

Title	Characterisation and applications of bacteriocinogenic Isolates from the deep-sea fish microbiome
Authors	Uniacke-Lowe, Shona
Publication date	2024
Original Citation	Uniacke-Lowe, S. M. 2024. Characterisation and applications of bacteriocinogenic Isolates from the deep-sea fish microbiome. PhD Thesis, University College Cork.
Type of publication	Doctoral thesis
Rights	© 2024, Shona Maria Uniacke-Lowe. - https://creativecommons.org/licenses/by-nc-nd/4.0/
Download date	2026-11-01 11:14:07
Item downloaded from	https://hdl.handle.net/10468/16393

Ollscoil na hÉireann, Corcaigh
National University of Ireland, Cork



**Characterisation and Applications of
Bacteriocinogenic Isolates from the Deep-sea
Fish Microbiome**

Thesis presented by

Shona Maria Uniacke-Lowe, BSc. (Hons)

for the degree of

Doctor of Philosophy

Teagasc Food Research Centre, Moorepark, Ireland

APC Microbiome Ireland, University College Cork

School of Microbiology, University College Cork

Head of School: Prof. David Clarke

Supervisors: Prof. R. Paul Ross, Prof. Colin Hill & Prof. Catherine
Stanton

2024



Table of Contents

Table of Contents	i
Declaration	iii
List of Publications	iv
Thesis Abstract	v
Chapter 1	1
Abstract.....	2
1.1. Introduction	3
1.2. The Diversity of the Marine Fish Gut Microbiome	5
1.3. Bacteriocins from Marine Fish Gut Microbiota	8
1.4. ‘Extreme’ Marine Environment Impact on Microbial Products	12
1.5. Applications of Marine Fish -Derived Bacteriocins	14
1.6. Challenges, Metagenomics and Future Prospects	15
1.7. Conclusions	18
1.8. References	19
1.9. Chapter 1 Figures and Tables.....	28
Chapter 2	33
Abstract.....	34
2.1. Introduction	35
2.2. Materials and Methods	38
2.3. Results.....	42
2.4. Discussion.....	48
2.5. Conclusions	53
2.6. References	54
2.7. Chapter 2 Figures and Tables.....	61
2.8. Chapter 2 Supplementary Material.....	76
Chapter 3	91
Abstract.....	92
3.1. Introduction	93
3.2. Materials and Methods	95

3.3. Results	100
3.4. Description of <i>Winogradskyella bathintestinalis</i> sp. nov.	106
3.5. References	108
3.6. Chapter 3 Figures and Tables.....	113
3.7. Chapter 3 Supplementary Material.....	122
Chapter 4	130
Abstract.....	131
4.1. Introduction	132
4.2. Materials and Methods	134
4.3. Results	139
4.4. Description of <i>Planococcus notacanthi</i> sp. nov.....	143
4.5. References	145
4.7. Chapter 4 Supplementary Material.....	156
Appendix to Chapter 4.....	163
Chapter 5	169
Abstract.....	170
5.1. Introduction	171
5.2. Materials and Methods	173
5.3. Results	177
5.4. Discussion	182
5.5. Conclusions	186
5.6. References	187
5.7. Chapter 5 Figures and Tables.....	192
5.8. Chapter 5 Supplementary Material.....	206
Chapter 6	211
References	222
Acknowledgements	224

Declaration

This is to certify that the work I am submitting is my own and has not been submitted for another degree, either at University College Cork or elsewhere. All external references and sources are clearly acknowledged and identified within the contents. I have read and understood the regulations of University College Cork concerning plagiarism and intellectual property.

List of Publications

Uniacke-Lowe, S., Collins, F. W. J., Hill, C. & Ross, R. P. (2023). Bioactivity Screening and Genomic Analysis Reveals Deep-Sea Fish Microbiome Isolates as Sources of Novel Antimicrobials. *Marine Drugs*, 21, 444.

Uniacke-Lowe S., Johnson C.N., Stanton C., Hill C., Ross P. (2023). *Winogradskyella bathintestinalis* sp. nov., isolated from the intestine of the deep-sea loosejaw dragonfish, *Malacosteus niger*. *International Journal of Systematic and Evolutionary Microbiology*;73(10).

Uniacke-Lowe S., Stanton C., Hill C., Ross P. (2024). *Planococcus notacanthi* sp. nov., isolated from the skin of a deep-sea snub-nosed spiny eel. *International Journal of Systematic and Evolutionary Microbiology*;74(3).

Thesis Abstract

One of the major grand challenges confronting humanity in the coming decades is the increased risk of infection, stemming from the alarming surge in antimicrobial resistance among pathogenic bacteria. Consequently, there is an urgent need to uncover new antimicrobials and mechanisms, including antimicrobial bacterial strains, which could potentially substitute for some of the widely employed antibiotics. In this respect, the main goal of this thesis work was initially to explore deep-sea fish microbiomes as a source of antimicrobials and their prospective clinical applications. Our hypothesis was that these microbiomes represent some of the last bastions of microbial communities that have been negligibly impacted by human activity and as such are untapped for this purpose.

Although still very limited, there is an accumulating number of studies where marine fish microbiomes are being characterised and assessed for their bacteriocinogenic potential (Chapter 1). In many cases these microbiomes are dominated by Gram-negative Pseudomonadota and Gram-positive Bacillota and Actinomycetota. They can be important hosts to uncommon species that are not found in the surrounding waters, as well as to novel taxa. Of particular interest, in terms of their antimicrobial production, are bacilli, lactic acid bacteria and actinobacteria. Already, a limited number of novel bacteriocins from marine fish gut isolates have been identified and characterised. These include Formicin, a class I (lanthipeptide) bacteriocin from mackerel, several class IIa bacteriocins, and CAMT6, a very recently identified class IIId bacteriocin. These novel bacteriocins cover a range of molecular structures, with varying spectra of activity – some of which were capable of targeting important pathogenic strains, such as *Clostridium difficile* and *Listeria monocytogenes*. Bacteriocins from marine fish microbiota also have diverse applications, not only for therapeutic, but also as food preservatives and probiotics. This provides great scope for the potential of finding novel bacteriocins from deep-sea fish microbiomes.

Chapter two outlines an exploration of bacterial isolates from the gut and skin of deep-sea fish for antimicrobial production. *In vitro* screening was initially used to

screen for antimicrobial producers with activity against various indicators including foodborne pathogenic bacteria. Subsequent *in silico* screening of the genomes of selected producers revealed the presence of a wide range of secondary metabolite biosynthetic gene clusters (BGCs) as well as potentially novel bacteriocins, including class I (lanthipeptide), class IIc (circular) and class IIId bacteriocins. Genomic screening also revealed the almost complete absence of antimicrobial resistance genes.

Chapter three outlines the description of a novel species, *Winogradskyella bathintestinalis* APC 3343, which was isolated from the intestine of a deep-sea dragon fish, *Malacosteus niger*, as part of the screening study. This yellow pigment-producing strain was differentiated from related type strains of *Winogradskyella* species based on a combination of 16S, genomic, biochemical, and phenotypic characterisation. This study represents one of first to characterise a novel *Winogradskyella* species isolated from a deep-sea fish.

Chapter four describes the characterisation of a novel *Planococcus* species, isolated from the skin of a deep-sea ‘snub nose spiny eel’ (*Notacanthus chemnitzii*), for which the name *Planococcus notacanthi* was proposed. This strain, APC 4016, had been shown to inhibit the growth of Gram-positive target strains and was subsequently found encode a potentially novel lanthipeptide with similarity to cerecidin of *Bacillus cereus*. Based on phylogenetic analysis and its distinctive phenotypic characteristics, such as a tolerance of a wide range of growth conditions and the presence of menaquinone 5 as a respiratory quinone, *Planococcus notacanthi* was deemed a novel species, with APC 4016 as the type strain.

We also explored the application of these deep-sea isolates and their DNA as controls for determining relative and/or absolute quantification of bacteria, in terms of cell numbers, in human metagenomic sequencing experiments. Quantified genomic DNA from two deep-sea species (crucially that have not been identified in human microbiome samples) was applied to human faecal samples and subsequently extracted and detected using 16S rDNA sequencing and qPCR, alongside the microbiome DNA. The preliminary data shows that the “spiked-in” deep-sea DNA has the potential to be used as a marker for relative abundance of bacterial cell numbers.

Chapter five outlines the identification of an agarolytic (agar-degrading) isolate from the skin of a deep-sea fish and the subsequent molecular characterisation of one β -agarase and one putative B agarase encoded by the isolate. Sequence analysis of the two agarase proteins, Aga4017A and Aga4017B, showed that they shared structural similarities and key characteristics to GH16 and GH50 family β -agarases, respectively. Recombinant pCR2.1-TOPO vectors encoding *aga4017A* or *aga4017B* were created and cloned into *E. coli* TOP10. Transformants containing this pCR2.1-TOPO_aga4017A vector were able to express the agarase protein and exhibited agarolytic activity. The predicted 3D model of Aga4017A was also elucidated using Alphafold2 and revealed it to be a globular protein with distinctive modules corresponding to a signal peptide, a catalytic GH16 β -agarase domain and a ricin B lectin -like carbohydrate-binding domain.

Overall, this work demonstrates the biodiscovery potential of the microbiome of deep-sea fish, in terms of its microbial inhabitants, the antimicrobials they produce, and also their functionalities, which may have applications for utilisation in the field of microbiome research.

Chapter 1

The Marine Fish Gut Microbiome as a Source of Novel Bacteriocins

Abstract

The marine environment is the largest ecological habitat on Earth, albeit one of the least explored, particularly in terms of its microbial inhabitants. The marine fish gut is host to a diverse microbial community from which diverse bioactive molecules can be sourced. Due to the unique environmental pressures these microbial communities experience, the bioactive molecules they produce often evolve unique adaptations that give them diverse structures and activities, differentiating them from terrestrial homologues. Of particular interest, due to their structural and functional diversity, are the ribosomally-synthesised antimicrobial peptides (bacteriocins). With increasing pressure from emerging antibiotic-resistant disease and industrial demand for novel therapeutics, the marine fish gut microbiome represents a relatively untapped resource of novel bacteriocins that could prove beneficial to human health and aquaculture. This review presents an overview of the marine fish gut microbiome and explores its potential as a source of bacteriocins for human health with considerations for applications and future research in this area.

1.1. Introduction

The marine environment is the largest ecological habitat on the planet. The Earth's oceans cover over 70 % of the planet's surface and reach depths of over 10,000m (Greenaway et al., 2021). Each layer, or zone, of the ocean is its own ecological niche that is distinguishable based on factors such as temperature, salinity and sunlight availability (National Oceanic and Atmospheric Administration). The top layers (epipelagic and mesopelagic zones) sustain a high abundance of diverse life and are very metabolically active. In contrast, the deep ocean layers (bathypelagic, abyssal and hadal zones) are subject to low temperatures, high hydrostatic pressure and lack of sunlight (Somero, 1992). From hydrothermal vents to the deep, dark trenches, marine fish inhabit extreme environments and have as a result acquired diverse and unique physiological and molecular adaptations.

The marine environment is also a rich source of diverse natural bioactive compounds with therapeutic and economic potential. To date, there are 15 clinically approved drugs derived from marine organisms, ranging from peptides to antibodies and covering a wide variety of biological activities. Many more are in various stages of clinical trials (Approved Marine Drugs, 2023). We are in an age when there is a huge demand for novel antimicrobials and alternative treatments to antibiotics. Antimicrobial resistance (AMR) is an extremely concerning and global threat to human health. It has been estimated that in 2019 alone, AMR was associated with 4.95 million deaths (Antimicrobial Resistance Collaborators, 2022). Traditional antibiotic treatments have become less effective, multidrug-resistant pathogens are more common yet the development of new antibiotic-based treatments has slowed dramatically over the past few decades (Frieri et al., 2017). From 2014 to 2018 over 250 preclinical antimicrobial compounds from marine organisms were reported (Mayer et al., 2020, Mayer et al., 2021, Mayer et al., 2022). The majority of these compounds were polyketides (35 %), terpenoids (26 %), alkaloids (17 %) and peptides (14 %), as well as lipids, lipopeptides and several other molecule types. Just over half of these molecules were produced by bacteria or fungi, and, significantly, these producers were isolated from higher organisms, particularly from marine sponges.

The mammalian gut microbiome is a well-established source of beneficial bacteria and compounds with therapeutic potential (Halloran and Underwood, 2019, Wang et al., 2019), and undoubtedly, this scenario is paralleled in the fish gut microbiome. There is growing interest in bioprospecting the fish gut microbiome to find novel bioactive compounds with therapeutic potential and for applications in aquaculture (Wanka et al., 2018). The gut microbiome of fish has crucial biological functions in digestion, immune system modulation, stress response, and protection from pathogens/infection (Butt and Volkoff, 2019). These functions are often mediated by microorganisms and microbial metabolites, such as bacteriocins.

Bacteriocins are low molecular weight (<10 kDa), ribosomally-synthesised peptides with antimicrobial activity produced by bacteria that are immune to their own bacteriocin. The antagonistic activity of bacteriocins can be bacteriocidal or bacteriostatic, and the range of susceptible target organisms can be described as broad or narrow. Bacteriocins are usually classified into two major classes, based mainly on the extent to which they are modified. Class I bacteriocins are those which undergo extensive post-translational modification (e.g., lantibiotics or lanthipeptides), and class II bacteriocins are unmodified or cyclic (Cotter et al., 2013). The class I lanthipeptides, for example, are synthesised as prepeptides with unique lanthionine and/or β -methyllanthionine residues that undergo extensive posttranslational modification resulting in the formation of signature (methyl)lanthionine rings. Lanthipeptides can be further divided into subclasses I–V based on the enzyme(s) involved in their biosynthesis, modification and mode of action (Montalbán-López et al., 2021). The class II bacteriocins are subdivided into four groups (classes IIa-d): the anti-listerial pediocin-like peptides (IIa), two-component peptides (IIb), cyclic/circular bacteriocins (IIc) and the linear and non-pediocin-like peptides (IId) (Nissen-Meyer et al., 2009). Bacteriocins are emerging as promising alternatives to conventional antibiotics due to their spectrum of activity – which can be broad or narrow, thermostability and their capacity for bioengineering and generation of variants with value-added properties (Cotter et al., 2013).

