characterization of the abortive infection genetic determinant abiD isolated from pBF61 of Lactococcus lactis subsp. lactis KR5. Appl Environ Microbiol 61:2023–2026.

50. Cheng R, Huang F, Wu H, Lu X, Yan Y, Yu B, Wang X, Zhu B. 2021. A nucleotide-sensing endonuclease from the Gabija bacterial defense system. Nucleic Acids Res 49:5216–5229.

51. Tal N, Morehouse BR, Millman A, Stokar-Avihail A, Avraham C, Fedorenko T, Yirmiya E, Herbst E, Brandis A, Mehlman T, Oppenheimer-Shaanan Y, Keszei AFA, Shao S, Amitai G, Kranzusch PJ, Sorek R. 2021. Cyclic CMP and cyclic UMP mediate bacterial immunity against phages. Cell 184:5728-5739.e16.

52. Millman A, Melamed S, Leavitt A, Doron S, Bernheim A, Hör J, Garb J, Bechon N, Brandis A, Lopatina A, Ofir G, Hochhauser D, Stokar-Avihail A, Tal N, Sharir S, Voichek M, Erez Z, Ferrer JLM, Dar D, Kacen A, Amitai G, Sorek R. 2022. An

expanded arsenal of immune systems that protect bacteria from phages. Cell Host Microbe 30:1556-1569.e5.

53. Grice EA, Segre JA. 2011. The skin microbiome. Nature Reviews Microbiology 9:244–253.

54. Oh J, Freeman AF, Comparative N, Program S, Park M, Sokolic R, Candotti F, Holland SM, Segre JA, Kong HH. 2013. The altered landscape of the human skin microbiome in patients with primary immunodeficiencies. Genome research 23:2103–2114

55. Findley K, Oh J, Yang J, Conlan S, Deming C, Meyer JA, Schoenfeld D, Nomicos E, Park M, Becker J, Benjamin B, Blakesley R, Bouffard G, Brooks S, Coleman H, Dekhtyar M, Gregory M, Guan X, Gupta J, Han J, Hargrove A, Ho SL, Johnson T, Legaspi R, Lovett S, Maduro Q, Masiello C, Maskeri B, McDowell J, Montemayor C, Mullikin J, Riebow N, Schandler K, Schmidt B, Sison C, Stantripop M, Thomas J, Thomas P, Vemulapalli M, Young A, Kong HH, Segre JA. 2013. Topographic diversity of fungal and bacterial communities in human skin. Nature 2013 498:7454 498:367–370.

56. Kashaf SS, Proctor DM, Deming C, Saary P, Hölzer M, Comparative N, Program S, Taylor ME, Kong HH, Segre JA, Almeida A, Finn RD. 2022. Integrating cultivation and metagenomics for a multi-kingdom view of skin microbiome diversity and functions. Nature Microbiology 7:169–179.

57. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng JF, Darling A, Malfatti S, Swan BK, Gies EA, Dodsworth JA, Hedlund BP, Tsiamis G, Sievert SM, Liu WT, Eisen JA, Hallam SJ, Kyrpides NC, Stepanauskas R, Rubin EM, Hugenholtz P, Woyke T. 2013. Insights into the phylogeny and coding potential of microbial dark matter. Nature 499:431–437.

58. Bernard KA, Pacheco AL, Burdz T, Wiebe D, Bernier AM. 2020. Corynebacterium godavarianum jani et al. 2018 and corynebacterium hadale wei et al. 2018 are both later heterotypic synonyms of corynebacterium gottingense atasayar et al. 2017,

proposal of an emended description of corynebacterium gottingense atasayar et al. 2017. Int J Syst Evol Microbiol 70:3534–3540.

59. Kolmogorov M, Yuan J, Lin Y, Pevzner PA. 2019. Assembly of long, error-prone reads using repeat graphs. Nature Biotechnology 37:540–546.

60. Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30:2068–2069.

61. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL. 2009. BLAST+: Architecture and applications. BMC Bioinformatics 10:1–9.

62. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, Philippon A, Allesoe RL, Rebelo AR, Florensa AF, Fagelhauer L, Chakraborty T, Neumann B, Werner G, Bender JK, Stingl K, Nguyen M, Coppens J, Xavier BB, Malhotra-Kumar S, Westh H, Pinholt M, Anjum MF, Duggett NA, Kempf I, Nykäsenoja S, Olkkola S, Wieczorek K, Amaro A, Clemente L, Mossong J, Losch S, Ragimbeau C, Lund O, Aarestrup FM. 2020. ResFinder 4.0 for predictions of phenotypes from genotypes. J Antimicrob Chemother 75:3491–3500.

63. Richter M, Rosselló-Móra R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. Proc Natl Acad Sci U S A 106:19126–19131.

64. Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. 2015. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. Analytical Methods 8:12–24.

65. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, Koren S, Phillippy AM. 2016. Mash: Fast genome and metagenome distance estimation using MinHash. Genome Biol 17:1–14.

66. Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL. 2004. Versatile and open software for comparing large genomes. Genome Biol 5:1–9.

67. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0. Syst Biol 59:307–321.

68. Anisimova M, Gascuel O. 2006. Approximate Likelihood-Ratio Test for Branches: A Fast, Accurate, and Powerful Alternative. Syst Biol 55:539–552.

69. Letunic I, Bork P. 2013. Interactive Tree of Life (iTOL) v6: recent updates to the phylogenetic tree display and annotation tool. Nucleic Acids Res 1:13–14.

70. Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. 2011. BLAST Ring Image Generator (BRIG): Simple prokaryote genome comparisons. BMC Genomics 12:1–10.

71. Cantalapiedra CP, Herņandez-Plaza A, Letunic I, Bork P, Huerta-Cepas J. 2021. eggNOG-mapper v2: Functional Annotation, Orthology Assignments, and Domain Prediction at the Metagenomic Scale. Mol Biol Evol 38:5825–5829.

72. Huerta-Cepas J, Szklarczyk D, Heller D, Hernández-Plaza A, Forslund SK, Cook H, Mende DR, Letunic I, Rattei T, Jensen LJ, Von Mering C, Bork P. 2019. eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. Nucleic Acids Res 47:D309–D314.

73. Kanehisa M, Sato Y, Kawashima M. 2022. KEGG mapping tools for uncovering hidden features in biological data. Protein Sci 31:47–53.

74. Kanehisa M, Sato Y. 2020. KEGG Mapper for inferring cellular functions from protein sequences. Protein Sci 29:28–35.

75. Payne LJ, Meaden S, Mestre MR, Palmer C, Toro N, Fineran PC, Jackson SA. 2022. PADLOC: a web server for the identification of antiviral defence systems in microbial genomes. Nucleic Acids Res 50:W541–W550.

Chapter 5

The combined roles of staphylococci and corynebacteria in thioalcohol malodour formation

5.1 Manuscript information

Malodour formation can be driven by cooperation between *Staphylococcus* sp. and *Corynebacterium* sp. within the same community

Reyme Herman[1,5], Lauren Cozens[1,2], Robert Cornmell[3], Anthony J. Wilkinson[4], Barry Murphy[3], Gavin H. Thomas[1,5*]

[1] Department of Biology, University of York, Wentworth Way, York, YO10 5DD, UK
[2] School of Biochemistry, Biomedical Sciences Building, University Walk, Bristol, BS81TD, UK
[3] Unilever Research & Development, Port Sunlight Laboratory, Quarry Road East, Bebington, Wirral, Merseyside, CH63 3JW, UK
[4] York Structural Biology Laboratory, Department of Chemistry, University of York, Wentworth Way, York, YO10 5DD, UK
[5] York Biomedical Research Institute, University of York, YO10 5DD, UK

Manuscript intended to be submitted.

Co-authors have given permission for this manuscript to be used in this thesis.

5.2 Context within thesis (Abstract)

The work presented in this thesis culminates in Chapter 5 through the exploration of the different roles members of the axillary microbiome may have in thioalcohol-based malodour generation. The isolates identified in Chapters 3 and 4 were combined to assess the possible roles of all 23 *Staphylococcus* and *Corynebacterium* isolates derived from the axilla of the same high-malodour volunteer. These isolates were assessed for biotransformation capabilities of both the precursor molecule Cys-Gly-3M3SH and the intermediate Cys-3M3SH to find all possible routes to 3M3SH formation between isolates of the two major skin genera.

5.3 Abstract

Human malodour has long been attributed to microbial activity on the skin, driven by the metabolism of sweat gland-derived compounds. The genus *Corynebacterium*, a dominant member of the skin microbiome, was initially implicated in the formation of thioalcohol-based malodour. However, the focus of research later shifted toward a monophyletic (closely related) group of cutaneous *Staphylococcus* species including *S. hominis*. This group can internalise the apocrine gland-derived malodour precursor Cys-Gly-3M3SH before breaking it down to Cys-3M3SH followed by further hydrolysis to produce the odourous compound 3M3SH. In this study, we demonstrate that *Corynebacterium amycolatum* can also biotransform Cys-Gly-3M3SH, potentially contributing to human malodour. Despite the well-documented complexity of the skin microbiome, previous studies of malodour pathways have mainly focused on individual organisms. Here, we have explored malodour formation through cross-species and cross-genus interactions between *Corynebacterium* and *Staphylococcus* isolates collected from the axilla of a high odour individual. We show that both genera can partially hydrolyse Cys-Gly-3M3SH, which is then further metabolised by *C. amycolatum* to complete malodour production in a process we term "tandem biotransformation." Surprisingly, we found that 3M3SH production was significantly enhanced when two *Corynebacterium* isolates worked together in tandem surpassing the levels produced by a single isolate. Together, our findings restore the role of corynebacteria in malodour formation. We also unveil collaborations between microbes within the same niche in an ancient biochemical process previously attributed to single organisms.

5.4 Introduction

Human body odour, or malodour, arises primarily due to microbial activity on the skin where bacteria transform odourless precursor molecules into volatile odourants[1–3]. These precursors are secreted by apocrine glands which are uniquely associated with hair follicles and found in specific body regions such as the underarms and genital area. It is known that several key malodour precursors, including S-(1-hydroxy-3-methylhexan-3-yl)-L-cysteinylglycine (Cys-Gly-3M3SH)[4] are secreted from these glands. Extensive evidence indicates that these precursors are rapidly broken down by the resident skin microbiome[5,6].

The human skin microbiome including the underarm (axilla) microbiota is primarily composed of staphylococci and corynebacteria[7–9]. The conversion of the odourless precursor Cys-Gly-3M3SH into the pungent thiol odourant 3M3SH is a multi-step process involving its uptake, followed by enzymatic breakdown by a dipeptidase and a cysteine-S-conjugate β-lyase (C-S β-lyase)[10–14] (Fig 1). The dipeptidase and C-S β-lyase required for this transformation were initially identified in an axillary *Corynebacterium* isolate[12,13]. Recent studies have focused on some *Staphylococcus* species, revealing their key role in 3M3SH production. Whole-cell assays using the Cys-Gly-3M3SH precursor have implicated a specific clade of axillary staphylococci, particularly *Staphylococcus hominis* in this process[4,11,15]. A complete biochemical pathway for 3M3SH production in *S. hominis* has now been elucidated, involving the transporter DtpT[10], the dipeptidase PepV[16], and the PLP-dependent C-S β-lyase PatB[11]. While DtpT and PepV are commonly found in cutaneous staphylococci like *S. epidermidis*, PatB is uniquely found in the specific staphylococcal clade containing *S. hominis* and related species.

The genus *Corynebacterium* form another dominant group of skin commensals and are present in both high- and low-odour individuals[17]. Several commensal *Corynebacterium* species, including *C. tuberculostearicum* and *C. amycolatum*, have been identified on the skin[6,18]. Axillary isolates of *C. amycolatum* have demonstrated moderate (~40% relative to *S. hominis*) ability to biotransform Cys-Gly-3M3SH[19], yet this species has received little consideration in recent malodour studies. Interestingly, early molecular biology research on human malodour focused on corynebacterial

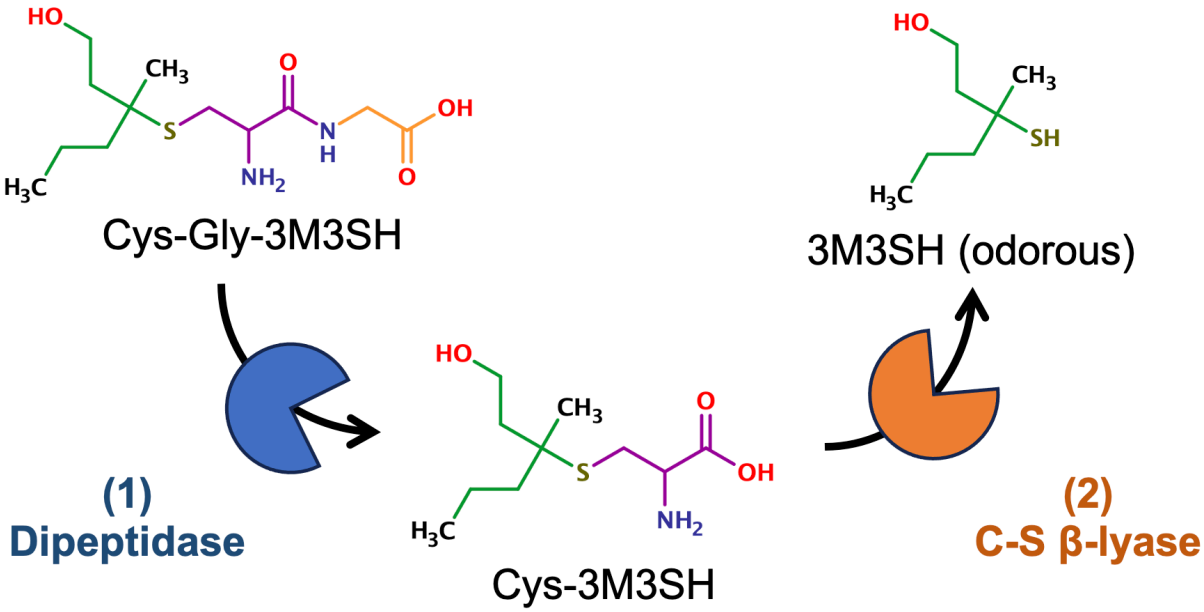


Figure 1. **The odorous molecule 3M3SH is derived from the precursor Cys-Gly-3M3SH.** (1) A dipeptidase first removes the glycine moiety from Cys-Gly-3M3SH to generate the intermediate Cys-3M3SH. (2) A C-S β-lyase then further processes the intermediate to generate the odorous molecule 3M3SH that is a major contributor to thioalcohol-based malodour.

proteins, identifying the dipeptidase TpdA[20] and the C-S β-lyase AecD[21] as enzymes capable of converting Cys-Gly-3M3SH to 3M3SH, although the corresponding uptake pathway remains unidentified.

Here we propose and test an alternative hypothesis for malodour formation involving a division of labour between bacterial species. According to this hypothesis the initial step in the biotransformation process (Cys-Gly-3M3SH to Cys-3M3SH) occurs in organisms such as *Staphylococcus epidermidis* that do not express PatB, while the subsequent conversion of Cys-3M3SH to 3M3SH is carried out by other bacteria including *Corynebacterium* species. We term this process "tandem biotransformation," which describes a complete Cys-Gly-3M3SH biotransformation pathway involving cross-species and even cross-genus interactions. To test this hypothesis, we analysed 23 genetically diverse axillary isolates of *Staphylococcus*[22] and *Corynebacterium*[23] from the York Skin Microbiome (YSM) collection, all derived from the same individual who was assessed to have high malodour. This approach allowed us to assess microbial interactions representative of a true skin microbiome community. Using this diverse isolate library, we evaluated their contributions to thioalcohol-based malodour formation, identifying *C. amycolatum* as a competent Cys-Gly-3M3SH biotransformer and uncovering multiple tandem biotransformation routes to malodour involving both *Staphylococcus* and *Corynebacterium* species.

5.5 Results

5.5.1 Axillary *S. hominis* and *C. amycolatum* isolates are capable of biotransforming both Cys-Gly-3M3SH and Cys-3M3SH

In this study, we explore the roles of genetically diverse axillary isolates of *Staphylococcus* and *Corynebacterium* in thioalcohol-based malodour formation. This subset of YSM isolates is made up of 10 species including the known malodour producer *S. hominis*. To establish these roles, we initially tested the capabilities of 23 isolates to biotransform both synthetic Cys-Gly-3M3SH and Cys-3M3SH (Fig 2, Supplementary Fig 4). Generation of 3M3SH was quantified by measuring the concentration of thiols in the reaction mixture using a standard curve (Supplementary

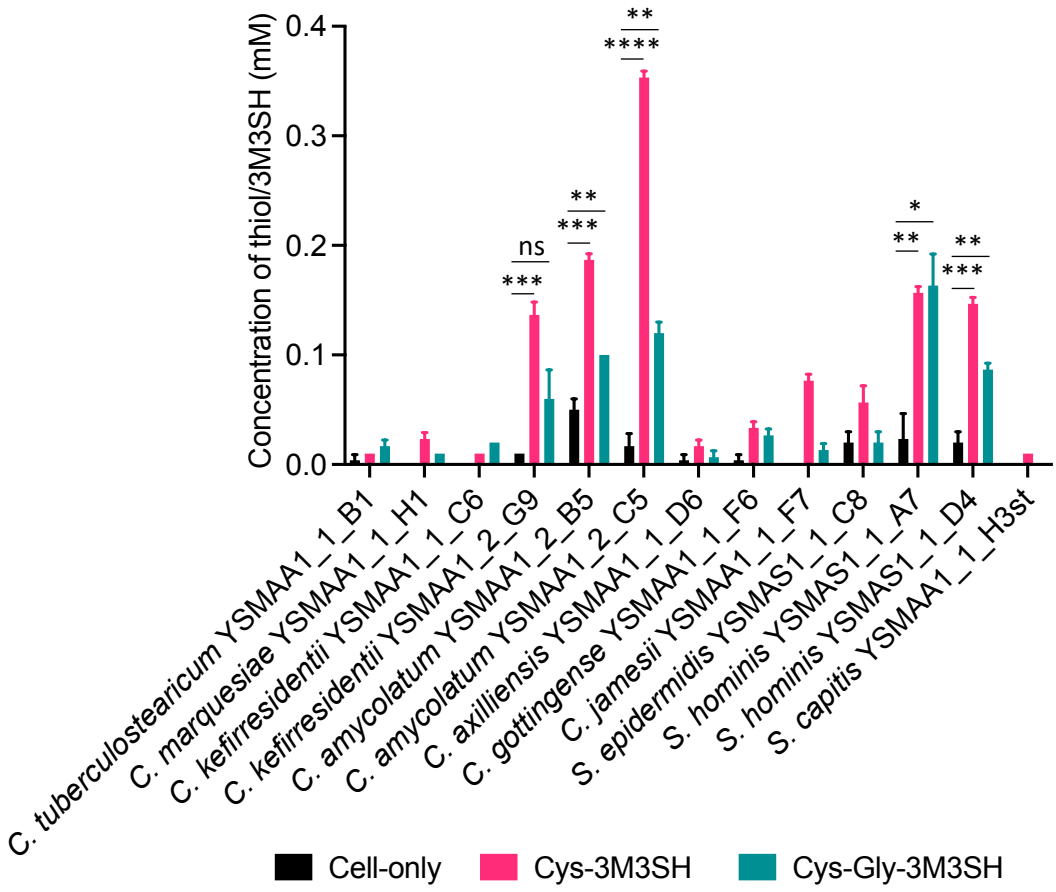


Figure 2. Isolates of *S. hominis* and *C. amycolatum* are able to efficiently biotransform Cys-Gly-3M3SH and Cys-3M3SH. (A) *Staphylococcus* and *Corynebacterium* isolates were incubated with either Cys-Gly-3M3SH, Cys-3M3SH or PBS (Cell-only). *C. amycolatum* YSMAA1_1_C5 is able to biotransform Cys-3M3SH most efficiently. Data shown are the mean (± s.d.) A412 values of biological triplicates measured after development with DTNB. (*P* = >0.05(ns), 0.01(*), 0.001(**), 0.0001(***), <0.0001(****)).

Fig 1-3). Differences in the measured concentrations of thiols between cell-only controls and biotransformation reaction mixtures directly correlate with the generation of 3M3SH.

As expected, we observe significant Cys-Gly-3M3SH biotransformation activity by *S. hominis* isolates (Fig 2). Further, the two tested *C. amycolatum* isolates were also efficient at biotransforming Cys-Gly-3M3SH. We also tested all isolates for Cys-3M3SH activity with a particular focus on *Corynebacterium sp.* isolates as species within this genus are commonly found with both enzymes required for malodour generation (Fig 2, Supplementary Fig 4). We observed minimal activities with the exception of *C. kefirresidentii* YSMAA1_2_G9, *C. amycolatum* YSMAA1_2_B5 and *C. amycolatum* YSMAA1_2_C5 with the latter being the most effective Cys-3M3SH biotransformer. Although we observe moderate Cys-3M3SH biotransformation with *C. kefirresidentii* YSMAA1_2_G9, the other tested isolates of the same species were found to be weak biotransformers of this intermediate. Hence, we focused downstream testing on *C. amycolatum* isolates as strong biotransformation capabilities of both precurors were found in both genetically unique isolates. Unexpectedly, we also observed equally moderate Cys-3M3SH biotransformation activities for the two *S. hominis* isolates. This finding provides another route by which *S. hominis* could contribute to malodour formation. By cataloguing biotransformation activities of this set of 23 genetically diverse isolates from the same individual, we were able to identify isolates that could participate in the full biotransformation reaction (Cys-Gly-3M3SH to 3M3SH) and also the second step in a tandem biotransformation process (Cys-3M3SH to 3M3SH).

5.5.2 Cross-genera cooperation between *Staphylococcus* and *Corynebacterium* in converting Cys-Gly-3M3SH to 3M3SH

We then sought to identify isolates that could perform the first step (Cys-Gly-3M3SH to Cys-3M3SH) of the tandem reaction. In this study we propose roles for non-PatB *Staphylococcus* like *S. epidermidis* to perform this role due to the widespread presence of the transporter DtpT and dipeptidase PepV. We tested this hypothesis against the 7 staphylococci from this collection, including 4 genetically diverse *S. epidermidis*, by first incubating the cells with Cys-Gly-3M3SH then incubating the

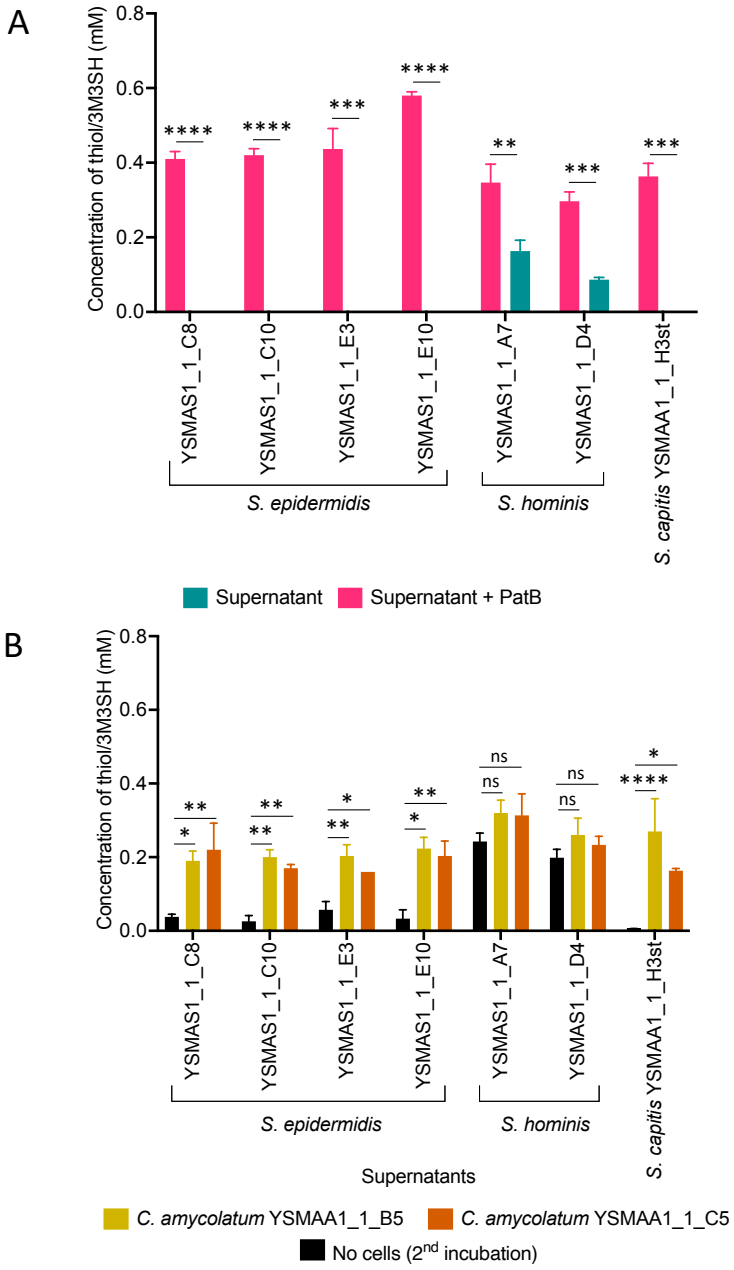


Figure 3. Cross-genus tandem biotransformation of Cys-Gly-3M3SH. (A) All *Staphylococcus* isolates tested released Cys-3M3SH, as measured by the amount of 3M3SH generated following treatment with recombinant PatB (magenta) or without PatB addition (teal). (B) Supernatants from *Staphylococcus* isolates lacking PatB incubated with Cys-Gly-3M3SH were further converted to 3M3SH by *Corynebacterium amycolatum*. Isolates YSMAA1_1_B5 (gold) and YSMAA1_1_C5 (bronze) successfully biotransformed Cys-3M3SH from staphylococcal supernatants. Controls lacking a second incubation step (black) are shown for comparison. Data represent mean (±s.d.) of thiol concentrations in supernatants from biological triplicates. (*P* = >0.05(ns), 0.01(*), 0.001(**), 0.0001(***), <0.0001(****)).

supernatant with recombinant *S. hominis* PatB[24] which is highly specific for Cys-3M3SH. The exogenously added *S. hominis* PatB then liberates 3M3SH to allow for indirect quantification of Cys-3M3SH using DTNB and the standard curve (Fig 3A, Supplementary Fig 1-2).