Attention is now turning to fish microbiomes as novel sources of bioactive molecules such as bacteriocins. Conventionally, the diversity of fish microbiomes has been determined using culture-dependent methods and has very likely been

underestimated in terms of its abundance and richness. However, ongoing advances in culture-independent methodologies, such as metagenomic sequencing, are enabling more accurate analysis of this microbial niche. Because of the huge expanse of the marine environment, marine microorganisms are among the least explored and least accessible and thus the application of culture-independent metagenomic analysis becomes even more valuable. The marine fish gut microbiome is essentially an untapped reservoir for novel antimicrobials.

In this review, we discuss the diversity of the gut microbiomes of marine fish, with a focus on finfish: Chondrichthyes (cartilaginous fish) and Osteichthyes (bony fish). We also discuss current research on the discovery of bacteriocins from marine fish gut microbiota, their applications, and the current prospects of this microbial niche as a source of novel antimicrobials.

1.2. The Diversity of the Marine Fish Gut Microbiome

The marine fish gut is host to a diverse microbial community largely consisting of bacteria, with reported counts ranging from 10^4 – 10^9 CFU/g (Egerton et al., 2018), as well as fungi, archaea and viruses. This community begins to form at the fish larval stage, with early colonisers originating from the egg surface, the surrounding water and the first feed (Roeselers et al., 2011). The diversity of the fish gut microbiome is then influenced by biotic and abiotic factors such as host phylogeny, trophic level (diet) and environmental salinity (Sullam et al., 2012), as well as, though to a lesser extent, environmental pH and temperature (Ward et al., 2009, Egerton et al., 2018). The fish gut microbiome is an important and unique niche within the marine environment, with reports that the composition of microorganisms found in marine fish gut can differ from that of the surrounding waters and can even contain species that are rarely, if at all, found in the surrounding water (Troussellier et al., 2017).

The composition of the marine fish gut microbiome has already been extensively reviewed (Egerton et al., 2018, Romero et al., 2014, Ou et al., 2021). Early studies on fish microbiome diversity were limited to the culturable organisms (Yano et al., 1995, Ohwada et al., 1980), however, with the emergence of next-generation sequencing (NGS) and metagenomic technologies there is now a growing number of culture-

independent studies that have been able to provide a more comprehensive description of the fish gut microbiome (Ghanbari et al., 2015). Findings from more recent studies which have utilised culture-independent methods are given in Table 1. The most frequently reported phylum is Pseudomonadota (formerly Proteobacteria), followed by Bacillota (Firmicutes), Actinomycetota (Actinobacteria) and Bacteroidota (Bacteroidetes). Other abundant phyla include Mycoplasmatota (Tenericutes), Cyanobacteriota (Cyanobacteria), Fusobacteriota (Fusobacteria) and Chloroflexota (Chloroflexi). At a lower taxonomic level, Vibrionales are frequently detected, particularly species of *Vibrio* and *Photobacterium*, as well as members of the families *Moraxellaceae*, *Pseudomonadaceae* and *Micrococcaceae* (Table 2). Host diet is a significant determinant of the predominant taxa detected within the gut microbiome. Results reported by Sullam et al. (2012) suggested that Bacteroidales and Clostridiales dominate the gut microbiomes of herbivorous marine fish, whilst Vibrionales and Alteromonadales are dominant in carnivorous and omnivorous marine fish. Similarly, Huang et al. (2020) demonstrated that host feeding habits could be differentiated by “indicators”, specific bacterial taxa. It was suggested that these indicators were associated with specific symbiotic functional activities such as the production of bioactive enzymes, thus aiding in digestion within the gut.

The majority of studies of marine fish gut microbiota have focused on 16S rRNA sequence data, profiling the bacterial portion of this community. However, one recent study analysed whole-metagenome shotgun sequencing data from the intestines of various deep-sea fish of the northern Atlantic Ocean (Collins 2019). Overall, the most abundant phyla detected were Pseudomonadota, Bacillota, Actinomycetota and Bacteroidota, in agreement with previous studies. However, cluster analysis showed that the proportions of these phyla varied significantly between some samples. Furthermore, in one cluster a significant proportion of Eukaryota was detected, specifically Ascomycota, Basidiomycota and Euryarchaeota. In terms of the eukaryotic species, there have been suggestions that yeasts are commensals of the fish microbiome with species such as *Debaryomyces hansenii*, *Saccharomyces cerevisiae* and red-pigmented *Rhodotorula* dominating (Romero et al., 2014, Andlid et al., 1995). The main archaeal groups reported in studies of marine

fish gut microbiota are Crenarchaeota and Euryarchaeota (2016, van der Maarel et al., 1999). Euryarchaeota have also been identified in deep-sea fish microbiota samples (Collins, 2019). These are two of the major groups of planktonic archaea found in marine environments and are key contributors to nutrient cycling in pelagic waters and deep-sea sediment (Vuillemin et al., 2019). It has been suggested that the digestive tract of fish is an important habitat for archaeal groups within the marine environment, particularly the obligate anaerobic Euryarchaeota (van der Maarel et al., 1999).

Viruses are the most abundant entities on earth, and it is no different in the marine environment. It has been estimated that there are between 10^6 and 10^8 viruses per ml of seawater (Suttle, 2005). A recent study by Geoghegan et al. (2021) used a metagenomic-based approach to characterise the viromes of several marine fish species: They identified viral sequences that represented 11 viral families, with *Astroviridae*, *Picornaviridae*, *Arenviridae*, *Reoviridae* and *Hepadnaviridae* dominating. They suggested that many of the identified viruses were diet or microbiome - associated and that host phylogenetics is a significant determinant of virome diversity.

Efforts have been made to define a “core microbiome” for marine fish, however this has proven a difficult task given the diversity of marine fish and the factors that affect their microbiome diversity. Fish gut microbiota members can also be transient (Egerton et al., 2018). Instead, it has been suggested that, perhaps, functional diversity is more important than taxonomic/phylogenetic diversity (Ghanbari et al., 2015, Huang et al., 2020). Furthermore, many of these studies have noted that a high percentage of operational taxonomic units (OTUs) could not be taxonomically assigned at the genus level, indicating the novelty and high biodiscovery potential of the marine fish gut microbiomes. For example, Huang et al. (2020) reported that over 70 % of OTUs from coastal fish gastrointestinal samples were unassigned at genus level, whilst Johny et al. (2018a) reported over 90 % of the deep-sea fish gut OTUs could not be assigned at genus level.

The gut microbiome of marine fish is involved in the regulation of host processes including digestion, stress and immune responses, reproduction, and metabolism

(Sullam et al., 2012, Butt and Volkoff, 2019). Many of these processes are mediated by the production of microbial metabolites, including antimicrobials, and bioactive enzymes.

1.3. Bacteriocins from Marine Fish Gut Microbiota

Bioactive molecules representing several different bacteriocin classes and bacteriocin-like peptides have been identified in isolates of marine fish gut microbiomes, as outlined in Table 3. These peptides exhibit activity against a wide range of Gram-positive and Gram-negative bacteria, and even against fungi in some cases. Though in many of these studies the chemical structure of these bacteriocins/peptides is yet to be determined, it again highlights the biodiscovery potential and novelty of this microbial niche. Notably, these bacteriocins are all from bacilli or lactic acid bacteria (LAB) strains, some of which are novel species or strains.

1.3.1. Bacteriocins from LAB

Lactic acid bacteria (LAB) are important members of the fish gut microbiome with cultivable species from marine fish including *Carnobacterium* spp., *Lactobacillus* spp., *Lactococcus* spp., *Enterococcus* spp. and *Pediococcus* spp. (Ringø et al., 2018). Studies have shown that beneficial LAB in the fish gut may have a significant role in promoting host fish health by improving resistance to infection through immune system modulation and the production of antimicrobial molecules, such as bacteriocins (Ringø et al., 2018).

Pilet et al. (1995) screened for antimicrobial-producing strains from the intestines of salmon and trout using *Listeria* species as target organisms. They identified two antimicrobial-producing species, *Carnobacterium piscicola* V1 and *Carnobacterium divergens* V41. *C. piscicola* V1 was found to produce two class IIa bacteriocins, piscicocins V1a and V1b. The purified peptides were active against the same spectrum of Gram-positive indicator strains, which included *C. divergens*, *E. faecalis* and *L. monocytogenes*. Furthermore, piscicocin V1a was significantly more potent than V1b and was deemed to be a novel bacteriocin based on amino acid sequence analysis (Bhugaloo-Vial et al., 1996). In a subsequent study, *C. divergens* V41 was found to produce divercin V41, also a novel class IIa bacteriocin. The amino acid sequence of divercin V41 was shown to be most similar to those of pediocin PA-1 and

enterocin A (Metivier et al., 1998). Furthermore, piscicocins V1a and V1b and divercin V41 were optimally produced at temperatures below 20 °C and in the presence of up to 4 % NaCl (Duffes et al., 1999), likely an adaptation to the marine environment.

Enterococcus mundtii Tw56 was isolated from the intestine of an *Odontesthes platensis* specimen from the Patagonian region of Argentina, an area subject to relatively low temperatures. Cell-free supernatant (CFS) from this strain was able to inhibit Gram-positive and Gram-negative strains including *Enterococcus* spp., *Listeria* spp., *Pseudomonas aeruginosa* and *Shewanella putrefaciens*. It was deduced through thermal and chemical stability assays and PCR that the CFS contained mundticin KS, a class IIa bacteriocin (Schelegueda et al., 2015).

Two separate studies found nisin Z -producing *Lactococcus lactis* strains from marine fish intestines with activity against the fish pathogens, *Streptococcus iniae* and *Lactococcus garvieae*, respectively (Heo et al., 2012, Sequeiros et al., 2015). Nisin Z is a natural variant of the class I bacteriocin (lanthipeptide) nisin A, differing only by a single amino acid (His27Asn) (Mulders et al., 1991) and contains the characteristic lanthionine and methyllanthionine rings of this bacteriocin class (Figure 1). As with other nisin peptides, nisin Z acts by binding to and forming pores in the membrane of target cells leading to cell death (Breukink et al., 1997). Sequeiros et al. (2015) demonstrated this in their study when they exposed *L. garvieae* to supernatant from the nisin Z-producing fish isolate (*L. lactis* TW34) and reported a drop in viable cell counts by 6 logs (CFU/mL) within one hour but without cell lysis, thus indicating bactericidal activity. Heo et al. (2012) prepared partially purified nisin Z in a solution containing varying concentrations of NaCl from 0–4 % w/v. They showed that nisin Z prepared in 3.5 % NaCl, a concentration similar to that of seawater, was the most effective against *S. iniae*. This demonstrates the increased effectiveness of the marine-derived nisin Z under native conditions.

More recently, Li et al. (2023) isolated a novel class IIId bacteriocin, CAMT6 from the marine fish isolate *Enterococcus durans* YQ-6. CFS from this strain possessed a broad spectrum of activity, targeting both Gram-positive and Gram-negative bacteria, and also inhibited fungi. Measurement of CFS conductivity and cell morphology imaging of *L. monocytogenes* after exposure to CAMT6 demonstrated that the mode of action of CAMT6 involved disruption of the cell membrane, causing leakage of cell contents.

Furthermore, the purified CAMT6 peptide exhibited anti-*Listeria* activity in a chicken breast model and anti-biofilm activity, demonstrating a potential application in food preservation. The amino acid sequence of CAMT6 is unusual in that, at 12 amino acids in length, it is much shorter than previously reported bacteriocins from enterococci and did not share homology with bacterial-derived proteins, but, rather, was most similar to antimicrobial peptides from humans and sheep.

1.3.2. Bacteriocins from Bacilli

Bacilli isolated from marine fish have also been found to produce highly stable, and often novel, bacteriocins that can target a broad range of target strains:

Bindiya et al. (2015) screened isolates from the gut contents of a deep-sea shark (*Centroscyllium fabricii*) against various indicator strains. Amongst the antimicrobial producers, they identified *Bacillus amyloliquefaciens* BTSS-3, a strain with activity against pathogenic bacteria including *Salmonella enterica* Typhimurium, *Clostridium perfringens*, *S. aureus*, *Proteus vulgaris*, and several *Bacillus* species (Bindiya et al., 2015). The purified antimicrobial molecule designated BaCf3, was subsequently shown to possess characteristics of being a bacteriocin: it had a mass of approximately 3000 Da, was pH tolerant (pH 2.0–13.0) and highly thermostable (4–100 °C) and retained activity even after autoclaving (Bindiya et al., 2019). Sequence compositional analysis found BaCf3 to be rich in hydrophobic amino acids, such as cysteine and glycine, and was predicted to contain at least one disulfide bridge. The predicted tertiary structure consisted of three anti-parallel β -sheets and resembled that of laterosporulin, a defensin-like bacteriocin from *Brevibacillus* sp. strain GI-9 (Singh et al., 2012). Based on its hydrophobic nature and structural similarity to defensin-like bacteriocins, it was predicted that BaCf3 targeted the cell membrane. This was confirmed by microscopic analysis of *Bacillus circulans* target cells after exposure to BaCf3, whereby leakage of intracellular material and disruption to the cell membrane was observed (Saidumohamed et al., 2022). Furthermore, BaCf3 was shown to also have anti-biofilm and anti-cancer activities (Bindiya et al., 2019, Saidumohamed et al., 2022). This is one of the first studies that has purified a novel bacteriocin from a deep-sea fish gut isolate, and, significantly, it retained

antimicrobial activity under low temperatures (4 °C) and after being subjected to high pressure and heat (autoclaving).

Formicin, a novel two-component lanthipeptide (class I bacteriocin), was discovered by screening bacterial isolates from the intestine of Atlantic mackerel (*Scomber scombrus*) for antimicrobial activity. The peptide, produced by a *Bacillus paralicheniformis* isolate, was found to inhibit a range of Gram-positive indicator organisms including clinically relevant pathogens such as *Listeria monocytogenes*, *Staphylococcus aureus*, *Streptococcus mutans*, several species of *Enterococcus* and Clostridia, including *Clostridioides difficile* (Collins et al., 2016). Formicin differs from other two-component lanthipeptides in that the α -peptide contains fewer hydrophobic amino acids and has an overall positive charge of +2, and the β -peptide uniquely contains a negatively charged C-terminus due to the presence of an aspartate residue in the penultimate position (Figures 2a and 2b). The discovery of formicin and its contribution to scientific discovery in Ireland was commemorated by the Irish national postal service (An Post) together with Science Foundation Ireland by the release of a special issue stamp in 2018 (Figure 2c).