Using this approach, we found that isolates of PatB-negative species *S. epidermidis* and *S. capitis* were capable of generating and releasing Cys-3M3SH into the supernatant (Fig 3A). This finding supports our hypothesis that these species of staphylococci can partially hydrolyse Cys-Gly-3M3SH resulting in freed Cys-3M3SH.

With this assay, we are also able to determine that *S. hominis* alone is able to generate 0.1 – 0.2 mM 3M3SH directly from Cys-Gly-3M3SH, comparable with the findings of Bawdon *et al*[19]. Our data also unexpectedly revealed production and release of Cys-3M3SH by *S. hominis* suggesting that not all of the Cys-Gly-3M3SH was converted to 3M3SH intracellularly. Rather, some intermediate Cys-3M3SH may be expelled into the supernatant. Together, data in Fig 3A suggest a potential role of all tested *Staphylococcus*, including *S. hominis*, in the proposed tandem biotransformation reaction.

We then addressed our overarching aim of testing the proposed tandem biotransformation reaction by evaluating whether *Staphylococcu*s and *Corynebacterium* isolates could cooperatively convert Cys-Gly-3M3SH to 3M3SH. For this purpose, we resuspended cells of *C. amycolatum* YSMAA1_1_B5 and *C. amycolatum* YSMAA1_1_C5 with the supernatants generated from the incubation of all staphylococci isolates with Cys-Gly-3M3SH and further incubated the suspension before deriving the concentration of free thiols (Fig 3B).

Based on the data in Fig 2, incubation of *C. amycolatum* YSMAA1_1_B5 and YSMAA1_1_C5 with Cys-Gly-3M3SH alone yielded 3M3SH concentrations of 0.100 ± 0.003 mM and 0.119 ± 0.005 mM respectively. These values serve as baseline levels for 3M3SH production via direct biotransformation of the precursor. In the tandem biotransformation assay, any additional thiol production beyond these levels would indicate hydrolysis of Cys-3M3SH present in the supernatants. Taking the aforementioned baseline levels into account, data in Fig 3B suggest that ~0.1 mM to

~0.2 mM 3M3SH was generated by both *C. amycolatum* isolates from freed Cys-3M3SH in the staphylococcal supernatants. However, 3M3SH generation in the tandem biotransformation reactions was less than expected as data in Fig 2A suggest that Cys-3M3SH produced by PatB-deficient staphylococci ranges from 0.363 ± 0.035 mM for *S. capitis* YSMAA1_1_H3st to 0.580 ± 0.010 mM for *S. epidermidis* YSMAS1_1_E10. As we do not observe complete conversion, we hypothesise that some metabolites found in the staphylococcal supernatants may be inhibiting the function of the tested corynebacteria. Both isolates of *S. hominis* were able to generate Cys-3M3SH (Fig 3A) but we did not observe significant differences ($P > 0.05$) when *S. hominis* supernatants were incubated with either of the two *C. amycolatum* isolates. Nevertheless, data presented in Fig 3 collectively support the proposed cross-genus tandem biotransformation pathway, particularly involving *Staphylococcus* species lacking PatB.

5.5.3 Pairs of corynebacteria are able to work in tandem to completely biotransform Cys-Gly-3M3SH

Given the absence of a known transporter for Cys-Gly-3M3SH, presence of both the peptidase and lyase and the lack of biotransformation activity with Cys-Gly-3M3SH (Fig 2), it was considered unlikely that *Corynebacterium* species other than *C. amycolatum* can convert Cys-Gly-3M3SH to Cys-3M3SH. Nonetheless, to rigorously evaluate this hypothesis, we tested the ability of all *Corynebacterium* isolates to generate Cys-3M3SH from Cys-Gly-3M3SH (Fig 4A, Supplementary Fig 5 and 6A).

Surprisingly, all isolates of corynebacteria were able to convert Cys-Gly-3M3SH to Cys-3M3SH (Fig 4A), in some cases more efficiently than staphylococci isolates (Fig 3A). While no consistent species-specific trends were observed, the levels of Cys-3M3SH produced ranged between 0.161 ± 0.040 mM for *C. amycolatum* YSMAA1_2_C5 and 1.053 ± 0.304 mM for *C. tuberculostearicum* YSMAA1_1_D8. Unlike *Staphylococcus* species lacking PatB, the underlying mechanism of this biotransformation remains unclear for *Corynebacterium*. Due to the lack of known corynebacterial Cys-Gly-3M3SH transport activity, we suggest the presence of a ubiquitous corynebacterial surface or secreted peptidase that is able to hydrolyse the

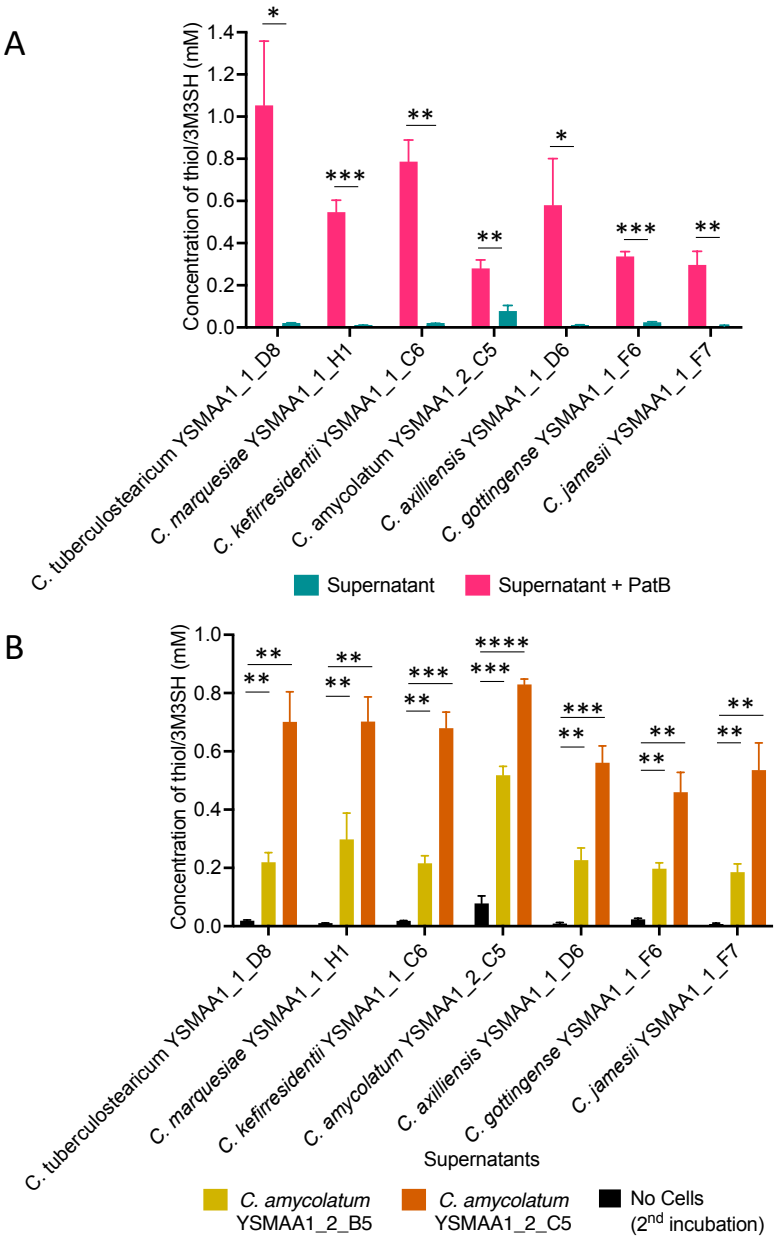


Figure 4. Tandem biotransformation of Cys-Gly-3M3SH by pairs of *Corynebacterium* isolates. (A) *Corynebacterium* isolates tested released Cys-3M3SH, as measured by the amount of 3M3SH generated following treatment with recombinant PatB (magenta) or without PatB addition (teal). (B) *C. amycolatum* YSMAA1_1_C5 efficiently converts Cys-3M3SH precursors present in supernatants from other *Corynebacterium* isolates to 3M3SH. Supernatants from overnight *Corynebacterium* incubations with Cys-Gly-3M3SH were mixed with *C. amycolatum* YSMAA1_1_B5 (gold) or YSMAA1_1_C5 (bronze). Controls lacking a second incubation step (black bars) are shown for comparison. Data represent mean (± s.d.) of thiol concentrations in supernatants from biological triplicates. (P = >0.05(ns), 0.01(*), 0.001(**), 0.0001(***), <0.0001(****)).

glycine from Cys-Gly-3M3SH. Regardless of the mode of action, we have now demonstrated another source of Cys-3M3SH which could be further broken down by organisms capable of performing the second step of the tandem biotransformation reaction. To assess potential *Corynebacterium-Corynebacterium* tandem biotransformation, we incubated supernatants from overnight Cys-Gly-3M3SH-treated cultures of *Corynebacterium* isolates with *C. amycolatum* YSMAA1_1_B5 and YSMAA1_1_C5. While *C. amycolatum* YSMAA1_1_B5 was not more efficient at completing the tandem biotransformation process than when incubated with staphylococcal supernatants, *C. amycolatum* YSMAA1_1_C5 was found to be up to 4-fold more effective in the same process with corynebacteria supernatants (Fig 4B, Supplementary Fig 5 and 6B). These tandem biotransformation processes were also ~4-fold more effective than with *S. hominis* alone. Interestingly, *C. amycolatum* YSMAA1_1_C5 performing the full process on its own only yielded 0.119 ± 0.005 mM of 3M3SH which if done in tandem, yielded 0.829 ± 0.019 mM of 3M3SH. Similarly, *C. amycolatum* YSMAA1_1_B5 is able to yield 0.100 ± 0.003 mM 3M3SH directly from Cys-Gly-3M3SH in a single incubation step but if done in tandem with itself, yielded 0.389 ± 0.034 mM of 3M3SH.

5.6 Discussion

The formation of thioalcohol-based human malodour has traditionally been associated with single organisms like *S. hominis*. However, the skin microbiome is a complex community of microbes spanning multiple genera and species. It is therefore plausible that multiple species contribute to malodour formation in some way. Our YSM library of axilla isolates from the same individual provides a platform to perform cross-species and cross-genus testing of different routes to malodour formation. Using this resource, we not only characterised more species of skin microbes capable of biotransforming the malodour precursor Cys-Gly-3M3SH but also proposed an alternative route to malodour formation using a phenomenon we term “tandem biotransformation”.

To assess the role of individual species in malodour formation, we first evaluated the ability of *C. amycolatum* isolates from a high malodour individual to biotransform the thioalcohol precursor Cys-Gly-3M3SH. These isolates produced 3M3SH at levels approximately 50% of that observed with *S. hominis*, indicating partial but significant

activity ($P < 0.05$). This result is consistent with previous findings by Bawdon *et. al.*[19] who reported comparable yields for *C. amycolatum* under similar assay conditions (Fig 3A, Fig 4A). While *S. hominis* remains the most efficient single-species converter of Cys-Gly-3M3SH, the observed activity in *C. amycolatum* suggests it may contribute to malodour formation via a single-organism biotransformation route, particularly in individuals with high relative abundance of this species.

The complete conversion of Cys-Gly-3M3SH to the volatile thiol 3M3SH is thought to proceed via two enzymatic steps: (1) cleavage of Cys-Gly-3M3SH to Cys-3M3SH by a dipeptidase, followed by (2) cleavage of Cys-3M3SH to 3M3SH by a C-S β-lyase (Fig 1). While previous models assumed both steps occurred within a single organism, our data support alternative community-driven mechanisms. Fig 2 describes data supporting one such mechanism where all tested *Staphylococcus* isolates performed step (1), and *C. amycolatum* carrying out step (2) in a cross-genus tandem biotransformation reaction.

Unexpectedly, we observed that all tested *Corynebacterium* isolates were also capable performing step (1) despite no prior evidence of Cys-Gly-3M3SH transport. Hence, we propose that this step could be occurring in the extracellular environment through in an as-yet unknown mechanism. When supernatants from *Corynebacterium*-treated Cys-Gly-3M3SH were subsequently incubated with *C. amycolatum* YSMA1_1_C5, the generated 3M3SH concentration was at least 3-fold higher than for *S. hominis* alone. This phenomenon is also observed when *C. amycolatum* YSMA1_1_C5 was used for both steps of the tandem biotransformation reaction, further challenging our current understanding of malodour formation. Together, data from Fig 3 and Fig 4 not only support the proposed tandem biotransformation pathway but also reiterate the potential role of *Corynebacterium* in human malodour formation.

Multiple studies found a positive correlation between high odour individuals and the proportion of axillary corynebacteria[9,25] although until now, the basis for that correlation was unclear. Could the corynebacteria-corynebacteria tandem biotransformation of Cys-Gly-3M3SH provide an explanation for this? The corynebacteria involved in this process could stand to benefit from nutrients produced

as by-products of the pathway towards malodour formation and hence allowing proliferation of this genus. It is not surprising that the routes to malodour formation may be more complex than single-organism pathways. These members of the skin microbiome have likely evolved alongside humans and their ancestors for millions of years and are continuously exposed to skin metabolites (e.g. Cys-Gly-3M3SH). In nutrient limiting environments like the skin, it stands to reason that these organisms will find ways to utilise any metabolite they can access.

To the best of our knowledge, this and all other studies to date testing malodour formation by bacterial isolates are conducted using *in vitro* methods that do not mimic the axillary environment. For example, the physiologically relevant concentration range of Cys-Gly-3M3SH or Cys-3M3SH in the underarm is still unknown and hence, the concentrations used *in vitro* assays may be inaccurate. Further, the distribution of commensal axillary species within different individuals could vary greatly. Although there are some positive correlations between malodour and relative abundance of axillary *Corynebacterium*, species we identified as potential major malodour contributors (i.e. *C. amycolatum*) may not always be in a high enough abundance to recapitulate our findings in actual axillary communities. Alternatively, the abundance of *S. hominis* could be high enough to dominate Cys-Gly-3M3SH biotransformation routes in some individuals, removing the need for a tandem biotransformation process altogether. Nevertheless, through a community-based approach, we identified multiple routes to malodour formation within this community of axillary microbes. While it is difficult to ascertain the true nature of diverse axillary microbiomes, these alternative pathways could play important roles in malodour formation in at least some individuals.

Together, our experiments support the presence of cellular machinery involved in malodour formation across both *Staphylococcus* and *Corynebacterium* including the capacity for tandem biotransformation of Cys-Gly-3M3SH. These findings open a clear path for future studies aimed at elucidating the molecular basis of this cooperative process, particularly the transport of Cys-3M3SH and the enzymatic mechanisms by which corynebacteria cleave Cys-Gly-3M3SH. By adopting a community-based perspective, our work offers new insight into how a diverse population of axillary bacteria may interact to drive this evolutionarily persistent human trait.

5.7 Materials and methods

5.7.1 Bacterial strains and growth conditions

Axillary isolates of *Staphylococcus* and *Corynebacterium* tested in this study (Supplementary Table 1) were derived from the same individual and previously reported in two studies[22,23]. In this study, isolates were grown aerobically in brain-heart infusion media supplemented with 0.2% Tween-80 at 37°C. Overnight cultures were prepared in triplicates in 96-well deep well plates for each biotransformation assay performed.

5.7.2 Generation of the 3M3SH standard curve

Different concentrations of synthetic Cys-3M3SH (Concept Life Sciences) were mixed with 1μM (excess) recombinantly expressed PatB[24] and the DTNB developing reagent (0.2mM DTNB (5,5'-dithio-bis-(2-nitrobenzoic acid)), 200mM Tris-HCl pH 8). The absorbance at 412nm (A_{412}) of the reaction mixes was measured in cuvettes (1 cm pathlengths). Every thiol interacts with 1 molecule of DTNB, releasing 1 molecule of TNB (Supplementary Fig 1). The concentration of thiols/3M3SH in each mixture were derived by converting A_{412} values to concentrations of released TNB using the extinction coefficient (14100 M^{-1} cm^{-1}). The same mixes were then transferred to Nunc 2.0 96-well plates in the same volume (100uL) as the development of whole cell assays performed in this body of work before measurement of A_{412} values for each mix in the Epoch2 plate reader. The resulting values were then plotted against the derived concentrations of thiols/3M3SH to generate the standard curve (Supplementary Fig 2). Data were analysed and plotted on GraphPad Prism10.

5.7.3 Single-organism biotransformation assay

Overnight cultures were washed with PBS and normalised to an OD_{600} of 1. Cells were either incubated with 10 mM Cys-Gly-3M3SH (Concept Life Sciences) or 0.05 - 10 mM Cys-3M3SH in 0.5 mL tubes, shaking at 100rpm for 24 hours at 37°C. Cells were then pelleted by centrifugation and 5 μL of the supernatants were developed in Nunc 2.0

96-well plates with the DTNB developing reagent (0.2 mM DTNB, 200 mM Tris-HCl pH 8) and where appropriate, excess exogenous PatB[24] was also added. A_{412} values were then measured in an Epoch2 plate reader. A_{412} measurements were converted to concentrations of thiol/3M3SH using the standard curve. The full workflow is described in Supplementary Fig 3. Data were analysed and plotted on GraphPad Prism10.

5.7.4 Tandem biotransformation assay

Overnight cultures of the strains used for the first step of the tandem assay were washed with PBS and normalised to an OD_{600} of 1. Cells were incubated with 10 mM Cys-Gly-3M3SH in 0.5 mL tubes, shaking at 100rpm for 24 hours at 37°C. Cells were then pelleted by centrifugation and the supernatants were retained. 5 μL of the supernatants were developed in Nunc 2.0 96-well plates with the DTNB developing reagent (0.2 mM DTNB, 200 mM Tris-HCl pH 8) and for the measurement of Cys-3M3SH, excess exogenous PatB[24] was also added. A_{412} values were then measured in an Epoch2 plate reader. A_{412} measurements were converted to concentrations of thiol/3M3SH using the standard curve. The full workflow is described in Supplementary Fig 5. Data were analysed and plotted on GraphPad Prism10.

The remaining volume of supernatants were used to resuspend washed and normalised (OD_{600} = 1) overnight cell pellets of the strains tested for the second step of the tandem assay. Mixtures were shaken at 100rpm for 24 hours at 37°C. As a control, aliquots of the supernatants were also incubated in the absence of bacteria cells for the second biotransformation step. The following day, reaction mixtures were then pelleted and 5 μL of the supernatants were developed in Nunc 2.0 96-well plates with the DTNB developing reagent (0.2 mM DTNB, 200 mM Tris-HCl pH 8). A_{412} values were then measured in an Epoch2 plate reader. A_{412} measurements were converted to concentrations of thiol/3M3SH using the standard curve. Data was analysed and plotted on GraphPad Prism10.

5.8 Ethics statement

This study was conducted in accordance with the ethical principles of Good Clinical Practice and the Declaration of Helsinki. The local ethics committees of Unilever R&D (Port Sunlight, UK) approved the protocols before commencement of the studies and all subjects gave written informed consent.

5.9 Conflict of interest

The authors declare that R. C. and B. M. are current employees of Unilever. All other authors declare that they have no conflicts of interest with the contents of this article.

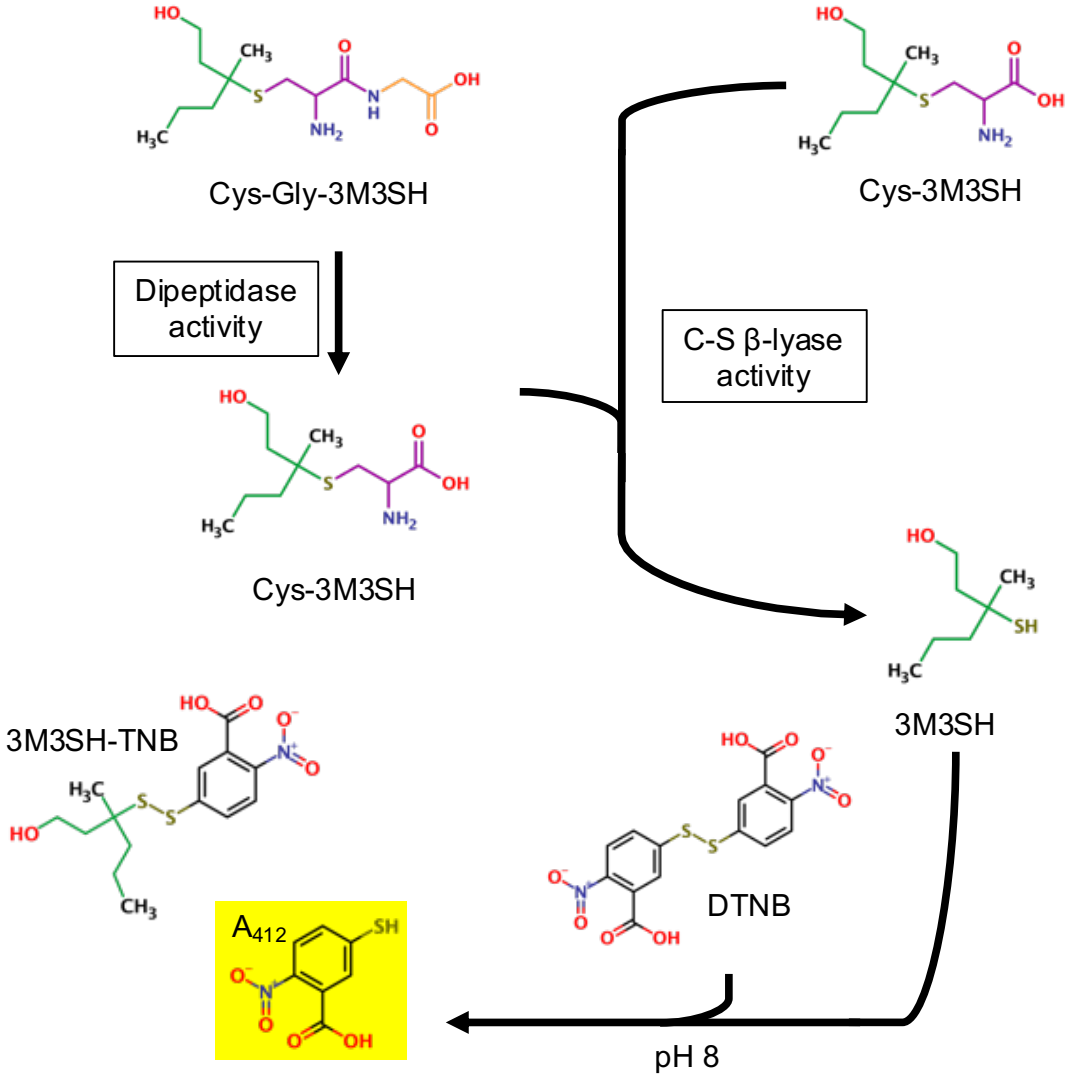
5.10 Funding

This work was supported by the Biotechnology and Biological Sciences Research Council grant BB/W510531/1 supporting R. H.