Novel bacteriocin-like molecules derived from bacilli of the marine fish gut microbiome have also been reported, such as CAMT2 (An et al., 2015) and BpSI14 (Saidumohamed et al., 2021), produced by *Bacillus amyloliquefaciens* ZJHD-06 and *Bacillus safensis*, respectively. Both of these molecules exhibited a broad spectrum of activity, targeting both Gram-negative and Gram-positive strains and even fungi. Elucidation of their chemical structures found that they were more similar to eukaryotic AMPs than to bacterial peptides.

Recent studies combining *in vitro* and genomic screening have also highlighted the antimicrobial potential of novel Bacilli isolated from fish (Emam and Dunlap, 2020, Yi et al., 2018). For example, *Bacillus* sp. GFP-2, a novel strain of *Bacillus velezensis*, was isolated from intestinal contents of whitespotted bamboo sharks (*Chiloscyllium plagiosum*) and found to inhibit the growth of *Bacillus subtilis* and *Escherichia coli*. Analysis of the genome sequence of strain GFP-2 found that it encoded genes for various AMPs, such as LCI and glyceraldehyde-3-phosphate dehydrogenase (GAPDH) -related peptides, various antimicrobial secondary metabolites including difficidin, bacilysin, macrolactin and fengycin, as well as lanthipeptides (Wu et al., 2018).

1.3.3. Actinobacteria

Actinobacteria are abundant in the marine environment, and those associated with higher organisms such as sponges, molluscs and fish are known to produce a wide variety of bioactive compounds with diverse biological activities (Valliappan et al., 2014). One of the most significant antimicrobials from a marine-derived actinomycete is anthracimycin – an antibiotic with potent activity against *Bacillus anthracis*, the pathogen responsible for causing anthrax. Anthracimycin was discovered by screening *Streptomyces* isolates from marine sediment (Jang et al., 2013). Indeed, *Streptomyces* are among the most frequently identified producers of bioactive compounds from marine environments, particularly those found in association with higher organisms (Valliappan et al., 2014). Several recent studies have screened Actinobacteria from marine fish microbiomes for antimicrobial activity: Vignesh et al. (2019) isolated Actinobacteria from marine fish gut samples that had antimicrobial activity against *Salmonella enterica*, *S. aureus* and *E. coli*. Furthermore, one of the isolates, a *Streptomyces* strain, also demonstrated anti-fungal and quorum-sensing inhibitory activity. More recently, Vadivel et al. (2021) isolated antimicrobial and anti-quorum sensing *Streptomyces* from marine fish gut samples. Several studies have highlighted that the key to identifying novel bioactive and antimicrobial compounds may now be in the “rare” Actinobacteria – those that are not as easily cultivated. These include genera such as *Salinispora*, *Actinomadura*, *Microbacterium* and *Micrococcus* (Valliappan et al., 2014, Subramani and Sipkema, 2019). Sanchez et al. (2012) identified various Actinobacteria, including *Rhodococcus*, *Microbacterium* and *Micromonospora* species (which would be considered part of the “rare” Actinobacteria group) from marine fish digestive tract samples and detected antimicrobial activity against pathogenic *Vibrio* spp., *B. subtilis*, *S. aureus* and *E. faecium*. Though these studies didn’t identify specific antimicrobial molecules, they highlight the biodiscovery and antimicrobial potential of marine fish gut-derived Actinobacteria.

1.4. ‘Extreme’ Marine Environment Impact on Microbial Products

Another interesting aspect of the marine environment is the spectrum of ‘extremes’ to which it is subjected, for example, fluctuating temperatures (from hydrothermal

vents to near-freezing deep waters), sunlight exposure (or complete lack thereof), salinity and increasing hydrostatic pressure with depth (approx. one atmosphere every 10m). Such conditions have effects on the molecular structure and functionality of proteins found in these niches. Cold-adapted enzymes, for example, may require more flexible intramolecular bonding (such as electrostatic and hydrophobic interactions) as a result of the effect of low temperature on entropy. Such 'flexible' interactions can operate at a relatively reduced energy cost. By comparison, mesophilic enzymes require more thermostable intramolecular bonding (Gerday et al., 1997).

However, the challenges of protein production in the deep sea are different. In short, these effects include compression of the protein and consequently a reduction in protein volume. Secondly, high pressure can lead to protein unfolding as a result of the penetration of water into the protein structure (Gerday et al., 1997). High rigidity and stability are therefore required in order to negate the denaturing effects of high pressure. Adaptations include the production of piezolytes, cold-shock proteins, certain transcription factors and chaperones, as well as changes within protein structures themselves (Mykityczuk et al., 2012, Martin et al., 2002). However, there are reports that deep-sea proteins are highly active under, if not at least tolerant to, high pressure. Lactate dehydrogenase (LDH) from the deep-sea fish, *Coryphaenoides armatus*, for example, was shown to be more tolerant to high temperature and high pressure than a homologous LDH from a shallow-water fish species. Sequence analysis of the LDH proteins found differences in over 20 amino acid residues in the deep-sea LDH that were suggested to enhance protein stability and enzyme activity (Brindley et al., 2008). Studies of various other proteins from bathypelagic and abyssal fish species have reported that this tolerance to high pressure is primarily achieved through changes in protein primary structure, mainly amino acid substitutions which result in less flexibility sites that are consequently more resistant to compression (i.e. changes in volume) and can therefore maintain protein functionality (Porter et al., 2016, Morita, 2003). Lemaire et al. (2018) analysed hydrophobic ligand binding proteins from an abyssal fish species and after aligning the protein sequences with their non-deep-sea homologs they suggested that post-translational modification, particularly glycosylation of ligand binding domains, is also

important for high-pressure tolerance in deep-sea proteins. There is also strong evidence that osmolytes, in particular Trimethylamine N-oxide (TMAO), are crucial to maintaining protein integrity at great depths (Yancey et al., 2002, Somero, 2003). Martin et al. (2002) found a *Photobacterium* species that produced a novel class of osmolytes, termed “piezolytes”, the intracellular concentrations of which responded to changes in hydrostatic and osmotic pressures.

In the context of bacteriocins from deep-sea bacteria, they may possess unique adaptations in their transcription, structure, and functionality as a consequence of the deep-sea environmental conditions, such as resistance to high hydrostatic pressure and low temperatures.

1.5. Applications of Marine Fish -Derived Bacteriocins

One of the applications for marine-derived bacteriocins is as an alternative to antibiotics in the treatment of fish diseases within the aquaculture industry (Desriac et al., 2010, Sahoo et al., 2016). Fish and fish products are significant, globally traded commodities, with millions of tonnes consumed each year, either directly as food, as fishmeal or for fish oil. Outbreaks of disease in commercial fish significantly impact global supply, economics and human health (FAO, 2020).

Many studies have also explored the application of bacteriocins from marine fish microbiomes, or the producing strains, in biopreservation of food products, such as seafood products (Schelegueda et al., 2015, Duffes et al., 1999) and chicken (Li et al., 2023). Schelegueda et al. (2015), for example, assessed the potential use of mundticin KS from *E. mundtii* Tw56 in food biopreservation. The bacteriocins exhibited antimicrobial activity against *Listeria innocua* over a range of pH (2.0–10.0), and after heat treatment (up to 121 °C x 15 min). Furthermore, CFS from the producing strain retained full activity after storage at –30 °C for 1 year, –indicating its potential for use in frozen foods. Duffes et al. (1999) studied the anti-*Listeria* activity of the bacteriocinogenic strains, *C. piscicola* V1 and *C. divergens* V41, in vacuum-packed cold smoked salmon over a period of up to 4 weeks. In co-culture assays at 8 °C, both strains reduced the viability of *L. monocytogenes* by 5–7-fold (log CFU/g) compared to *L. monocytogenes* alone. The inhibitory effects were even greater at 4 °C, with *C. piscicola* V1, for example, reducing the counts of *L. monocytogenes* to less

than 10 CFU/g after 4 weeks. Crude bacteriocin extracts from the *Carnobacterium* strains also demonstrated effective anti-*Listeria* activity, particularly at 4 °C whereby *L. monocytogenes* was below detectable levels after 1 week. Significantly, the crude extracts were more effective at inhibiting *Listeria* than nisin under these conditions. Bacteriocinogenic isolates from marine fish are generating great interest in the aquaculture industry for use as probiotics. For example, Shastry et al. (2021) demonstrated that *Enterococcus lactis* RS5 (Table 3) exhibited not only bacteriocin production but also resistance to bile salts, low pH (53 % viability at pH 1.5) and protease digestion - key attributes of probiotic strains. Nguyen et al. (2017) explored the probiotic potential of the nisin Z-producing *L. lactis* WFLU12, originally isolated from the gut of olive flounder (Heo et al., 2012). They reported that fish on a diet supplemented with the antimicrobial strain, WFLU12 (approx. 10⁹ CFU/g), exhibited increased growth and were more resistant to infection by *Streptococcus parauberis*, compared to the control group.

There is also the potential for use as treatments for infection in humans; indeed, many of the bacteriocins listed in Table 3 are capable of targeting human pathogens. Formicin, for example, exhibited effective activity against several clinically relevant species including *C. difficile* (Collins et al., 2016). This is particularly important in cases of antimicrobial resistance whereby treatment with conventional antibiotics has become less effective and novel alternatives are sought.

1.6. Challenges, Metagenomics and Future Prospects

One of the (first) challenges in searching for novel bacteriocins, and other useful metabolites, from novel marine sources is the cultivation of the producing microorganisms. Without the correct conditions, potentially novel marine microorganisms, and thus their metabolites, may be lost when attempting to culture *in vitro* (Deming and Baross, 2000). Environmental pressure, temperature and dissolved oxygen concentration are just a few factors to be taken into account as they can all affect the diversity of isolated microorganisms and the metabolites they produce (Choi et al., 2015, López et al., 2012, Nakayama et al., 1994, Zeng et al., 2004). Often, modified or selective media is required, and long incubation times of days, or even weeks, for antimicrobial production to occur. Sanchez et al. (2012), for

example, implemented the use of selective media for the isolation of bioactive marine Actinobacteria from fish. Vadivel et al. (2021) also explored optimisation of antimicrobial production in *Streptomyces maritimus* SQA4 (from squid) using the “one factor at a time” method, whereby one factor or variable in the cultivation step is modified at a time. They observed that substituting carbon, nitrogen and salt sources, as well as altering pH, had varying effects on antimicrobial activity. Furthermore, production of the bioactive metabolite was detected when the strain was cultivated on solid medium after two days but was not detected when in broth culture until after 10 days of incubation. The growth of some marine bacteria may also be dependent on the presence of specific signalling molecules found in their natural environment (Nichols et al., 2008).

Advances in genome-mining methods for bacteriocins and other antimicrobial molecules have already been extensively reviewed (Zhong et al., 2020). Genomic DNA screening has been used to great effect by several groups in the identification of bacteriocin genes. Genome sequence analysis, using BAGEL3 bacteriocin mining software, allowed for the identification of the formicin biosynthetic operon, from which the masses of the core bacteriocin peptides could be predicted. This was a critical step in the purification and characterisation of formicin, as the peptide masses found using colony mass spectrometry did not match any previously characterised bacteriocins (Collins et al., 2016).

As mentioned above, culture-independent and metagenomic methods are being used to great effect in characterising the taxonomic diversity of the marine fish gut microbiome. Several studies have also employed the use of metagenomic sequencing to characterise functionality, the presence of antimicrobial resistance genes and bioactive metabolite genes in marine fish microbiota (Johnny et al., 2021, Collins et al., 2021, Stevenson et al., 2022), including those from rarer deep-sea fish species (Figure 3). The culture-independent nature of metagenomics allows for a more comprehensive understanding of these gut communities and removes the bias of sampling only culturable microbiota. Few studies, however, have used metagenomics approaches for screening marine fish microbiomes for bacteriocins specifically. Yi et al. (2019) used metagenomic sequencing to study functional dynamics in the gut microbiota of several aquatic animals, including marine and freshwater fish. By

comparative gene sequence analysis against multiple protein databases, they identified numerous bacteriocin-associated genes, including secretion systems, and immunity-related membrane transporter systems. However, they identified few core peptide biosynthetic genes, suggesting a role in antimicrobial resistance rather than production by the microbiota.

Whilst *in silico* genomics screening can identify bacteriocin biosynthetic gene clusters, there is still a need for peptide purification but expression in the native host can present another challenge. Another emerging technology is bacteriocin “reincarnation”. This method involves cloning of the biosynthetic operon of an “inactive” antimicrobial into a heterologous host, thereby enabling controlled expression of the antimicrobial gene. Collins et al. (2018) utilised this method to reincarnate pediocin-like bacteriocin structural genes that had been identified during *in silico* analysis of the *Lactobacillus* pangenome, yet these strains did not exhibit *in vitro* antimicrobial activity. Ten pediocin-like bacteriocin genes were cloned with a (pediocin) signal sequence and heterologously expressed in *E. coli* and *Lactobacillus paracasei* and displayed antimicrobial activity against *L. innocua*. This study demonstrates the potential for such methods in the expression of such “inactive” bacteriocins, that may be identified in genomic sequences of marine-fish-derived bacteria, and allow for their heterologous expression under more familiar, terrestrial, conditions.