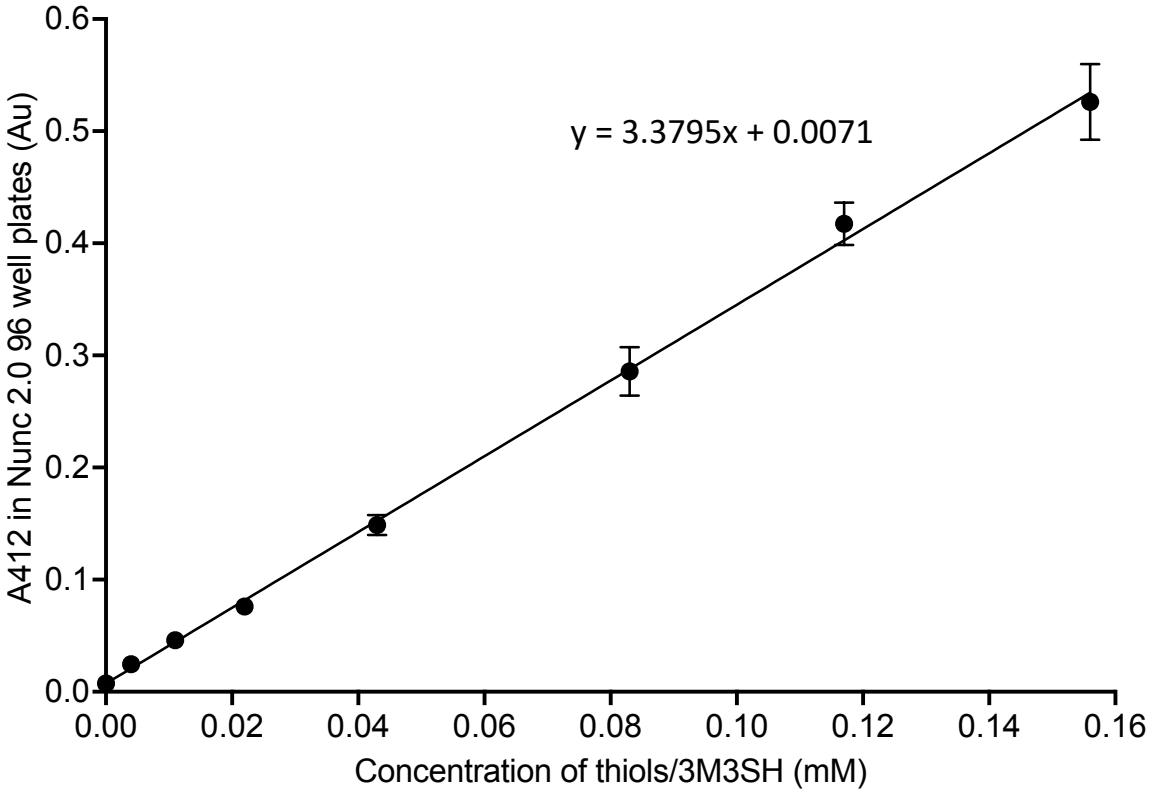
5.11 Author contributions

R. H., B. M. and G. H. T. conceptualization; R. H., and L. C. validation; R. H., L. C., and G. H. T. formal analysis; R. H. and L. C. investigation; R. H. and L. C. data curation; R. H. writing–original draft; R. H., L. C., R. C., B. M., A. J. W. and G. H. T. writing–review & editing; R. H. visualization; A. J. W., B. M., and G. H. T. supervision; R. H., B. M., and G. H. T. project administration; B. M. and G. H. T. funding acquisition.

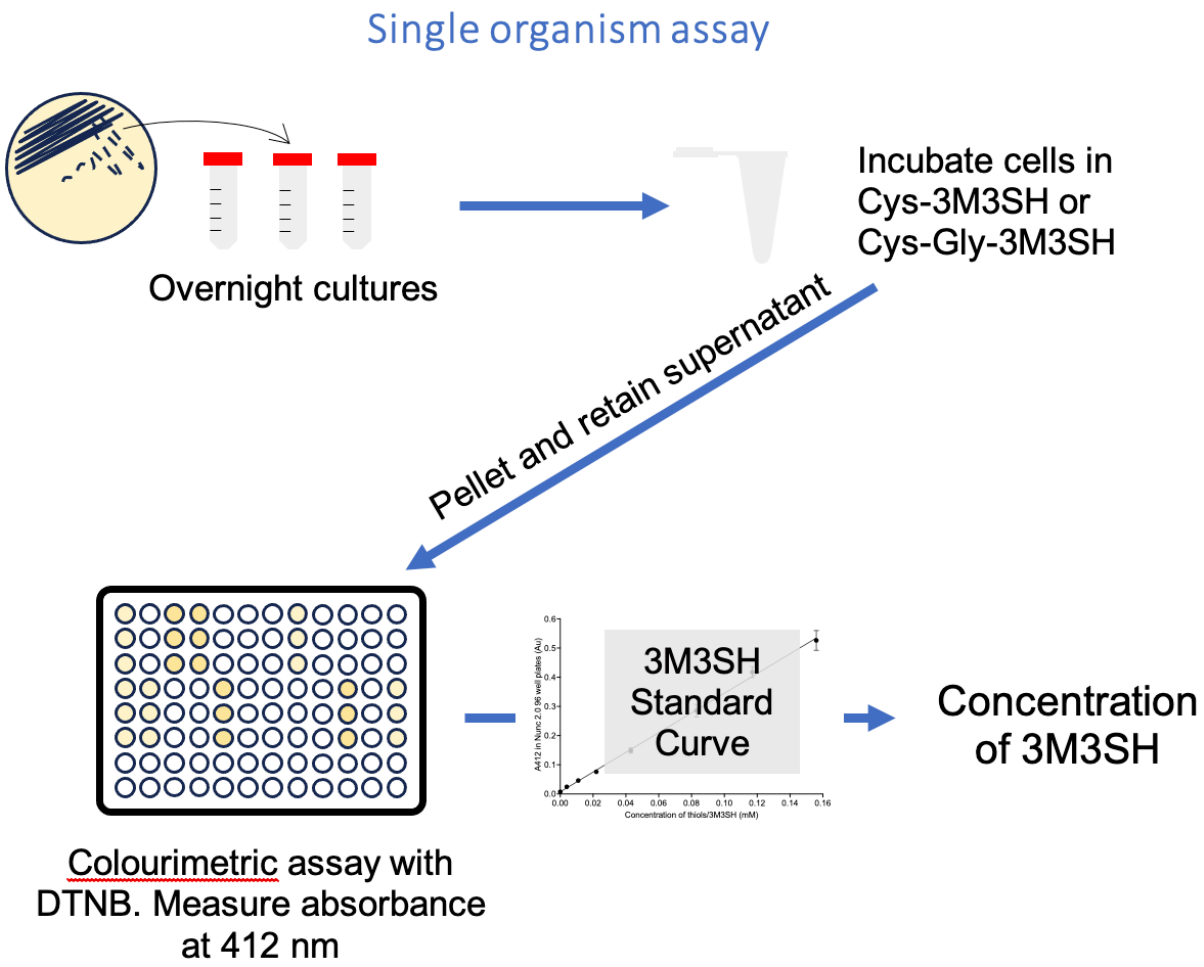
5.12 Supporting data



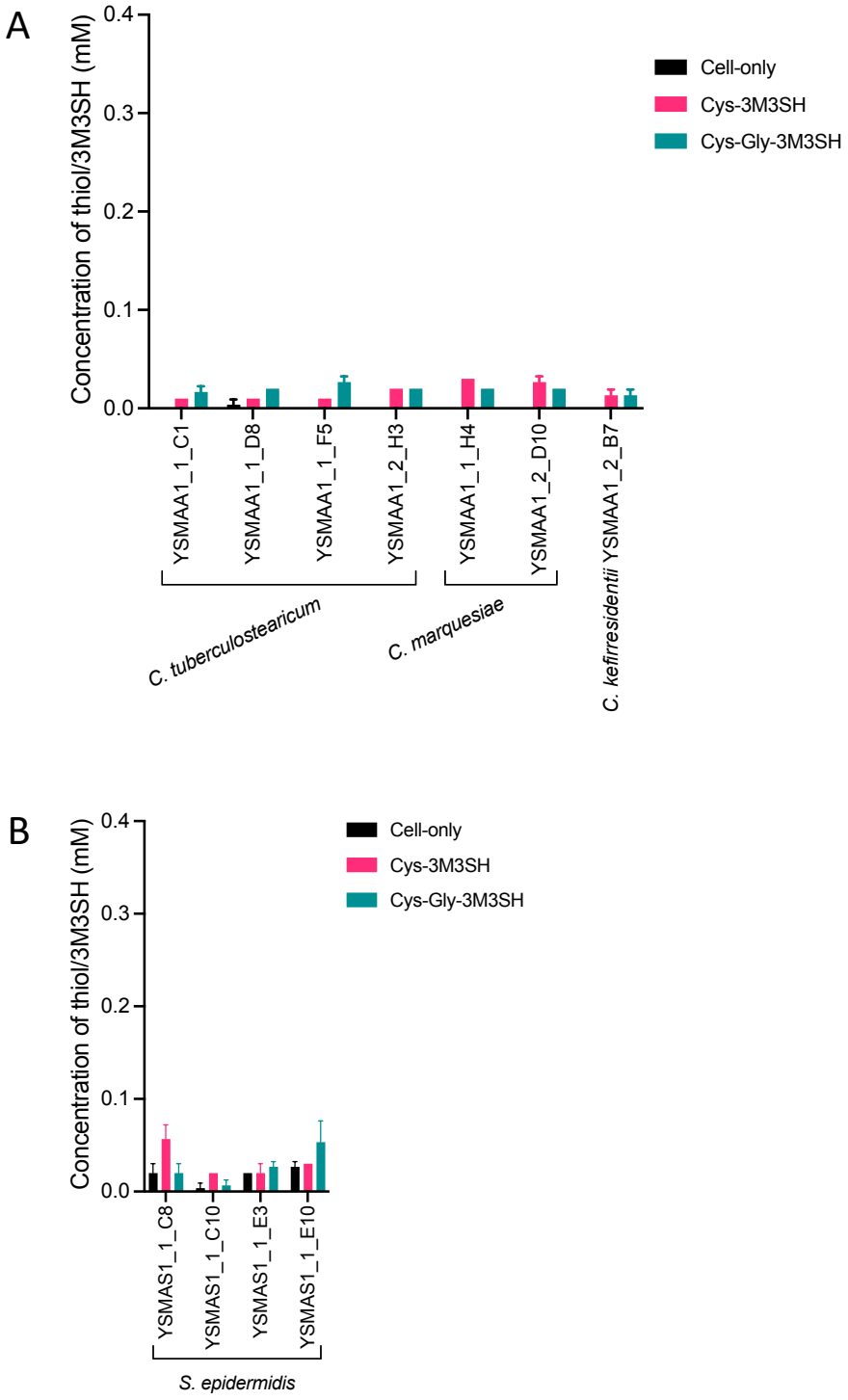
Supplementary Figure 1. The DTNB-based Cys-Gly-3M3SH/Cys-3M3SH biotransformation assay. Cys-Gly-3M3SH and Cys-3M3SH were incubated with various strains of *Corynebacterium* or *Staphylococcus* to biotransform into 3M3SH through dipeptidase and C-S β-lyase activities. The resulting 3M3SH in turn interacts with labelling molecule 5,5′-dithiobis-(2-nitrobenzoic acid) (DTNB) at pH 8. This results in the formation of TNB which absorbs at 412 nm (visible yellow colour). In some cases, recombinant PatB was added to process any resultant Cys-3M3SH to form 3M3SH as a means to quantify the dipeptidase activity.