1.7. Conclusions

With the increasing demand for novel antimicrobial therapeutics to combat antimicrobial resistance, there is a pressing need to source novel bioactive compounds. The marine environment is vast, with a multitude of factors affecting the diversity of marine life and their biomolecules. The marine fish gut microbiome is no exception, and it is emerging as a productive source of antimicrobial molecules, such as bacteriocins. Key aspects of marine life, such as high salinity, hydrostatic pressure and a range of environmental temperatures all have an impact on the structural and functional diversity of marine antimicrobial molecules, including bacteriocins, often distinguishing them from terrestrial homologs. Characterising the true diversity of microorganisms in marine fish gut microbiomes and the bioactive molecules they produce still presents challenges but may be key in determining the true potential of this niche. Recent applications of culture-independent technologies, such as metagenomics and next-generation sequencing, have begun to bridge this knowledge gap and are already being used to great effect to resolve the taxonomic diversity of the fish gut microbiome. Bacteriocins from the marine fish gut can be used to combat disease, either directly or indirectly as products of probiotic strains, and for the biopreservation of food products. The marine fish gut microbiome is a potential treasure trove of novel bacteriocins and other antimicrobials but limitations in accessibility and technology have meant its true bioprospective potential has yet to be truly explored.

1.8. References

- ANTIMICROBIAL RESISTANCE COLLABORATORS 2022. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *The Lancet*, 399, 629-655.
- APPROVED MARINE DRUGS [Online]. *Marine Pharmacology*. Available at: <https://www.marinepharmacology.org/approved> (Accessed on August 14, 2023).
- AN, J., ZHU, W., LIU, Y., ZHANG, X., SUN, L., HONG, P., WANG, Y., XU, C., XU, D. & LIU, H. 2015. Purification and characterization of a novel bacteriocin CAMT2 produced by *Bacillus amyloliquefaciens* isolated from marine fish *Epinephelus areolatus*. *Food Control*, 51, 278-282.
- ANDLID, T., JUÁREZ, R.-V. & GUSTAFSSON, L. 1995. Yeast colonizing the intestine of rainbow trout (*Salmo gairdneri*) and turbot (*Scophthalmus maximus*). *Microbial Ecology*, 30, 321-334.
- BHUGALOO-VIAL, P., DOUSSET, X., METIVIER, A., SOROKINE, O., ANGLADE, P., BOYAVAL, P. & MARION, D. 1996. Purification and amino acid sequences of piscicocins V1a and V1b, two class IIa bacteriocins secreted by *Carnobacterium piscicola* V1 that display significantly different levels of specific inhibitory activity. *Applied and Environmental Microbiology*, 62, 4410-4416.
- BINDIYA, E. S., TINA, K. J., RAGHUL, S. S. & BHAT, S. G. 2015. Characterization of deep sea fish gut bacteria with antagonistic potential, from *Centroscyllium fabricii* (Deep Sea Shark). *Probiotics and Antimicrobial Proteins*, 7, 157-63.
- BINDIYA, E. S., TINA, K. J., SASIDHARAN, R. S. & BHAT, S. G. 2019. BaCf3: highly thermostable bacteriocin from *Bacillus amyloliquefaciens* BTSS3 antagonistic on food-borne pathogens. *3 Biotech*, 9, 136.
- BREUKINK, E., VAN KRAAIJ, C., DEMEL, R. A., SIEZEN, R. J., KUIPERS, O. P. & DE KRUIJFF, B. 1997. The C-terminal region of nisin is responsible for the initial interaction of nisin with the target membrane. *Biochemistry*, 36, 6968-6976.
- BRINDLEY, A. A., PICKERSGILL, R. W., PARTRIDGE, J. C., DUNSTAN, D. J., HUNT, D. M. & WARREN, M. J. 2008. Enzyme sequence and its relationship to hyperbaric stability of artificial and natural fish lactate dehydrogenases. *PloS one*, 3, e2042-e2042.
- BURTSEVA, O., KUBLANOVSKAYA, A., FEDORENKO, T., LOBAKOVA, E. & CHEKANOV, K. 2021. Gut microbiome of the White Sea fish revealed by 16S rRNA metabarcoding. *Aquaculture*, 533, 736175.
- BUTT, R. L. & VOLKOFF, H. 2019. Gut microbiota and energy homeostasis in Fish. *Frontiers in endocrinology*, 10, 9-9.

- CHOI, EUN J., NAM, S.-J., PAUL, L., BEATTY, D., KAUFFMAN, C. A., JENSEN, P. R. & FENICAL, W. 2015. Previously uncultured marine bacteria linked to novel alkaloid production. *Chemistry & Biology*, 22, 1270-1279.
- COLLINS, F. W. J. 2019. *An investigation into antimicrobial production in the Lactobacillus genus and the fish microbiome*. PhD Thesis, University College Cork, Cork, Ireland.
- COLLINS, F. W. J., MESA-PEREIRA, B., O'CONNOR, P. M., REA, M. C., HILL, C. & ROSS, R. P. 2018. Reincarnation of bacteriocins from the *Lactobacillus* pangenomic graveyard. *Frontiers in Microbiology*, 9, 1298.
- COLLINS, F. W. J., O'CONNOR, P. M., O'SULLIVAN, O., REA, M. C., HILL, C. & ROSS, R. P. 2016. Formicin - a novel broad-spectrum two-component lantibiotic produced by *Bacillus paralicheniformis* APC 1576. *Microbiology*, 162, 1662-1671.
- COLLINS, F. W. J., WALSH, C. J., GOMEZ-SALA, B., GUIJARRO-GARCÍA, E., STOKES, D., JAKOBSDÓTTIR, K. B., KRISTJÁNSSON, K., BURNS, F., COTTER, P. D., REA, M. C., HILL, C. & ROSS, R. P. 2021. The microbiome of deep-sea fish reveals new microbial species and a sparsity of antibiotic resistance genes. *Gut Microbes*, 13, 1-13.
- COTTER, P. D., ROSS, R. P. & HILL, C. 2013. Bacteriocins — a viable alternative to antibiotics? *Nature Reviews Microbiology*, 11, 95-105.
- DEMING, J. W. & BAROSS, J. A. 2000. Survival, dormancy, and nonculturable cells in extreme deep-sea environments. In: *Nonculturable Microorganisms in the Environment*. COLWELL, R. R. & GRIMES, D. J. (eds.). Boston, MA: Springer USA. 2000; pp. 147–197.
- DESRIAC, F., DEFER, D., BOURGOUGNON, N., BRILLET, B., LE CHEVALIER, P. & FLEURY, Y. 2010. Bacteriocin as weapons in the marine animal-associated bacteria warfare: inventory and potential applications as an aquaculture probiotic. *Marine Drugs*, 8, 1153-1177.
- DUFFES, F., LEROI, F., BOYAVAL, P. & DOUSSET, X. 1999. Inhibition of *Listeria monocytogenes* by *Carnobacterium* spp. strains in a simulated cold smoked fish system stored at 4°C. *International Journal of Food Microbiology*, 47, 33-42.
- EGERTON, S., CULLOTY, S., WHOOLEY, J., STANTON, C. & ROSS, R. P. 2018. The Gut Microbiota of Marine Fish. *Frontiers in microbiology*, 9, 873-873.
- EMAM, A. M. & DUNLAP, C. A. 2020. Genomic and phenotypic characterization of *Bacillus velezensis* AMB-y1; a potential probiotic to control pathogens in aquaculture. *Antonie van Leeuwenhoek*, 113, 2041-2052.
- FAO 2020. *The State of World Fisheries and Aquaculture 2020*. Sustainability in Action; Report No.: 978-92-5-132692-3 Contract No.: 63; FAO: Rome, Italy, 2020.

- FRIERI, M., KUMAR, K. & BOUTIN, A. 2017. Antibiotic resistance. *Journal of Infection and Public Health*, 10, 369-378.
- GEOGHEGAN, J. L., DI GIALONARDO, F., WILLE, M., ORTIZ-BAEZ, A. S., COSTA, V. A., GHALY, T., MIFSUD, J. C., TURNBULL, O. M., BELLWOOD, D. R. & WILLIAMSON, J. E. 2021. Virome composition in marine fish revealed by meta-transcriptomics. *Virus Evolution*, 7, veab005.
- GERDAY, C., AITTALEB, M., ARPIGNY, J. L., BAISE, E., CHESSA, J.-P., GARSOUX, G., PETRESCU, I. & FELLER, G. 1997. Psychrophilic enzymes: a thermodynamic challenge. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology*, 1342, 119-131.
- GHANBARI, M., KNEIFEL, W. & DOMIG, K. J. 2015. A new view of the fish gut microbiome: Advances from next-generation sequencing. *Aquaculture*, 448, 464-475.
- GREENAWAY, S. F., SULLIVAN, K. D., UMFRESS, S. H., BEITTEL, A. B. & WAGNER, K. D. 2021. Revised depth of the Challenger Deep from submersible transects; including a general method for precise, pressure-derived depths in the ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, 178, 103644.
- HALLORAN, K. & UNDERWOOD, M. A. 2019. Probiotic mechanisms of action. *Early Human Development*, 135, 58-65.
- HEO, W.-S., KIM, E.-Y., KIM, Y.-R., HOSSAIN, M. T. & KONG, I.-S. 2012. Salt effect of nisin Z isolated from a marine fish on the growth inhibition of *Streptococcus iniae*, a pathogen of streptococcosis. *Biotechnology Letters*, 34, 315-320.
- HUANG, Q., SHAM, R. C., DENG, Y., MAO, Y., WANG, C., ZHANG, T. & LEUNG, K. M. 2020. Diversity of gut microbiomes in marine fishes is shaped by host-related factors. *Molecular ecology*, 29, 5019-5034.
- JANG, K. H., NAM, S.-J., LOCKE, J. B., KAUFFMAN, C. A., BEATTY, D. S., PAUL, L. A. & FENICAL, W. 2013. Anthracimycin, a potent anthrax antibiotic from a marine-derived Actinomycete. *Angewandte Chemie International Edition*, 52, 7822-7824.
- JOHNY, T. K., PUTHUSSERI, R. M. & BHAT, S. G. 2021. Metagenomic landscape of taxonomy, metabolic potential and resistome of *Sardinella longiceps* gut microbiome. *Archives of Microbiology*, 204, 87.
- JOHNY, T. K., SAIDUMOHAMED, B. E., SASIDHARAN, R. S. & BHAT, S. G. 2018a. Inferences of gut bacterial diversity from next-generation sequencing of 16S rDNA in deep sea blind ray - *Benthobatis moresbyi*. *Ecological Genetics and Genomics*, 9, 1-6.
- JOHNY, T. K., SAIDUMOHAMED, B. E., SASIDHARAN, R. S. & BHAT, S. G. 2018b. Metabarcoding data of bacterial diversity of the deep sea shark, *Centroscyllium fabricii*. *Data in Brief*, 21, 1029-1032.

- JONES, J., DIBATTISTA, J. D., STAT, M., BUNCE, M., BOYCE, M. C., FAIRCLOUGH, D. V., TRAVERS, M. J. & HUGGETT, M. J. 2018. The microbiome of the gastrointestinal tract of a range-shifting marine herbivorous fish. *Frontiers in Microbiology*, 9, 2000.
- KORMAS, K., NIKOULI, E., KOUSTENI, V. & DAMALAS, D. 2023. Midgut bacterial microbiota of 12 fish species from a marine protected area in the Aegean Sea (Greece). *Microbial Ecology*, 86, 1405-1415.
- LEMAIRE, B., KARCHNER, S. I., GOLDSTONE, J. V., LAMB, D. C., DRAZEN, J. C., REES, J. F., HAHN, M. E. & STEGEMAN, J. J. 2018. Molecular adaptation to high pressure in cytochrome P450 1A and aryl hydrocarbon receptor systems of the deep-sea fish *Coryphaenoides armatus*. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, 1866, 155-165.
- LI, Q., CHEN, Q., WU, Y., CHEN, Z., LIU, Y., FANG, Z. & DENG, Q. 2023. Purification, characterization and structural identification of a novel bacteriocin produced by marine original *Enterococcus durans* YQ-6, and its inhibition of *Listeria monocytogenes*. *LWT*, 173, 114329.
- LÓPEZ, R., MONTEÓN, V., CHAN, E., MONTEJO, R. & CHAN, M. 2012. Oxygen limitation favors the production of protein with antimicrobial activity in *Pseudoalteromonas* sp. *Brazilian journal of microbiology : [publication of the Brazilian Society for Microbiology]*, 43, 1206-1212.
- MARTIN, D., BARTLETT, D. H. & ROBERTS, M. F. 2002. Solute accumulation in the deep-sea bacterium *Photobacterium profundum*. *Extremophiles*, 6, 507-514.
- MAYER, A. M. S., GUERRERO, A. J., RODRÍGUEZ, A. D., TAGLIALATELA-SCAFATI, O., NAKAMURA, F. & FUSETANI, N. 2020. Marine Pharmacology in 2014–2015: Marine compounds with antibacterial, antidiabetic, antifungal, anti-inflammatory, antiprotozoal, antituberculosis, antiviral, and anthelmintic activities; affecting the immune and nervous systems, and other miscellaneous mechanisms of action. *Marine Drugs*, 18, 5.
- MAYER, A. M. S., GUERRERO, A. J., RODRÍGUEZ, A. D., TAGLIALATELA-SCAFATI, O., NAKAMURA, F. & FUSETANI, N. 2021. Marine Pharmacology in 2016–2017: Marine compounds with antibacterial, antidiabetic, antifungal, anti-inflammatory, antiprotozoal, antituberculosis and antiviral activities; affecting the immune and nervous systems, and other miscellaneous mechanisms of action. *Marine Drugs*, 19, 49.
- MAYER, A. M. S., PIERCE, M. L., HOWE, K., RODRÍGUEZ, A. D., TAGLIALATELA-SCAFATI, O., NAKAMURA, F. & FUSETANI, N. 2022. Marine pharmacology in 2018: Marine compounds with antibacterial, antidiabetic, antifungal, anti-inflammatory, antiprotozoal, antituberculosis and antiviral activities; affecting the immune and nervous systems, and other miscellaneous mechanisms of action. *Pharmacological Research*, 183, 106391.

- METIVIER, A., PILET, M.-F., DOUSSET, X., SOROKINE, O., ANGLADE, P., ZAGOREC, M., PIARD, J.-C., MARLON, D., CENATIEMPO, Y. & FREMAUX, C. 1998. Divercin V41, a new bacteriocin with two disulphide bonds produced by *Carnobacterium divergens* V41: primary structure and genomic organization. *Microbiology*, 144, 2837-2844.
- MONTALBÁN-LÓPEZ, M., SCOTT, T. A., RAMESH, S., RAHMAN, I. R., VAN HEEL, A. J., VIEL, J. H., BANDARIAN, V., DITTMANN, E., GENILLOUD, O., GOTO, Y., GRANDE BURGOS, M. J., HILL, C., KIM, S., KOEHNKE, J., LATHAM, J. A., LINK, A. J., MARTÍNEZ, B., NAIR, S. K., NICOLET, Y., REBUFFAT, S., SAHL, H.-G., SAREEN, D., SCHMIDT, E. W., SCHMITT, L., SEVERINOV, K., SÜSSMUTH, R. D., TRUMAN, A. W., WANG, H., WENG, J.-K., VAN WEZEL, G. P., ZHANG, Q., ZHONG, J., PIEL, J., MITCHELL, D. A., KUIPERS, O. P. & VAN DER DONK, W. A. 2021. New developments in RiPP discovery, enzymology and engineering. *Natural Product Reports*, 38, 130-239.
- MORITA, T. 2003. Structure-based analysis of high pressure adaptation of α -actin. *Journal of Biological Chemistry*, 278, 28060-28066.
- MULDERS, J. W., BOERRIGTER, I. J., ROLLEMA, H. S., SIEZEN, R. J. & DE VOS, W. M. 1991. Identification and characterization of the lantibiotic nisin Z, a natural nisin variant. *European Journal of Biochemistry*, 201, 581-584.
- MYKYTCZUK, N. C. S., WILHELM, R. C. & WHYTE, L. G. 2012. *Planococcus halocryophilus* sp. nov., an extreme sub-zero species from high Arctic permafrost. *International Journal of Systematic and Evolutionary Microbiology*, 62, 1937-1944.
- NAKAYAMA, A., YANO, Y. & YOSHIDA, K. 1994. New method for isolating barophiles from intestinal contents of deep-sea fishes retrieved from the abyssal zone. *Applied and Environmental Microbiology*, 60, 4210-4212.
- NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION, U. S. D. O. C. *Layers of the Ocean* [Online]. Available at: <https://www.noaa.gov/jetstream/ocean/layers-of-ocean> (Accessed March 28, 2023).
- NGUYEN, T. L., PARK, C.-I. & KIM, D.-H. 2017. Improved growth rate and disease resistance in olive flounder, *Paralichthys olivaceus*, by probiotic *Lactococcus lactis* WFLU12 isolated from wild marine fish. *Aquaculture*, 471, 113-120.
- NICHOLS, D., LEWIS, K., ORJALA, J., MO, S., ORTENBERG, R., CONNOR, P., ZHAO, C., VOUIROS, P., KAEBERLEIN, T. & EPSTEIN, S. S. 2008. Short peptide induces an “uncultivable” microorganism to grow in vitro. *Applied and Environmental Microbiology*, 74, 4889.
- NISSEN-MEYER, J., ROGNE, P., OPPEGARD, C., HAUGEN, S. H. & KRISTIANSEN, E. P. 2009. Structure-function relationships of the non-lanthionine-containing peptide (class ii) bacteriocins produced by Gram-positive bacteria. *Current Pharmaceutical Biotechnology*, 10, 19-37.

- OHWADA, K., TABOR, P. S. & COLWELL, R. R. 1980. Species composition and barotolerance of gut microflora of deep-sea benthic macrofauna collected at various depths in the atlantic ocean. *Applied and Environmental Microbiology*, 40, 746-55.
- OU, W., YU, G., ZHANG, Y. & MAI, K. 2021. Recent progress in the understanding of the gut microbiota of marine fishes. *Marine Life Science & Technology*, 3, 434-448.
- PILET, M.-F., DOUSSET, X., BARRÉ, R., NOVEL, G., DESMAZEAUD, M. & PIARD, J.-C. 1995. Evidence for two bacteriocins produced by *Carnobacterium piscicola* and *Carnobacterium divergens* isolated from fish and active against *Listeria monocytogenes*. *Journal of Food Protection*, 58, 256-262.
- PORTER, M. L., ROBERTS, N. W. & PARTRIDGE, J. C. 2016. Evolution under pressure and the adaptation of visual pigment compressibility in deep-sea environments. *Molecular Phylogenetics and Evolution*, 105, 160-165.
- RINGØ, E., HOSEINIFAR, S. H., GHOSH, K., DOAN, H. V., BECK, B. R. & SONG, S. K. 2018. Lactic acid bacteria in finfish—An update. *Frontiers in microbiology*, 9, 1818.
- ROESELERS, G., MITTGE, E. K., STEPHENS, W. Z., PARICHY, D. M., CAVANAUGH, C. M., GUILLEMIN, K. & RAWLS, J. F. 2011. Evidence for a core gut microbiota in the zebrafish. *The ISME journal*, 5, 1595-1608.
- ROMERO, J., RINGØ, E. & MERRIFIELD, D. 2014. The gut microbiota of fish. *Aquaculture Nutrition: Gut Health, Probiotics and Prebiotics* 2014.75-100. <https://doi.org/10.1002/9781118897263.ch4>
- SAHOO, T. K., JENA, P. K., PATEL, A. K. & SESHADRI, S. 2016. Bacteriocins and their applications for the treatment of bacterial diseases in aquaculture: a review. *Aquaculture Research*, 47, 1013-1027.
- SAIDUMOHAMED, B. E., BABURAJ, A. P., JOHNY, T. K., SHEELA, U. B., SREERANGANATHAN, M. & BHAT, S. G. 2021. A magainin-2 like bacteriocin BpSl14 with anticancer action from fish gut *Bacillus safensis* SDG14. *Analytical Biochemistry*, 627, 114261.
- SAIDUMOHAMED, B. E., JOHNY, T. K., RAVEENDRAN, A. T., SHEELA, U. B., SREERANGANATHAN, M., SASIDHARAN, R. S. & BHAT, S. G. 2022. 3D structure elucidation and appraisal of mode of action of a bacteriocin BaCf3 with anticancer potential produced by Marine *Bacillus amyloliquefaciens* BTSS3. *Re:GEN Open*, 2, 45-56.
- SANCHEZ, L. M., WONG, W. R., RIENER, R. M., SCHULZE, C. J. & LININGTON, R. G. 2012. Examining the Fish Microbiome: Vertebrate-derived bacteria as an environmental niche for the discovery of unique marine natural products. *PLoS One*, 7, e35398.
- SCHELEGUEDA, L. I., VALLEJO, M., GLIEMMO, M. F., MARGUET, E. R. & CAMPOS, C. A. 2015. Synergistic antimicrobial action and potential application for fish

- preservation of a bacteriocin produced by *Enterococcus mundtii* isolated from *Odontesthes platensis*. *LWT - Food Science and Technology*, 64, 794-801.
- SEQUEIROS, C., GARCÉS, M. E., VALLEJO, M., MARGUET, E. R. & OLIVERA, N. L. 2015. Potential aquaculture probiont *Lactococcus lactis* TW34 produces nisin Z and inhibits the fish pathogen *Lactococcus garvieae*. *Archives of Microbiology*, 197, 449-458.
- SHASTRY, R. P., ARUNRENGANATHAN, R. R. & RAI, V. R. 2021. Characterization of probiotic *Enterococcus lactis* RS5 and purification of antibiofilm enterocin. *Biocatalysis and Agricultural Biotechnology*, 31, 101897.
- SINGH, P. K., CHITTPURNA, ASHISH, SHARMA, V., PATIL, P. B. & KORPOLE, S. 2012. Identification, purification and characterization of laterosporulin, a novel bacteriocin produced by *Brevibacillus* sp. strain GI-9. *PLoS One*, 7, e31498.
- SOMERO, G. N. 1992. Biochemical ecology of deep-sea animals. *Experientia*, 48, 537-543.
- SOMERO, G. N. 2003. Protein adaptations to temperature and pressure: complementary roles of adaptive changes in amino acid sequence and internal milieu. *Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology*, 136, 577-591.
- SONG, W., LI, L., HUANG, H., JIANG, K., ZHANG, F., CHEN, X., ZHAO, M. & MA, L. 2016. The Gut Microbial Community of Antarctic Fish Detected by 16S rRNA gene sequence analysis. *BioMed research international*, 2016, 3241529-3241529.
- STAR, B., HAVERKAMP, T. H., JENTOFT, S. & JAKOBSEN, K. S. 2013. Next generation sequencing shows high variation of the intestinal microbial species composition in Atlantic cod caught at a single location. *BMC Microbiology*, 13, 248.
- STEVENSON, S. J., LEE, K. C., HANDLEY, K. M., ANGERT, E. R., WHITE, W. L. & CLEMENTS, K. D. 2022. Substrate degradation pathways, conserved functions and community composition of the hindgut microbiota in the herbivorous marine fish *Kyphosus sydneyanus*. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 272, 111283.
- SUBRAMANI, R. & SIPKEMA, D. 2019. Marine Rare Actinomycetes: A promising source of structurally diverse and unique novel natural products. *Marine Drugs*, 17, 249.
- SULLAM, K. E., ESSINGER, S. D., LOZUPONE, C. A., O'CONNOR, M. P., ROSEN, G. L., KNIGHT, R., KILHAM, S. S. & RUSSELL, J. A. 2012. Environmental and ecological factors that shape the gut bacterial communities of fish: a meta-analysis. *Molecular ecology*, 21, 3363-3378.
- SUTTLE, C. A. 2005. Viruses in the sea. *Nature*, 437, 356-361.

- TROUSSELLIER, M., ESCALAS, A., BOUVIER, T. & MOUILLOT, D. 2017. Sustaining Rare Marine Microorganisms: Macroorganisms As Repositories and Dispersal Agents of Microbial Diversity. *Frontiers in microbiology*, 8, 947-947.
- VADIVEL, M., VENUGOPAL, G., ANGAMUTHU, V., MANIKKAM, R., JOSEPH, J. & ARUNI, W. Exploration of fish gut associated actinobacteria for its anti-microbial and anti-quorum sensing properties. International seminar on promoting local resources for sustainable agriculture and development (ISPLRSAD 2020), 2021. Atlantis Press, 126-139.
- VALLIAPPAN, K., SUN, W. & LI, Z. 2014. Marine actinobacteria associated with marine organisms and their potentials in producing pharmaceutical natural products. *Applied Microbiology and Biotechnology*, 98, 7365-7377.
- VAN DER MAAREL, M. J. E. C., SPRENGER, W., HAANSTRA, R. & FORNEY, L. J. 1999. Detection of methanogenic archaea in seawater particles and the digestive tract of a marine fish species. *FEMS Microbiology Letters*, 173, 189-194.
- VIGNESH, A., AISWARYA, S., GOPIKRISHNAN, V. & RADHAKRISHNAN, M. 2019. Bioactive potential of actinobacteria isolated from the gut of marine fishes *Indian Journal of Geo Marine Sciences*, 48, 1280-1285.
- VUILLEMIN, A., WANKEL, S. D., COSKUN, Ö. K., MAGRITSCH, T., VARGAS, S., ESTES, E. R., SPIVACK, A. J., SMITH, D. C., POCKALNY, R., MURRAY, R. W., D'HONDT, S. & ORSI, W. D. 2019. Archaea dominate oxic seafloor communities over multimillion-year time scales. *Science advances*, 5, eaaw4108.
- WANG, L., RAVICHANDRAN, V., YIN, Y., YIN, J. & ZHANG, Y. 2019. Natural products from mammalian gut microbiota. *Trends in Biotechnology*, 37, 492-504.
- WANKA, K. M., DAMERAU, T., COSTAS, B., KRUEGER, A., SCHULZ, C. & WUERTZ, S. 2018. Isolation and characterization of native probiotics for fish farming. *BMC microbiology*, 18, 119.
- WARD, N. L., STEVEN, B., PENN, K., METHÉ, B. A. & DETRICH, W. H. 2009. Characterization of the intestinal microbiota of two Antarctic notothenioid fish species. *Extremophiles*, 13, 679-685.
- WU, J., XU, G., JIN, Y., SUN, C., ZHOU, L., LIN, G., XU, R., WEI, L., FEI, H., WANG, D., CHEN, J., LV, Z. & LIU, K. 2018. Isolation and characterization of *Bacillus* sp. GFP-2, a novel *Bacillus* strain with antimicrobial activities, from Whitespotted bamboo shark intestine. *AMB Express*, 8, 84.
- YANCEY, P. H., BLAKE, W. R. & CONLEY, J. 2002. Unusual organic osmolytes in deep-sea animals: adaptations to hydrostatic pressure and other perturbants. *Comparative Biochemistry and Physiology Part A: Molecular & Integrative Physiology*, 133, 667-676.
- YANO, Y., NAKAYAMA, A. & YOSHIDA, K. 1995. Population sizes and growth pressure responses of intestinal microfloras of deep-sea fish retrieved from the abyssal zone. *Applied and Environmental Microbiology*, 61, 4480-4483.

- YI, Y., LIANG, L., WANG, Z., AI, P., YOU, X., BIAN, C., SHI, Q. & DONG, B. 2019. A comparative metagenomics study on gastrointestinal microbiota in amphibious mudskippers and other vertebrate animals. *Animals*, 9, 660.
- YI, Y., ZHANG, Z., ZHAO, F., LIU, H., YU, L., ZHA, J. & WANG, G. 2018. Probiotic potential of *Bacillus velezensis* JW: Antimicrobial activity against fish pathogenic bacteria and immune enhancement effects on *Carassius auratus*. *Fish Shellfish Immunology*, 78, 322-330.
- ZENG, X., XIAO, X., WANG, P. & WANG, F. P. 2004. Screening and characterization of psychrotrophic, lipolytic bacteria from deep-sea sediments. *Journal of Microbiology and Biotechnology*, 14, 952-958.
- ZHONG, Z., HE, B., LI, J. & LI, Y.-X. 2020. Challenges and advances in genome mining of ribosomally synthesized and post-translationally modified peptides (RiPPs). *Synthetic and Systems Biotechnology*, 5, 155-172.