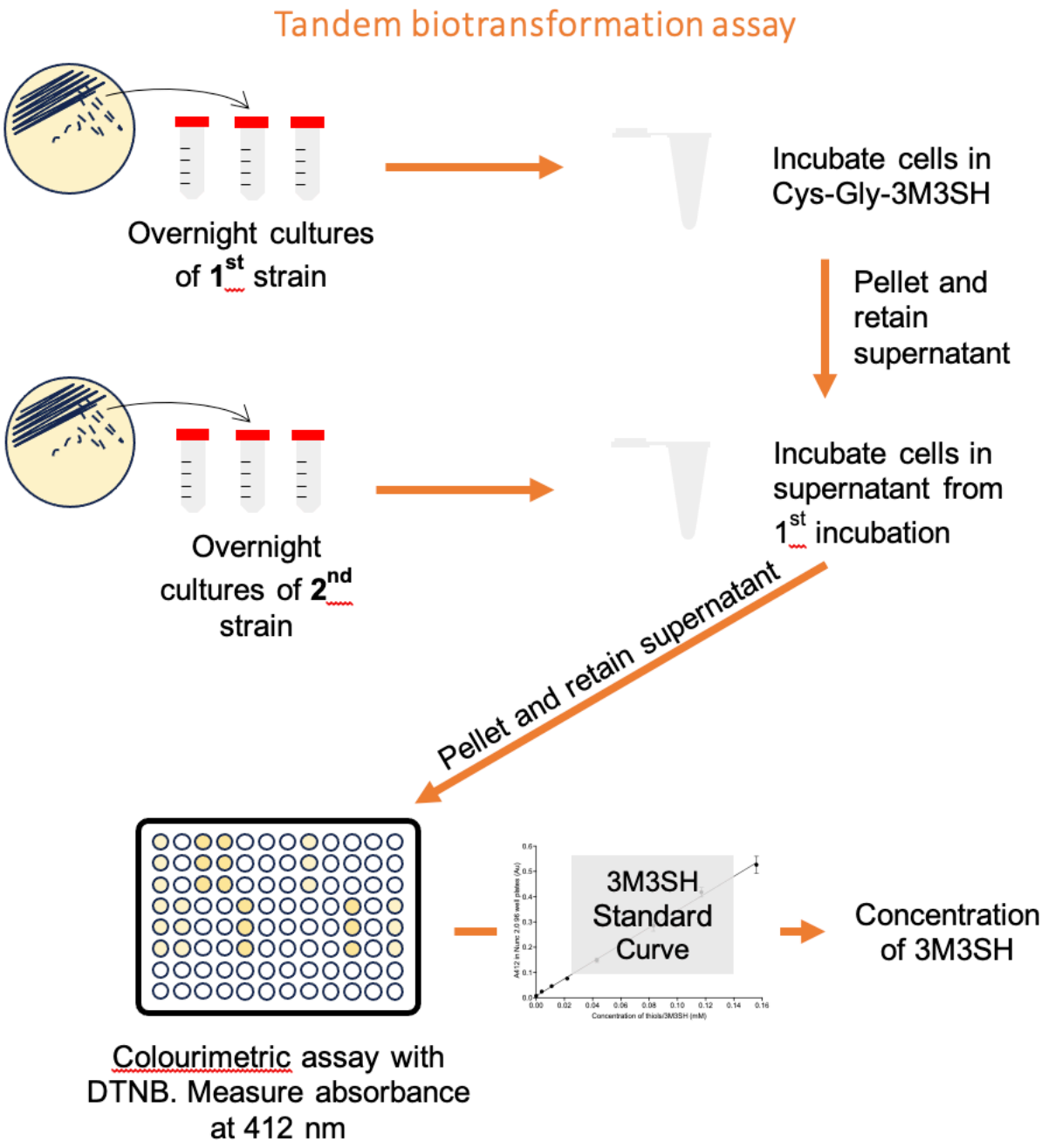
Supplementary Figure 2. Thiols/3M3SH standard curve. A standard curve was generated to allow for the direct quantification of thiols, including 3M3SH, present in reaction mixtures using A412 nm values measured in Nunc 2.0 96-well plates using the Epoch2 microplate spectrophotometer.



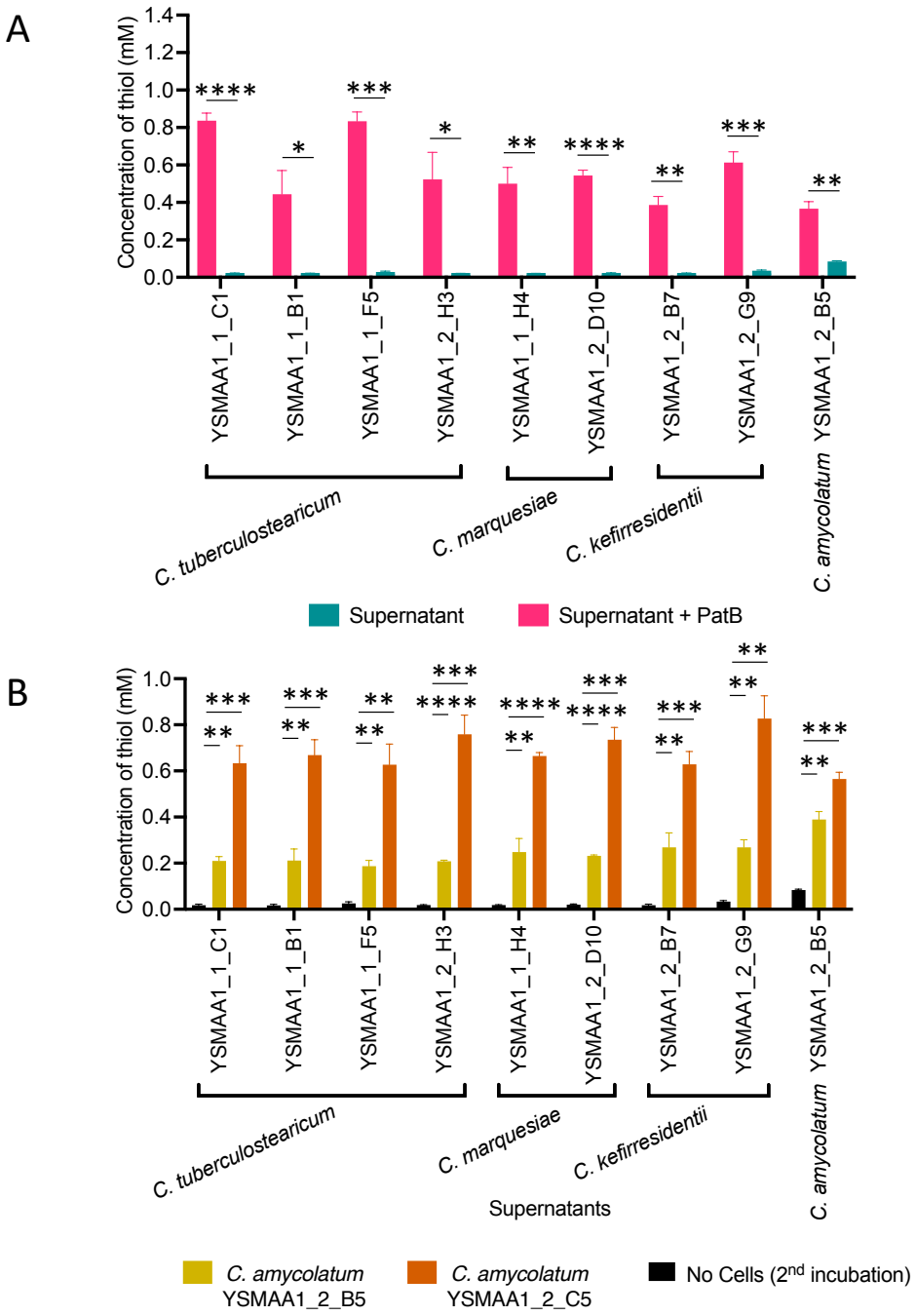
Supplementary Figure 3. The experimental workflow to assess the biotransformation capabilities of a single organism in malodour generation.



Supplementary Figure 4. Isolates of common cutaneous species were not able to biotransform Cys-Gly-3M3SH or Cys-3M3SH. (A) *Corynebacterium* and (B) *Staphylococcus* isolates were incubated with either Cys-Gly-3M3SH, Cys-3M3SH or PBS (Cell-only). Data shown are the mean (± s.d.) A412 values of biological triplicates measured after development with DTNB. (*P* = >0.05(ns), 0.01(*), 0.001(**), 0.0001(***), <0.0001(****)).



Supplementary Figure 5. The experimental workflow to assess the tandem biotransformation capabilities of pairs of strains in malodour generation.



Supplementary Figure 6. Tandem biotransformation of Cys-Gly-3M3SH by pairs of *Corynebacterium* isolates. (A) 9 *Corynebacterium* isolates tested released Cys-3M3SH, as measured by the amount of 3M3SH generated following treatment with recombinant PatB (magenta) or without PatB addition (teal). (B) Supernatants from overnight *Corynebacterium* incubations with Cys-Gly-3M3SH were mixed with *C. amycolatum* YSMAA1_1_B5 (gold) or YSMAA1_1_C5 (bronze). Controls lacking a second incubation step (black bars) are shown for comparison. Data represent mean (± s.d.) of thiol concentrations in supernatants from biological triplicates. (*P* = >0.05(ns), 0.01(*), 0.001(**), 0.0001(***), <0.0001(****)).

Supplementary Table 1. List of isolates used in this study.

Species	Strain
Corynebacterium tuberculostearicum	YSMAA1_1_B1
Corynebacterium tuberculostearicum	YSMAA1_1_C1
Corynebacterium tuberculostearicum	YSMAA1_1_D8
Corynebacterium tuberculostearicum	YSMAA1_1_F5
Corynebacterium tuberculostearicum	YSMAA1_2_H3
Corynebacterium marquesiae	YSMAA1_1_H1
Corynebacterium marquesiae	YSMAA1_1_H4
Corynebacterium marquesiae	YSMAA1_2_D10
Corynebacterium kefirresidentii	YSMAA1_1_C6
Corynebacterium kefirresidentii	YSMAA1_2_B7
Corynebacterium kefirresidentii	YSMAA1_2_G9
Corynebacterium amycolatum	YSMAA1_2_B5
Corynebacterium amycolatum	YSMAA1_2_C5
Corynebacterium axilliensis	YSMAA1_1_D6
Corynebacterium gottingense	YSMAA1_1_F6
Corynebacterium jamesii	YSMAA1_1_F7
Staphylococcus epidermidis	YSMAS1_1_C8
Staphylococcus epidermidis	YSMAS1_1_C10
Staphylococcus epidermidis	YSMAS1_1_E3
Staphylococcus epidermidis	YSMAS1_1_E10
Staphylococcus hominis	YSMAS1_1_A7
Staphylococcus hominis	YSMAS1_1_D4
Staphylococcus capitis	YSMAA1_1_H3st

5.13 References

1. James, A. G. The Axillary Microbiome and its Relationship with Underarm Odor. *Skin Microbiome Handbook: From Basic Research to Product Development* 79–129 (2020)
2. Shelley, W. B., Hurley, H. J. & Nichols, A. C. Axillary odor; experimental study of the role of bacteria, apocrine sweat, and deodorants. *AMA Arch Derm Syphilol* **68**, 430–46 (1953).
3. Natsch, A. & Emter, R. The specific biochemistry of human axilla odour formation viewed in an evolutionary context. *Philosophical Transactions of the Royal Society B* **375**, 20190269 (2020).
4. Troccaz, M. *et al.* 3-Methyl-3-sulfanylhexan-1-ol as a Major Descriptor for the Human Axilla-Sweat Odour Profile. *Chem Biodivers* **1**, 1022–1035 (2004).
5. Fredrich, E., Barzantny, H., Brune, I. & Tauch, A. Daily battle against body odor: towards the activity of the axillary microbiota. *Trends Microbiol* **21**, 305–312 (2013).
6. James, A. G. *et al.* Microbiological and biochemical origins of human axillary odour. *FEMS Microbiol Ecol* **83**, 527–40 (2013).
7. Grice, E. A. *et al.* Topographical and temporal diversity of the human skin microbiome. *Science (1979)* **324**, 1190–1192 (2009).
8. Grice, E. A. & Segre, J. A. The skin microbiome. *Nature Reviews Microbiology 2011 9:4* **9**, 244–253 (2011).
9. Taylor, D. *et al.* Characterization of the microflora of the human axilla. *Int J Cosmet Sci* **25**, 137–145 (2003).
10. Minhas, G. S. *et al.* Structural basis of malodour precursor transport in the human axilla. *Elife* **7**, e34995 (2018).
11. Rudden, M. *et al.* The molecular basis of thioalcohol production in human body odour. *Sci Rep* **10**, 12500 (2020).
12. Emter, R. & Natsch, A. The sequential action of a dipeptidase and a beta-lyase is required for the release of the human body odorant 3-methyl-3-sulfanylhexan-1-ol from a secreted Cys-Gly-(S) conjugate by Corynebacteria. *J Biol Chem* **283**, 20645–52 (2008).
13. Natsch, A., Schmid, J. & Flachsmann, F. Identification of Odoriferous Sulfanylalkanols in Human Axilla Secretions and Their Formation through Cleavage of Cysteine Precursors by a C□S Lyase Isolated from Axilla bacteria. *Chem Biodivers* **1**, 1058–1072 (2004).

14. Natsch, A., Gfeller, H., Gygax, P. & Schmid, J. Isolation of a bacterial enzyme releasing axillary malodor and its use as a screening target for novel deodorant formulations1. *Int J Cosmet Sci* **27**, 115–122 (2005).
15. Bawdon, D., Cox, D. S., Ashford, D., James, A. G. & Thomas, G. H. Identification of axillary Staphylococcus sp. involved in the production of the malodorous thioalcohol 3-methyl-3-sufanylhexan-1-ol. *FEMS Microbiol Lett* **362**, 1–10 (2015).
16. Herman, R. *et al.* Identification of a Staphylococcal dipeptidase involved in the production of human body odour. *Journal of Biological Chemistry* **0**, 107928 (2024).
17. Troccaz, M. *et al.* Mapping axillary microbiota responsible for body odours using a culture-independent approach. *Microbiome* **3**, 1–15 (2015).
18. Taylor, D. *et al.* Characterization of the microflora of the human axilla. *Int J Cosmet Sci* **25**, 137–145 (2003).
19. Bawdon, D., Cox, D. S., Ashford, D., James, A. G. & Thomas, G. H. Identification of axillary Staphylococcus sp. involved in the production of the malodorous thioalcohol 3-methyl-3-sufanylhexan-1-ol. *FEMS Microbiol Lett* **362**, 1–10 (2015).
20. Emter, R. & Natsch, A. The sequential action of a dipeptidase and a beta-lyase is required for the release of the human body odorant 3-methyl-3-sulfanylhexan-1-ol from a secreted Cys-Gly-(S) conjugate by Corynebacteria. *J Biol Chem* **283**, 20645–52 (2008).
21. Natsch, A., Schmid, J. & Flachsmann, F. Identification of Odoriferous Sulfanylalkanols in Human Axilla Secretions and Their Formation through Cleavage of Cysteine Precursors by a C□S Lyase Isolated from Axilla bacteria. *Chem Biodivers* **1**, 1058–1072 (2004).
22. Herman, R. *et al.* High-resolution Staphylococcus profiling reveals intra-species diversity in a single skin niche. *bioRxiv* 2025.04.01.646391 (2025) doi:10.1101/2025.04.01.646391.
23. Herman, R. *et al.* Revealing the diversity of commensal corynebacteria from a single human skin site. *biorxiv* (2024) doi:10.1101/2024.11.28.625817.
24. Rudden, M. *et al.* The molecular basis of thioalcohol production in human body odour. *Sci Rep* **10**, 12500 (2020).
25. Troccaz, M. *et al.* Mapping axillary microbiota responsible for body odours using a culture-independent approach. *Microbiome* **3**, 1–15 (2015).

Chapter 6

Conclusions and future directions

6.1 The potential roles of skin *Corynebacterium* in human malodour formation

Since the seminal study by Shelley *et al*.[1] that suggested a role for the skin microbiome in human malodour formation, different members of the skin microbiome have been implicated in this process. Although early studies in malodour formation of the two major odour classes (i.e. VFAs and thioalcohols) focused on the genus *Corynebacterium*, research later shifted to other skin genera. For example, the genus *Staphylococcus* has become the focus of the most recent thioalcohol malodour research and a pathway to malodour formation has been proposed including the work presented in Chapter 2 of this thesis[2–4].

The move away from *Corynebacterium* in malodour research may reflect stronger evidence for other skin genera, yet the narrow strain coverage in past *in vitro* studies could mean we are underestimating the role of this genus. The study in Chapter 4 discussed a high-throughput whole genome sequencing approach which enabled the identification of 16 genetically diverse *Corynebacterium* from a single high-odour individual, and these were subsequently characterised in the study presented in Chapter 5.

The findings from Chapter 5 suggest roles for *Corynebacterium* in thioalcohol-based malodour formation, in particular with a strain of *C. amycolatum*. Could the utilisation of a more comprehensive collection of skin *Corynebacterium* like that which was reported in Chapter 4 and 5 reveal potential roles in formation of other malodour classes like VFAs? It is certainly possible that some species/strains of this genus capable of contributing to VFA formation may have been overlooked in previous *in vitro* studies. The Corynebacterial enzyme capable of processing major VFA precursors Gln-3M2H and Gln-HMHA has been characterised in the strain *C.* sp. Ax20[5] and there are close homologues of this enzyme in the closely related species *C. glaucum* and across other species of corynebacteria. As the diversity within different microbiomes of the same skin site can vary significantly between individuals and change over time[6], it is difficult to predict what may be the major producer of malodour for a specific person at any one time. Hence, it may be beneficial to explore all possible routes to malodour formation, especially involving the genus *Corynebacterium*.

To date, neither the microbial benefit of biotransforming 3M3SH conjugates nor the origins of 3M3SH itself are clear. It is thought that 3M3SH is conjugated to glutathione (S-glutathionyl-3M3SH) in the apocrine secretory cells[7]. Glutathione is commonly used for detoxification within human cells via the mercapturic pathway [8] and hence, it is possible that 3M3SH was initially conjugated to glutathione for this purpose before being further processed by GGT1 and ABCC11 and released in apocrine secretions[9]. However, in East Asian populations, individuals are commonly found with a mutation on ABCC11 which renders the transporter incapable of exporting 3M3SH conjugates which likely accumulates in the apocrine secretory cells[3]. It is possible that intracellular peptidases could degrade these molecules but the fate of 3M3SH in this situation remains unclear.

6.2 Interactions between members of the skin microbiome community.

Prior to the work presented in this thesis, the biosynthetic route to thioalcohol-based malodour was attributed exclusively to a monophyletic lineage of cutaneous staphylococci, with the initial enzymatic step (dipeptidase activity) presumed to be mediated by an unidentified protein. Work carried out in this thesis not only completes the Cys-Gly-3M3SH biotransformation pathway in the monophyletic group of *Staphylococcus* (see Chapter 2) but also explored cross-species and even cross-genera routes to malodour formation (see Chapter 5).

Alongside the assessment of the isolates of corynebacteria, the potential inter-genera cooperation in malodour formation was explored in Chapter 5. The corresponding collection of genetically diverse axillary staphylococci from the same volunteer (see Chapter 3) was tested alongside the 16 *Corynebacterium* isolates. The conversion of Cys-Gly-3M3SH to 3M3SH in a tandem biotransformation process first involving *Staphylococcus* and then *Corynebacterium* was observed, albeit less strong than the tandem process involving *Corynebacterium* isolates in both steps.

Although the various proposed alternative routes to malodour formation may not always be present in different individual's axilla, findings from Chapter 5 suggest possible pathways that could contribute to malodour (Fig 1). It is perhaps unsurprising that processes contributing to malodour formation could be as complex as proposed.

The skin environment is known to be nutrient limited, and it stands to reason that skin commensals have adapted to utilise any metabolites made available by the host with the by-product of these process being odour molecules. Chapter 5 does not present the first known interactions between skin commensal staphylococci and corynebacteria. For example, it has been reported that *S. lugdunensis* can utilise a siderophore that is made by *C. tuberculostearicum* to support growth and hence, allow for the protection against pathogens like *S. aureus* through the production of the antimicrobial lugdunin[10]. As our understanding of the genetic diversity of the skin microbiome improves, more interactions like these will be revealed and contribute to the interaction networks in complex environments like the skin microbiome

6.3 The axillary microbiome and beyond

Improvements in sequencing technologies have enabled deeper genetic assessments of microbial communities through metagenomics and isolate genomics alike. The work presented in Chapters 3 and 4 utilises the latest high-accuracy long read sequencing on the Oxford Nanopore PromethION platform to provide subspecies/strain-level resolution of two major genera of the same skin site of the same individual.

Unlike *Staphylococcus*, the genus *Corynebacterium* is considered to be within 'microbial dark matter' where lack of available genetic data leads to the poor ability to accurately assign species and strains in sequencing studies. Although there are currently more than 110 species of *Corynebacterium*[11], the number of new skin species identified is continuously growing. This includes the most recent studies (since 2023): Ahmed *et al*. (novel species = 1)[12], Jensen *et al*. (novel species = 2)[13] and even the study reported in Chapter 4 (novel species = 2).

While the use of the selective media in Chapters 3 and 4 provided a wide range of species across both genera, isolates were clearly skewed to some species. For example, abundance of metagenomic reads from volunteer A1 (data not shown) that mapped to *S. epidermidis* was only 3.38x the number of *S. hominis* reads but the proportion of *S. epidermidis* isolates outnumber *S. hominis* by ~23x. Using the analysis performed in Chapters 3 and 4 (e.g. KEGG), further improvements on the selection

media could be made to enrich or reduce specific species to deepen the diversity within the YSM isolate collection for further studies.

Some aspects of the workflow used in Chapters 3 and 4 heavily relies on average nucleotide identities (ANI), including the clustering of isolate genomes to species (ANI $\geq$ 95%) and strain (ANI $\geq$ 99.5%) and subsequent selection of "representative" isolates which were then used in the malodour experiments in Chapter 5. While ANI provides a convenient and robust method to globally assess diversity within a microbial population or collection, this method does not capture total genetic diversity. ANI comparisons are performed with regions of the pair of genomes that can be aligned. This method will exclude any acquired DNA like mobile genetic elements which could have broader implications in understanding the roles of certain species/strains in a microbiome, either generally or in a defined context like malodour formation. One such analysis that could be affected is KEGG as the predicted metabolic profile of the "representative" isolate of a cluster could be different from another isolate of the same cluster. However, there are also limitations to performing KEGG analyses. Firstly, this method relies on the KEGG database of predefined descriptions of each known metabolic pathway and the enzyme families that are involved. Hence, this database is limited to known pathways and may not be frequently updated to include new findings. Further, the KEGG database does not take into account other factors of a functional microbiome like interbacterial warfare (e.g. quorum sensing) which may hamper efforts to holistically understand the microbiome.

Chapters 3 and 4 provide a general overview of the microbial communities that could existed on the human skin and with this information, further analyses should be performed to build on these findings. The same genetic characterisation workflow was also used to characterise other members of the same individual's skin microbiome community (data not presented) and also the axilla of other volunteers. Together, these genetically characterised isolates form the first members of the York Skin Microbiome (YSM) collection of isolates matched to discrete sources. While YSM was established and utilised in this project, the collection also actively facilitating other ongoing malodour projects and could be used in other microbiome studies to understand the complex networks that make up the skin environment.