1.9. Chapter 1 Figures and Tables

Table 1: The predominant phyla reported in recent culture-independent studies of marine fish gut microbiota. (phyla are listed in order of abundance where possible)

Fish species	Sample	Predominant Phyla	Ref.
<i>Gadus morhua</i> (Atlantic Cod)	Intestinal contents	Pseudomonadota, Bacteroidota, Bacillota	(Star et al., 2013)
<i>Siganus fuscescens</i> (Mottled spinefoot rabbitfish)	Intestinal contents	Pseudomonadota, Bacillota, Bacteroidota, Fusobacteriota, Mycoplasmatota, Cyanobacteriota	(Jones et al., 2018)
Various White Sea (arctic) fish	Posterior intestine	Pseudomonadota, Bacillota, Actinomycetota, Bacteroidota, Mycoplasmatota, Fusobacteriota	(Burtseva et al., 2021)
Various Mediterranean fish	Midgut	Pseudomonadota, Bacillota, Bacteroidota, Actinobacteriota, Patescibacteria, Fusobacteriota, Planctomycetota, and Dependientiae	(Kormas et al., 2023)
Coastal fish of Hong Kong	Gastrointestinal contents	Pseudomonadota, Bacillota, Mycoplasmatota	(Huang et al., 2020)
Various deep-sea fish of Atlantic Ocean	Intestinal contents	Pseudomonadota, Bacteroidota, Bacillota, Actinomycetota, Ascomycota, Basidiomycota, Euryarchaeota, Spirochaetes	(Collins, 2019)
<i>Centroscyllium fabricii</i> (Black dogfish shark)	Gut contents	Actinomycetota, Pseudomonadota, Acidobacteriota (Acidobacteria), Bacillota, Chloroflexota	(Johnny et al., 2018b)
<i>Benthobatis moresbyi</i> (Dark Blind Ray)	Gut contents	Actinomycetota, Pseudomonadota, Acidobacteriota, Chloroflexota, Bacillota	(Johnny et al., 2018a)

Ref. = reference

Table 2: Recent culture-independent studies of marine fish gut microbiome diversity. The most abundant/predominant taxa reported at family and genus level are given.

Fish	Sample	Abundant genera	Abundant Families	Ref.
<i>Centroscyllium fabricii</i> (Black dogfish shark)	Gut contents	<i>Acinetobacter</i> , <i>Thalassobacillus</i> , <i>Alteromonas</i> , <i>Leeuwenhoekiella</i> , <i>Corynebacterium</i> , <i>Pseudonocardia</i> , <i>Pseudomonas</i>	NR	(Johny et al., 2018b)
<i>Benthobatis moresbyi</i> (Dark Blind Ray)	Gut contents	<i>Acinetobacter</i>	<i>Moraxellaceae</i> , <i>Koribacteraceae</i> , <i>Nitrospiraceae</i>	(Johny et al., 2018a)
White Sea (arctic) fish	Posterior intestine	<i>Streptococcus</i> , <i>Sphingomonas</i> , <i>Micrococcus</i> , <i>Chthoniobacter</i> , <i>Pseudomonas</i> , <i>Corynebacterium</i> , <i>Staphylococcus</i> , <i>Acinetobacter</i> , <i>Propionibacterium</i> , <i>Vibrio</i> , <i>Photobacterium</i> , <i>Bacillus</i>	<i>Moraxellaceae</i> , <i>Vibrionaceae</i> , <i>Pseudomonadaceae</i> , <i>Propionibacteriaceae</i> , <i>Corynebacteriaceae</i> , <i>Micrococcaceae</i>	(Burtseva et al., 2021)
Various Mediterranean fish	Midgut	<i>Pseudoalteromonas</i> , <i>Bradyrhizobium</i> , <i>Diaphorobacter</i> , <i>Mycoplasma</i> , <i>Clostridium</i> , <i>Thaumasiovibrio</i> , <i>Microbulbifer</i>	<i>Xanthobacteraceae</i> , <i>Comamonadaceae</i> , <i>Pseudoalteromonadaceae</i> , <i>Clostridiaceae</i> , <i>Vibrionaceae</i> , <i>Propionibacteriaceae</i> , <i>Staphylococcaceae</i> , <i>Mycoplasmataceae</i> , <i>Flavobacteriaceae</i> , and <i>Peptostreptococcaceae</i>	(Kormas et al., 2023)
Various Antarctic fish		<i>Rhodococcus</i> , <i>Thermus</i> , <i>Acinetobacter</i> , <i>Propionibacterium</i> , <i>Streptococcus</i> , and <i>Mycoplasma</i>	NR	(Song et al., 2016)
Coastal fish of Hong Kong	Gastrointestinal contents	<i>Clostridium</i> , <i>Photobacterium</i> , <i>Ralstonia</i> , <i>Acinetobacter</i> , <i>Thermus</i> , <i>Ralstonia</i> ,	NR	(Huang et al., 2020)

NR = not reported, Ref. = reference

Table 3: Bacteriocins, bacteriocin-like inhibitory substances and other AMPs identified from marine fish gut microbiome isolates. Many class IIa bacteriocins have been isolated from this microbial niche, covering a wide spectrum of susceptible organisms.

Molecule	Producer	Host, source	Susceptible organism(s)	Ref.
Class I bacteriocins				
Formicin	<i>Bacillus paralicheniformis</i> APC 1576	<i>Scomber scombrus</i> , intestine	<i>Clostridia</i> spp., <i>Bacillus</i> spp., <i>Listeria</i> spp., <i>Enterococcus</i> spp., <i>Streptococcus mutans</i> , <i>M. luteus</i>	(Collins et al., 2016)
Nisin Z	<i>Lactococcus lactis</i> subsp. <i>lactis</i>	<i>Paralichthys olivaceus</i> , intestine	<i>Streptococcus iniae</i>	(Heo et al., 2012)
Nisin Z	<i>Lactococcus lactis</i> TW34	<i>Odontesthes platensis</i> , intestine	<i>Lactococcus garvieae</i>	(Sequeiros et al., 2015)
Class IIa bacteriocins				
Piscicocins Vla, Vlb	<i>Carnobacterium piscicola</i> V1	salmon/trout, intestine	<i>Listeria</i> spp.	(Pilet et al., 1995)
Divercin V41	<i>Carnobacterium divergens</i> V41	salmon or trout, intestine	<i>Carnobacterium piscicola</i> , <i>Listeria</i> spp.	(Metivier et al., 1998)
Mundticin KS	<i>Enterococcus mundtii</i> Tw56	<i>Odontesthes platensis</i> , intestine	<i>Enterococcus</i> spp., <i>Listeria</i> spp., <i>M. luteus</i> , <i>Pseudomonas aeruginosa</i> , <i>Shewanella putrefaciens</i>	(Schelegueda et al., 2015)
Enterocin R5	<i>Enterococcus lactis</i> RS5	<i>Sillago indica</i> , gut	<i>E. coli</i> , <i>S. enterica</i> Typhimurium, <i>S. aureus</i> , <i>P. aeruginosa</i> , <i>B. subtilis</i> , <i>B. cereus</i> , <i>Proteus vulgaris</i>	(Shastri et al., 2021)
Class IIb bacteriocins				
CAMT6	<i>Enterococcus durans</i> YQ-6	<i>Larimichthys polyactis</i> , NR	<i>S. aureus</i> , <i>Bacillus</i> spp., <i>S. haemolyticus</i> , <i>P. acnes</i> , <i>Salmonella paratyphi</i> , <i>V. parahaemolyticus</i> , <i>P. foulis</i> , <i>E. aerogenes</i> , <i>Fusarium sylvaticum</i> , <i>Aspergillus fumigatus</i>	(Li et al., 2023)
Other AMPs /bacteriocin-like inhibitory substances				
BaCf3	<i>Bacillus amyloliquefaciens</i> BTSS3	<i>Centroscyllium fabricii</i> , intestine	<i>Bacillus</i> spp., <i>Clostridium perfringens</i> , <i>Salmonella</i> Typhimurium, <i>Proteus vulgaris</i>	(Bindiya et al., 2019, Bindiya et al., 2015)

AMP = antimicrobial peptide; NR = not reported, Ref = reference

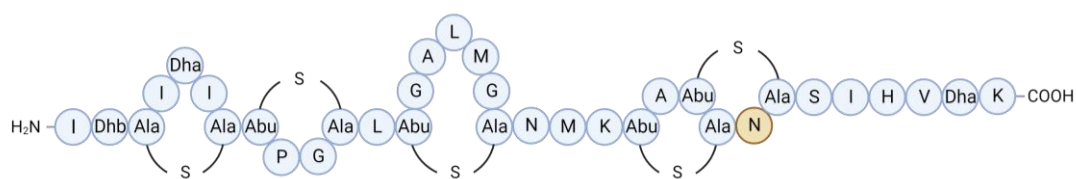


Figure 1: The structure of the class I bacteriocin (lanthipeptide) nisin Z, a natural nisin A variant, which differs by one amino acid (position 27 where Asn replaces Ala, highlighted in yellow). Nisin Z - producing *Lactococcus lactis* strains have been found in marine fish intestinal samples by two separate studies, Sequeiros et al. (2015) and Heo et al. (2012).

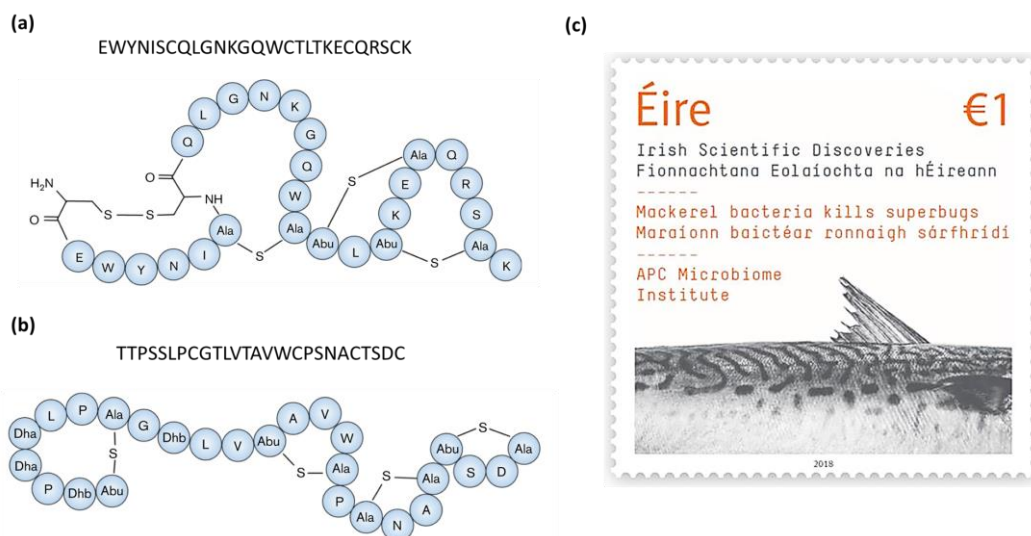


Figure 2: Predicted structures and core peptide sequences of the α (a) and β (b) peptides of Formicin, a novel two-component lanthipeptide produced by *Bacillus paralicheniformis* APC 1576, a mackerel gut isolate. Formicin's secondary structure differs from that of other similar lanthipeptides, yet it still retains a broad spectrum of activity against Gram-positive pathogens. Figure adapted from Collins et al. (2016). The discovery of formicin and its contribution to scientific research was commemorated by the release of a postal stamp released by An Post and Science Foundation Ireland, celebrating scientific discoveries in Ireland (c).

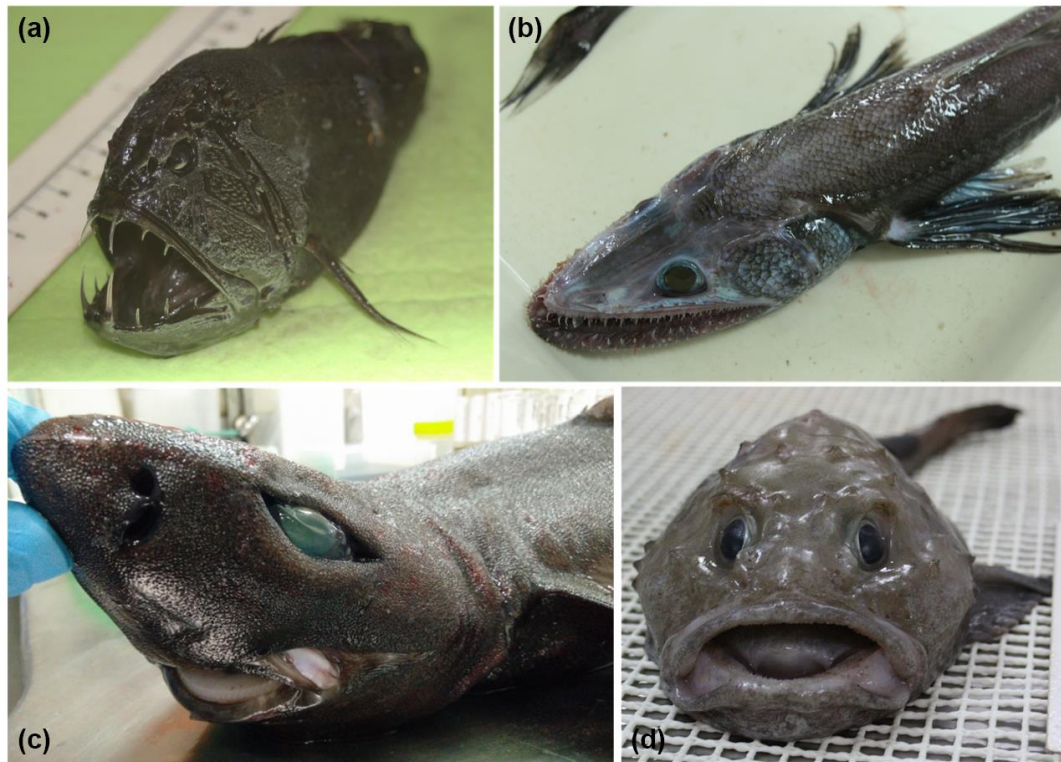


Figure 3: Culture-independent, metagenomics-based studies have enhanced our analysis of the microbiomes of marine fish, including rarer/unusual species such as *Anoplogaster cornuta* (a) *Bathysaurus ferox* (b) *Centroscyllium fabricii* (c) and *Cottunculus thomsonii* (d). Figure adapted from Collins et al. (2021).

Chapter 2

Bioactivity Screening and Genomic Analysis Reveals Deep-Sea Fish Microbiome Isolates as Sources of Novel Antimicrobials

Abstract

With the increase in antimicrobial resistance and the subsequent demand for novel therapeutics, the deep-sea fish microbiome can be a relatively untapped source of antimicrobials, including bacteriocins. Previously, bacterial isolates were recovered from the gut of deep-sea fish sampled from the Atlantic Ocean. In this study, we used *in vitro* methods to screen a subset of these isolates for antimicrobial activity, and subsequently mined genomic DNA from isolates of interest for bacteriocin and other antimicrobial metabolite genes. We observed antimicrobial activity against foodborne pathogens, including *Staphylococcus aureus*, *Listeria monocytogenes*, *Enterococcus faecalis* and *Micrococcus luteus*. In total, 147 candidate biosynthetic gene clusters were identified in the genomic sequences, including 35 bacteriocin/RiPP-like clusters. Other bioactive metabolite genes detected included non-ribosomal peptide synthases (NRPS), polyketide synthases (PKS; Types 1 and 3), beta-lactones and terpenes. Moreover, four unique bacteriocin gene clusters were annotated and shown to encode novel peptides: a class IIc bacteriocin, two class IIId bacteriocins and a class I lanthipeptide (LanM subgroup). Our dual *in vitro* and *in silico* approach allowed for a more comprehensive understanding of the bacteriocinogenic potential of these deep-sea isolates and an insight into the antimicrobial molecules that they may produce.

2.1. Introduction

The deep sea is one of the largest biomes on Earth, encompassing 95 % of the Earth's oceanic volume and reaching depths of over 10,000 m (Greenaway et al., 2021, Danovaro et al., 2017). It is a unique environment that is characterised by low temperatures, high hydrostatic pressure—which increases by one atmosphere with every 10 m increase in water depth—low oxygen, and no sunlight beyond 1000 m (Somero, 1992). These conditions are relatively constant, with little impact from ocean currents and seasonality (Pernice et al., 2016). However, it is estimated that less than 0.001 % of the deep sea has been explored, and relatively little is known about its microbial inhabitants compared to those of the terrestrial world (Danovaro et al., 2017). The deep sea is, perhaps, one of the last frontiers of biodiscovery and, as such, has been gaining attention as a source of potential novel microbial compounds.

Deep-sea microorganisms are uniquely adapted to the extreme conditions of their environmental niche. Notably, deep-sea microbial proteins are often optimally active at low temperatures (Zhang and Zeng, 2008) and can have unusual amino acid structures (Zhou et al., 2020), leading to an increased tolerance to high pressure (Penhallurick and Ichiye, 2021, Yan et al., 2009).

Some recently described antimicrobials are produced by novel deep-sea bacteria isolated from deep-sea sediment. These include marthiapeptide A, a cyclic peptide produced by *Marinactinospora thermotolerans* (Zhou et al., 2012); lobophorins E, F and K, antibiotics from *Streptomyces* spp. (Niu et al., 2011, Braña et al., 2017a); and anthracimycin, an anti-anthrax (*Bacillus anthracis*) antibiotic produced by *Streptomyces* sp. CNH365 (Jang et al., 2013). Other such antimicrobials have also been found in deep seawater, for example, brainimycins B and C (macrolides) produced by *Pseudonocardia carboxydivorans* (Braña et al., 2017b). Deep-sea bacteria found in association with higher organisms are also now gaining attention as producers of novel antimicrobials, particularly those from deep-sea coral (Braña et al., 2017a) and sponges (Back et al., 2021).

Bacteriocins are small (<10 kDa), ribosomally synthesised antimicrobial peptides produced by bacteria that target other bacteria, and to which the producer has

immunity. Bacteriocins can be divided into two main classes: class I (lantibiotics or lanthipeptides) and class II (non-lanthionine-containing bacteriocins) (Cotter et al., 2005). Lanthipeptides can be further subdivided into five subclasses, I–V, based on the biosynthetic enzyme(s) involved in their modification (Montalbán-López et al., 2021). The class II bacteriocins are subdivided into four groups: the anti-listerial pediocin-like peptides (class IIa), two-component peptides (class IIb), cyclic/circular bacteriocins (class IIc) and the linear and non-pediocin-like peptides (class IId) (Nissen-Meyer et al., 2009). Bacteriocins are emerging as alternatives to antibiotics due to their spectrum of activity—which can be broad or narrow—heat stability and their capacity for bioengineering to generate derivatives with value-added properties (Cotter et al., 2013). This is of particular importance, given the significant and growing threat that antibiotic-resistant organisms pose in the medical and food industries. According to a 2019 report by the Centres for Disease Control and Prevention (CDC), almost three million infections, including foodborne illness, caused by antibiotic-resistant organisms occur each year in the United States of America alone (CDC, 2019).

To date, very few studies have explored the production of bacteriocins from deep-sea fish microbial isolates. One example is BaCf3, a bacteriocin produced by *Bacillus amyloliquefaciens* BTSS3, isolated from the intestine of a deep-sea shark. This highly thermostable bacteriocin was initially identified through *in vitro* screening methods and showed inhibitory activity against pathogenic bacteria, including *Bacillus cereus*, *Clostridium perfringens*, *Staphylococcus aureus* and *Salmonella enterica* Typhimurium (Bindiya et al., 2019, Bindiya et al., 2015).

Previous work carried out in our laboratory generated metagenome-assembled genomes

(MAGs) from shotgun metagenomic sequencing of intestinal microbiomes samples from deep-sea fish from the Atlantic Ocean (Collins et al., 2021). The fish were collected as part of a groundfish survey. Analysis of the MAGS revealed information on the taxonomic and functional diversity of these samples, such as a predominance of Pseudomonadota (Proteobacteria), as well as a large abundance of genes involved in DNA repair, protein folding and motility. Bacterial isolates were recovered from the intestines and skin of a subset of these deep-sea fish and were preliminary

screened for bioactivity against the Gram-positive target organism, *Lactobacillus delbrueckii* subsp. *bulgaricus*. *L. bulgaricus* is commonly used in the assessment of antimicrobial activity due to its low-pH resistance and sensitivity to a wide range of bacteriocins (O'Sullivan et al., 2020, Sugrue et al., 2020, O'Shea et al., 2009). In this study, a cohort of these bacterial isolates were reassessed for bioactivity against the target strain, *Lactobacillus delbrueckii* subsp. *bulgaricus*, and were subsequently screened against a range of various target strains.

The aim of this study was to determine the bacteriocinogenic (ability to produce bacteriocins) potential of deep-sea fish microbial isolates (n = 65) through *in vitro* screening against various target strains, including pathogenic bacteria, and through *in silico* mining of the genomes of a selection of these isolates (n = 36) for putative bacteriocin biosynthetic gene clusters. We also looked for the presence of other secondary metabolite gene clusters and antimicrobial resistance (AMR) genes.

2.2. Materials and Methods

The workflow implemented in this study is illustrated in Figure 1.

2.2.1. Isolation and cultivation of antimicrobial-producing isolates

Fish were collected from sample sites in international waters off the Grand Banks of Newfoundland at depths ranging from 850 m to 1000 m, as previously described (Collins et al., 2021). An example of one the fish specimens and the sampling location is shown in Figure 2. Swabs were taken from the skin and intestines of the fish, plated onto BD Difco Marine

Agar 2216 (MA, Becton Drive, NJ, USA) and incubated at 4 °C for three weeks. Individual colonies (n = 374) were picked and stocked in 35 % glycerol and stored at –80 °C.

2.2.2. *In vitro* screening

Previous work in our laboratory (not published) conducted preliminary screening of a cohort of 374 stocked isolates against the indicator strain *Lactobacillus delbrueckii* subsp. *bulgaricus* LMG 6901 and identified 65 strains with antimicrobial activity. In this study, we reassessed the activity against *L. bulgaricus* by the 65 strains and examined the spectrum of antimicrobial activity of these strains against various target bacteria strains (listed in Table 1).

For analysis of activity against *L. bulgaricus*, isolates were plated on MA, brain heart infusion (BHI, Oxoid, Basingstoke, UK) agar and modified tryptic soy agar (mTSA, +0.5 g/L NaCl, Oxoid) and incubated at 5 °C (± 1 °C) for two weeks. The plates were then overlaid with sloppy media containing 0.8 % agar and 0.25 % of an overnight culture of *L. bulgaricus* in modified Difco MRS (mMRS, +0.5 g/L cysteine). Following incubation overnight, we measured the distance from the edge of the producing colony to the edge of the area of growth of the indicator organism and recorded the results qualitatively as follows: + (0.5–2.5 mm), ++ (2.6–5 mm), +++ (5.1–10 mm) and ++++ (>10 mm).

For subsequent colony overlay screening to assess the spectrum of activity, isolates were plated on MA and incubated at 5 °C (± 1 °C) for 2 weeks. The plates were then overlaid with indicator organisms in the relevant sloppy media, as listed in Table 1,

using the same method as mentioned above. Inhibition distances were measured after overnight incubation or after 48 h for indicators cultured at 20 °C.

2.2.3. 16S rRNA gene sequencing and taxonomy

Colony PCR was performed on all of the isolates. Cells were lysed by resuspending in 10 µL lysis buffer (20 mM Tris-HCL, pH 8.0, 2 mM EDTA, 1.2 % Triton X-100; reagents from Sigma-Aldrich/Merck, Burlington, Massachusetts, USA) and incubating at 95 °C for 10 min. PCR was performed in a total volume of 50 µL using 25 µL of BioMix™ Red (Meridian Bioscience®, London, UK), 16 µL of sterile molecular-grade water, 2 µL (at 10 pmol/µL) of the non-specific universal primers 27F (5'-AGAGTTTGATCCTGGCTCAG-3') and 1492R (5'-GGTTACCTTGTTACGACTT-3') and 5 µL of DNA template from lysed cells. The following amplification conditions were used: 95 °C for 1 min, followed by 32 cycles of 95 °C for 15s, 55 °C for 15 s and 72 °C for 32s with a final extension at 72 °C for 10 min. PCR products were run on a 1.5 % agarose gel, stained with MIDORI Green Advance and subject to gel electrophoresis. The gel was visualised using a UV-transilluminator to confirm the presence of amplified DNA. PCR products were purified using the GeneJET PCR purification Kit (Thermo Scientific, Waltham, MA, USA). DNA sequencing of both the forward and reverse strands was performed by GENEWIZ (Azenta Life Sciences, Leipzig, Germany) or Source BioScience (Waterford, Ireland). Where possible, sequence pairs were concatenated before analysis. Where available, 16S rRNA gene sequences were also extracted from subsequent gDNA sequencing data and used for inference. Taxonomy was inferred by searching the resulting 16S rRNA gene sequences for sequence similarity with NCBI's 16S rRNA database using the BLAST tool (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>, accessed on 3 November 2021).

2.2.4. Genomic DNA extraction and sequencing

Genomic DNA (gDNA) was extracted from 36 of the isolates using the GeneJET Genomic DNA Purification Kit (Thermo Scientific) and the relevant protocol, Gram-negative or Gram-positive, for the isolate. DNA quality was checked using agarose (1 %) gel electrophoresis. An overview of the sequencing and assembly information of each sample is given in Table S1a. In brief, the first set of gDNA samples were sequenced on an Illumina MiSeq platform (2 × 250 bp paired-end reads) at GenProbio

s.r.l. (Parma, Italy). The second set of samples were sequenced on the Illumina MiSeq platform (2 × 250 bp paired-end reads) by MicrobesNG (<http://www.microbesng.com>), which is supported by the BBSRC, (grant number BB/L024209/1). APC 4015 was sequenced by GENEWIZ (Azenta Life Sciences, Burlington, MA, USA) on the Illumina NovaSeq platform (2 × 150 bp paired end reads).

2.2.5. Genome assembly and annotation

Reads provided by MicrobesNG were processed and assembled by the sequencing company. For all other samples, read quality was assessed with FastQC (v0.11.9; <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>), trimmed with Trimmomatic (v0.36; Bolger et al., 2014) to a Phred score of 20 across a 4 bp sliding window. Reads less than 50 bp were discarded. Contigs (>500 bp) were assembled using SPAdes (v3.15.3; Bankevich et al., 2012), and assembly quality was assessed using QUAST (v4.4) (Gurevich et al., 2013) and CheckM (v1.0.18; Parks et al., 2015). Assemblies were submitted to Genbank and annotated using the NCBI Prokaryotic Annotation Pipeline (PGAP) upon submission (Tatusova et al., 2016). An overview of the assembly information, QUAST and CheckM statistics is given in Table S1. Additionally, the 16S rRNA gene sequences from the WGS data were searched against the SILVA rRNA gene project database (Quast et al., 2012). The resulting taxonomic descriptions and closest related reference strain(s) are reported in Table S2.

The datasets generated in this study are available under the BioProject accession number PRJNA883941.

2.2.6. *In silico* antimicrobial screening

Assembled genomes were screened for putative bacteriocin and secondary metabolite gene clusters using the genome mining tools BAGEL4 (BAGEL v4; van Heel et al., 2018) and AntiSMASH (v.5.1.2.; Blin et al., 2019), respectively.

2.2.7. Putative biosynthetic gene cluster (PBGC) analysis

Putative core precursor peptide genes were analysed using RiPPMiner's sequence similarity search (Agrawal et al., 2017) to query their similarity to known RiPPs from the database. Multiple alignments of the core peptide sequences and the RiPPMiner hits were made using Clustal Omega (Sievers et al., 2011) and visualised using Jalview (Waterhouse et al., 2009). Pairwise alignments between the precursor peptide

sequences and the RiPPMiner hits/reference sequences were generated using EMBOSS Needle (Madeira et al., 2022) using default parameters. The matrix used for each pairwise alignment is provided along with the results in Table S3. Genes within the PBGCs adjacent to the core peptide genes were queried and annotated using BLASTp to determine their potential function. The annotated PBGCs were visualised using Gene Graphics (Harrison et al., 2018).