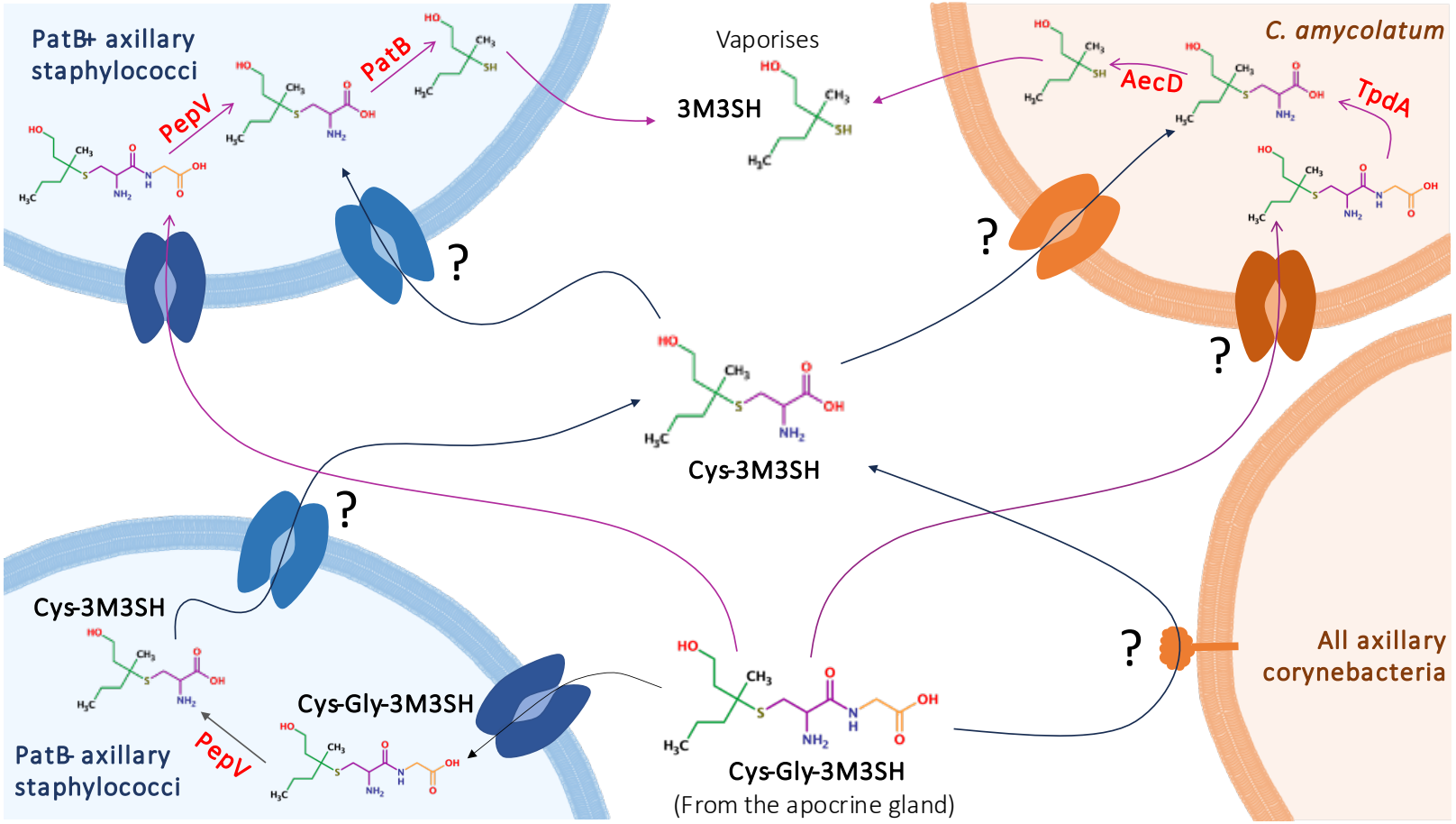


Figure 1. The proposed network contributing to malodour formation. Cys-Gly-3M3SH can be broken down to 3M3SH by PatB-containing/positive staphylococci including *S. hominis* and *C. amycolatum* (purple open head arrows) on their own. Alternatively, Cys-Gly-3M3SH could be partially broken down either by PatB-negative axillary staphylococci or corynebacteria to produce Cys-3M3SH (black arrows). The freed Cys-3M3SH could then be taken up by *C. amycolatum* or PatB-containing staphylococci to complete the biotransformation process. Unknown components of each pathway are indicated by a question mark.

6.4 Future directions

6.4.1 The roles of axillary *Corynebacterium* in other high-odour volunteers

The findings from Chapter 5 open new considerations when assessing the contributions by members of the skin microbiome to malodour production. However, it is acknowledged that this study was performed with axillary isolates from one individual and may not be representative of others. The findings, including the unexpected *Corynebacterium*-only Cys-Gly-3M3SH tandem biotransformation route, were derived from activities of likely skin species. As genetic differences can be expected between isolates of the same species from different individuals, the study in Chapter 5 should be replicated using isolates from other high-odour volunteers.

The YSM collection currently comprises isolates derived from the axilla of three other individuals, one of which was also assessed to have high-malodour. At the point of writing, only a small number of *Corynebacterium* isolates from this volunteer (see Chapter 4.6.1, volunteer A3) were genetically characterised. The same genetic characterisation approach should be extended to the remaining *Corynebacterium* isolates and to the *Staphylococcus* isolates before assessing the different possible routes to Cys-Gly-3M3SH biotransformation.

6.4.2 Identifying and characterising microbial components that could contribute to the alternative pathways to malodour formation

The discovery of alternative routes to thioalcohol malodour formation reveals new targets to uncover and understand the molecular basis of these pathways (Fig 1).

6.4.2.1 Single-organism biotransformation pathway – *C. amycolatum*

Previous studies in *C.* sp Ax20 of the two Corynebacterial enzymatic pathways suggest that *C. amycolatum* could break down Cys-Gly-3M3SH to 3M3SH using existing homologues of the intracellular enzymes TpdA and AecD. However, the transporter capable of Cys-Gly-3M3SH uptake in any *Corynebacterium* sp. remains unknown. A di-/tri-peptide permease (Dtp) transporter of the same family as the

Staphylococcal Cys-Gly-3M3SH transporter DtpT[2] was identified in *C. amycolatum*. This transporter could be similarly tested to assess Cys-Gly-3M3SH transport activity.

6.4.2.2 Tandem biotransformation – *Staphylococcus/Corynebacterium*

The *Staphylococcus*-*Corynebacterium* tandem biotransformation route requires the transport of Cys-Gly-3M3SH into the *Staphylococcus* cell for the dipeptidase step. The resulting Cys-3M3SH could then be actively exported into the environment before being taken up by the *Corynebacterium* or PatB-containing *Staphylococcus* to complete the biotransformation process. However, there has not been any reported Cys-3M3SH transporters for any organism for export or import. This molecule is, in essence, the amino acid L-cysteine with a heavily modified sidechain (Fig 1). Hence, cysteine transporters could be tested for the capability to accommodate this molecule. Like the characterised enzymes and transporters known to be involved in malodour formation, the Cys-3M3SH transporter is likely to possess a binding pocket capable of accommodating the largely hydrophobic 3M3SH moiety. Prior to experimental testing, the predicted structures (e.g. AlphaFold) of these putative transporters could be assessed *in silico* for the presence of this binding pocket.

6.4.2.3 Tandem biotransformation – *Corynebacterium/Corynebacterium*

Performing the first enzymatic step of the tandem biotransformation route will involve a dipeptidase. The Corynebacterial Cys-Gly-3M3SH dipeptidase TpdA is an intracellular dipeptidase but ubiquitous Corynebacterial Cys-Gly-3M3SH import has not been demonstrated. To rationalise the findings in Chapter 5, it was proposed that there may yet be another Cys-Gly-3M3SH dipeptidase but is either secreted or anchored to the cell surface (represented as surface anchored in Fig 1), removing the need for transport to carry out the first enzymatic step. If this proposed process is true, the resulting Cys-3M3SH is already in the environment and will then only need to be imported by *C. amycolatum* to complete the tandem biotransformation process. Hence, the identification of this secreted or cell surface anchored dipeptidase could be another avenue for exploration.

6.4.3 High-resolution genetics to understand the networks within a functional microbiome.

Through high-resolution genetics, Chapters 3 and 4 provide glimpses of the complexities just within a single skin microbiome community (i.e. the axilla of volunteer A1) through the assessment of all genes present in all isolates (~2300 predicted genes in isolates of both genera). Unlike other studies, both Chapters 3 and 4 employ non-discriminatory approaches to asses isolates of two genera and identified genes and metabolic pathways that could contribute to a functional microbiome. Clearly there is much that is unknown about interactions within a single microbial community. A much wider study could employ the approaches discussed in this thesis to truly investigate in great detail different individual microbiomes to understand how this essential component of the human skin maintains itself and the host. These findings could then be used to understand how these microbiomes could interact with the resident host cells of the respective skin niche, then to other body sites of the same individual and maybe even understanding how this external facing microbiome could be affected by or play a role in human-human interactions beyond malodour[14].

6.5 Final reflections

Our understanding of the skin microbiome only scratches the surface. The development and improvement of sequencing technologies have enabled higher resolution genetics, paving the way to better understand the diversity and networks within this complex environment. The work presented in this thesis demonstrates the power of combining culturomics and genomics to help construct a collection of isolates from the same source which were used to reveal novel routes to thioalcohol-based malodour formation and restore the genus *Corynebacterium* as possible contributors of this ancient human trait.

6.6 Bibliography

1. Shelley, W. B., Hurley, H. J. & Nichols, A. C. Axillary odor; experimental study of the role of bacteria, apocrine sweat, and deodorants. *AMA Arch Derm Syphilol* **68**, 430–46 (1953).
2. Minhas, G. S. *et al.* Structural basis of malodour precursor transport in the human axilla. *Elife* **7**, e34995 (2018).
3. Rudden, M. *et al.* The molecular basis of thioalcohol production in human body odour. *Sci Rep* **10**, 12500 (2020).
4. Herman, R. *et al.* Identification of a Staphylococcal dipeptidase involved in the production of human body odour. *Journal of Biological Chemistry* **300**, 107928 (2024).
5. Natsch, A., Gfeller, H., Gygax, P., Schmid, J. & Acuna, G. A Specific Bacterial Aminoacylase Cleaves Odorant Precursors Secreted in the Human Axilla. *Journal of Biological Chemistry* **278**, 5718–5727 (2003).
6. Grice, E. A. *et al.* Topographical and temporal diversity of the human skin microbiome. *Science* **324**, 1190–1192 (2009).
7. Baumann, T. *et al.* Glutathione-conjugated sulfanylalkanols are substrates for ABCC11 and γ-glutamyl transferase 1: A potential new pathway for the formation of odorant precursors in the apocrine sweat gland. *Exp Dermatol* **23**, 247–252 (2014).
8. Cooper, A. J. L. & Hanigan, M. H. Metabolism of Glutathione S-Conjugates: Multiple Pathways. *Comprehensive Toxicology* **10**, 363 (2017).
9. Yoshiura, K. I. *et al.* A SNP in the ABCC11 gene is the determinant of human earwax type. *Nat Genet* **38**, 324–330 (2006).
10. Rosenstein, R. *et al.* The Staphylococcus aureus-antagonizing human nasal commensal Staphylococcus lugdunensis depends on siderophore piracy. *Microbiome* **12**, 1–17 (2024).
11. Oliveira, A. *et al.* Insight of genus *Corynebacterium*: Ascertaining the role of pathogenic and non-pathogenic species. *Front Microbiol* **8**, 286540 (2017).
12. Ahmed, N. *et al.* Genomic characterization of the *C. tuberculostearicum* species complex, a prominent member of the human skin microbiome. *mSystems* **8**, e00632-23 (2023).
13. Jensen, M. G. *et al.* Species- and strain-level diversity of Corynebacteria isolated from human facial skin. *BMC Microbiol* **23**, 1–14 (2023).

14. Lax, S. *et al.* Longitudinal analysis of microbial interaction between humans and the indoor environment. *Science (1979)* **345**, 1048–1052 (2014).

List of Abbreviations

Below is a list of abbreviations used throughout the thesis:

3M2H	3-methyl-2-hexanoic acid
3M3SH	3-methyl-3-sulfanylhexan-1-ol
A_{412}	Absorbance at 412 nm
ABC	ATP-binding cassette
AEX	Anion exchange
ANI	Average nucleotide identity
ASV	Amplicon sequence variant
BGC	Biosynthetic gene cluster
BHIT	Brain-heart infusion broth supplemented with 1% Tween-80
COG	Clusters of Orthologous Groups
DTNB	5-5'-dithiobis-(2-nitrobenzoic acid)
EDTA	Ethylenediaminetetraacetic acid
EUCAST	European committee on antimicrobial susceptibility testing
GMQE	Global Model Quality Estimation
GTDB	Genome taxonomy database
HIC	Hydrophobic interactions chromatography
HMHA	3-hydroxy-3-methylhexanoic acid
KEGG	Kyoto Encyclopedia of Genes and Genomes
KPi	Potassium phosphate
LB	Luria-bertani
MALDI-MS/MS	Matrix-assisted laser desorption/ionisation tandem mass spectrometry
MLST	Multi-locus sequence typing
NAPAA	Non-alpha poly-amino acids
NRPS	Non-ribosomal polyketide synthase
ns	Not significant
OD	Optical density
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction

PDB	Protein data bank
PKS	Polyketide synthase
PLP	Pyridoxal 5’-phosphate
POT	Proton-dependent oligopeptide transporter
QMEANDisCo	Qualitative Model Energy Analysis – Distance Constraints
rRNA	Ribosomal ribonucleic acid
RM	Restriction modification
s.d.	Standard deviation
SEC	Size exclusion chromatography
ST	Sequence type
T1PKS	Type 1 polyketide synthase
T3PKS	Type 3 polyketide synthase
TCEP	Tris(2-carboxyethyl)phosphine
TLC	Thin-layer chromatography
TNB	2-nitro-5-sulfanylbenzoic acid
TSB	Tryptic soy broth
VFA	Volatile fatty acid
WGS	Whole genome sequencing
WT	Wild-type
YSM	York Skin Microbiome