2.2.8. Antimicrobial resistance gene screening

Genomes were screened for the presence of antimicrobial resistance genes using ABRicate (<https://github.com/tseemann/abricate>). The databases used were ARG-ANNOT (Gupta et al., 2014), CARD (Jia et al., 2017), EcOH (Ingle et al., 2016), MEGARES (Doster et al., 2020), NCBI (Feldgarden et al., 2019), PlasmidFinder (Carattoli et al., 2014), Resfinder (Zankari et al., 2012) and VFDB (Chen et al., 2016).

2.3. Results

For this study, we included bacteria that had been isolated from the skin and intestines of deep-sea fish and assessed them for antimicrobial activity against various target strains, including foodborne pathogenic bacteria. Those producing antimicrobial activity were identified by 16S rRNA gene analysis. Isolates of interest were selected for whole genome sequencing and subsequent *in silico* screening for putative bacteriocin gene clusters and putative biosynthetic gene clusters (PBGCs) encoding secondary metabolites. Strains were also assessed for the presence of antimicrobial resistance (AMR) genes.

2.3.1. Recovery and selection of bioactive isolates

The 65 bacterial isolates came from eight different deep-sea fish, namely: *Alepocephalus bairdii*, *Anoplogaster cornuta*, *Apristurus* sp., *Bathysaurus ferox*, *Centroscymnus coelolepis*, *Malacosteus niger*, *Notacanthus chemnitzii*, and *Simenchelys parasitica*. The majority (63 %) of the isolates came from skin swabs, with most isolates recovered from the deep-sea fish *N. chemnitzii* and *C. coelolepis* (Figure 3).

We confirmed that 44 of the 65 isolates were active against *L. bulgaricus* (Table 2). All the isolates grew on marine agar (MA) with minimal colony swarming, though modified tryptic soy agar (mTSA) resulted in slightly greater inhibitory activity. Twenty-one and eighteen of the isolates did not grow on brain heart infusion (BHI) or mTSA, respectively, under these conditions. In instances where inhibition could not be measured because of colony swarming or poor growth of the indicator, the result was recorded as “no data” (nd; Table 2).

2.3.2. Taxonomic diversity

The majority of the strains belonged to the class Gammaproteobacteria (of the phylum Pseudomonadota, at 72.3 %) followed by Bacilli (phylum Bacillota 16.9 %), Actinomycetia (phylum Actinomycetota, 9.2 %) and Flavobacteriia (phylum Bacteroidota, 1.5 %). The most abundant orders were Alteromonadales (49.2 %), Pseudomonadales (20 %) and Bacillales (16.9 %). Other represented orders were Micrococcales, Mycobacteriales, Flavobacteriales and Vibrionales. The most

abundant genus was *Pseudoalteromonas* (family *Pseudoalteromonadaceae*, 51.9 %), followed by *Psychrobacter* (family *Moraxellaceae*, 21.5 %), *Planococcus* (family *Planococcaceae*, 9.2 %) and *Staphylococcus* (family *Staphylococcaceae*, 7.7 %). The remaining genera were represented only by a single isolate; these included *Winogradskyella*, *Rhodococcus*, *Arthrobacter*, *Kocuria*, *Curtobacterium*, *Photobacterium* and *Micrococcus*. An overview of the taxonomic diversity of the isolates is shown in Figure 4.

2.3.3. Spectrum of activity

Following the initial screen against *L. bulgaricus*, the isolates were tested against a range of different target strains, including foodborne pathogens, food spoilage bacteria and marine bacteria. MA was chosen as the culture medium for the spectrum of activity screen as it supported the growth of the widest range of the marine isolates as compared to BHI and mTSA. Representative zones of inhibition are shown in Figure 5. Overall, 13 of the 22 target strains were inhibited (Figure 6). Four of the deep-sea strains that initially showed little or no activity against *L. bulgaricus* after culturing on marine agar were able to inhibit several of the other strains. Isolates of the genus *Planococcus* (strains APC 3900, APC 4015, APC 4016, 26D.a_f and 26D.b_f) were most effective at inhibiting the more closely related indicators: *Planococcus* sp. APC 3906 (of the same genus) and those of the order Micrococcales: *Arthrobacter* sp. APC 3897, *Microbacterium* sp. APC 3901 and *Micrococcus luteus*. *Planococcus* sp. APC 4016 additionally showed limited activity against *E. faecalis* and *S. pyogenes*. *Arthrobacter* sp. APC 3897 showed the greatest activity against *L. innocua* and *L. monocytogenes*. This isolate also showed activity against the *Planococcus* sp. and *Microbacterium* sp. indicators. *Pseudoalteromonas* sp. APC 4017 had the broadest spectrum of activity, inhibiting nine of the twenty-two indicators. None of the isolates inhibited *B. cereus*, *S. intermedius*, *Rhodococcus* sp. APC 3903 or *Kocuria* sp. APC 4018. None of the Gram-negative target strains (*E. coli*, *P. aeruginosa*, *S. Typhumurium*, *V. fischerii*, *Psychrobacter* sp. APC 3276) were inhibited. Thirteen of the isolates had no inhibitory effects against any of the indicators under these conditions. It was observed that several strains from the same sample site and the same host produced identical inhibition patterns and are likely to be identical strains.

For example, *Psychrobacter* sp. strains APC 3350 and APC 3277, both from the skin of *Centroscymnus coelolepis*, only inhibited *L. bulgaricus*; similarly, *Pseudoalteromonas* sp. strains APC 4020 and APC 4023 from the intestine of deep-sea fish *Bathysaurus ferox* both also inhibited only *L. bulgaricus*.

2.3.4. *In silico* antimicrobial screening

A selection of isolates that demonstrated *in vitro* antimicrobial activity and a representative of the identified taxonomic genera were chosen for whole genome sequencing. A total of 36 genomes were sequenced to enable genome-based *in silico* screening to identify antimicrobial-encoding genes and to identify potential bacteriocin producers that may not have shown *in vitro* activity under our tested conditions. The genome assemblies were cross-checked against the MAG sequence data (from Collins et al. [2021]), and it was confirmed that the isolate sequences were not present in the MAGs.

The BAGEL4 mining software searches for potential motifs and core peptides of ribosomally synthesised and post-translationally modified peptides (RiPPs), such as bacteriocins, as well as larger antimicrobial peptides (>10 kDa), and presents each output as an area of interest (AOI; van Heel et al., 2018). Nineteen AOIs were predicted by BAGEL4 across 13 of the 36 genomes consisting of fourteen class I bacteriocins, two class II bacteriocins and three larger antimicrobial proteins (>10 kDa) (Table 3 and Figure 7). Sactipeptides were the most abundant sub-type (seven hits), followed by lanthipeptides (six hits). Of the lanthipeptide hits, three were of the lanthipeptide subclass II, sharing 50–56 % similarity to the bacteriocins mersacidin and cerecidin (Table 3). The remaining three lanthipeptides were subclass III; however, they lacked an identifiable core peptide gene within the predicted gene cluster. The remaining bacteriocin/RiPP sub-types included one linear-azol(in)e-containing peptide (LAP), one class II circular bacteriocin with similarity to circularin A and one class II bacteriocin with similarity to lactococcin 972. Four AOIs each belonged to *Planococcus* strains, *Pseudoalteromonas* strains and the *Rhodococcus* sp. strain APC 3903. The remaining AOIs belonged to *Arthrobacter* sp. APC 3897, *Curtobacterium* sp. APC 3903, *Kocuria* sp. APC 4018, *Psychrobacter* sp. APC 3350 and *Winogradskyella* sp. APC 3343.

AntiSMASH identified 141 putative biosynthetic gene clusters (PBGCs) encoding secondary metabolites, with PBGCs being found in all 36 genomes (Table 4 and Figure 8). The most frequent PBGC types identified were RiPP-like (29 hits), non-ribosomal peptide synthase (NRPS, 26 hits), siderophore (20 hits), terpene (18 hits) and betalactone (13 hits). Additionally, PBGCs encoding resorcinol, redox-cofactors, type 3 and type 1 polyketide synthases (PKSs), non-alpha poly-amino acids (NAPAAAs), linear azol(in)e-containing peptides (LAPs), ectoine, butyrolactone and aryl polyene were also found. Proportionally, the *Rhodococcus* strain APC 3903 had the greatest number of predicted secondary metabolite gene clusters at 28 (18 % of the total). The presence of terpene BGCs is an indication that the orange colour observed in several of the colonies (Figure 5), for example, the *Planococcus* spp. is the result of terpene/carotenoid production. As with the BAGEL4 findings (Table 3), many of the predicted RiPP-like hits lacked an identifiable core biosynthetic gene, in which case the PBGC prediction may have been based on the presence of bacteriocin or RiPP-associated motifs or genes; this is indicated in where either no annotation was provided, or no known cluster hit was identified (Table 4).

The most promising candidates for further analysis, based on the presence of an identified core peptide/biosynthetic gene, were the putative cyclic bacteriocin in *Arthrobacter* APC 3897, the *Planococcus* spp. cerecidin-like lanthipeptides and the lactococcin 972-like peptides of *Curtobacterium* sp. APC 4022.

2.3.4.1. Class I LanM-group bacteriocins (Lanthipeptides)

Isolates APC 3900, 4015 and 4016 of the genus *Planococcus* all encoded identical sequences, which corresponded to a cerecidin-like lanthipeptide (Figure 9A). The gene cluster encoded two premature/pro-peptides, a large modification protein (LanM-like), a transporter, a regulation protein, a protease and ATP-binding/ABC transport proteins. Several small open reading frames (sORFs) were also identified, but their function could not be determined. The two peptide precursors, denoted Pep1 and Pep2 for clarity, shared 52.94 % identity with each other. When queried with RiPPMiner's sequence similarity search tool, the top hits were cerecidin A1–6 and cerecidin A7 (*B. cereus*). Pep1 shared 46.27 % and 44.78 % identity with cerA7 and cerA1–6, respectively. Pep2 shared 54.29 % and 55.71 % id to cerA7 and cerA1–

6, respectively. Other similar peptides from the database were cytolysin LI (*E. faecalis*), carnolysin A2 (*Carnobacterium maltaromaticum*), gardimycin (*Actinoplanes liguriae*), mersacidin (*B. amyloliquefaciens*) and cytolysin Ls (*E. faecalis*). The multiple sequence alignment (MSA) of the *Planococcus* and RiPP hit sequences in Figure 9B shows only a few conserved residues within the region between residues 50 and 60. As with cerecidin, the Pep1 and Pep2 peptide sequences appear to possess a leader sequence, denoted by the GG/GA motif, followed by six additional amino acids which precede a potential core peptide initiated by TT (Wang et al., 2014). The putative *Planococcus* modification gene shared 45.49 % identity to LanM of cerecidin (Genbank accession: AHJ59542.1), indicating that the *Planococcus* peptides are class II lanthipeptides (Arnison et al., 2013). The statistics generated using EMBOSS Needle of the pairwise alignments of the precursor peptides, and the reference peptides are given in Table S3. Figures of the pairwise alignments are given in Figures S1 and S2.

2.3.4.2. Class IIc (circular) bacteriocin

The cyclic bacteriocin gene cluster predicted in *Arthrobacter* isolate APC 3897 was annotated using BLASTp. The gene cluster consists of ten genes: a peptide precursor, a transporter/cleavage protein, a DUF95 family protein, three ATP-binding/transport proteins, an immunity protein and three hypothetical proteins of unknown function (Figure 10A). The predicted core precursor peptide sequence was queried using RiPPMiner: The closest related sequence was enterocin AS-48, a circular bacteriocin from *Enterococcus faecalis*, with which the peptide precursor sequences shared 71.3 % identity. The other top bacteriocin/RiPP hits were butyriovibriocin (30.56 %ID; *Butyrovibrio fibrisolvens*), circularin A (25.37 %; *Clostridium beijerinckii*), enterocin NKR-5-3b (32.47 %; *Enterococcus faecium*) and amylocyclicin (28.32 %; *Bacillus amyloliquefaciens*). The peptide sequences were aligned and are illustrated in Figure 10B, which is coloured according to shared identity of the aligned amino acid residues and sequences are arranged according to pairwise identity. The pairwise alignment statistics and Figures generated using EMBOSS Needle are provided in Table S3 and Figure S3, respectively.

2.3.4.3. Class IId bacteriocins

Curtobacterium sp. APC 4022 was found to contain two PBGCs located on two separate contigs that each encoded a lactococcin 972-like peptide (Figure 11A). The predicted gene cluster of the first peptide, denoted Pep1, consists of a core peptide gene, two ATP-binding/ABC-transport genes and two hypothetical genes of unknown function. The second gene cluster consists of four genes: a core peptide (Pep2) gene, a hypothetical gene of unknown function, an ATP-binding/ABC-transporter gene and a gene encoding an exonuclease family protein. When queried with RiPPMiner's sequence similarity tool, the most similar hit to both core peptide sequences was mutacin II (*Streptococcus mutans*). The two *Curtobacterium* peptide sequences were aligned with mutacin II and lactococcin 972 (Figure 11B). Pep1 and Pep2 share 20.97 % identity with each other. Pep1 shared 15.32 % and 12.4 % identity with lactococcin 972 and mutacin II, respectively. Pep2 shared 26 % and 15 % identity with lactococcin 972 and mutacin II, respectively. The pairwise alignment statistics generated using EMBOSS Needle are provided in Table S3. Figures of the pairwise alignments of *Curtobacterium* precursor peptides with the reference sequences are given in Figures S4 and S5. The pairwise alignment of Pep1 and Pep2 is also shown in Figure S4.

2.3.5. Antimicrobial resistance genes

The 36 deep-sea genomes were screened for AMR genes using ABRicate. Hits to resistance genes were only detected in the genome of *Rhodococcus* sp. APC 3903 (Table 5). Two genes corresponding to rifamycin/rifampin resistance were detected: (rif)iri, which encodes a rifampin (inactivating) monooxygenase and rpoB2, which encodes a resistant variant of the beta-subunit of RNA polymerase. A gene encoding a multidrug efflux protein (mtrA) was also detected.

2.4. Discussion

This study aimed to determine the bacteriocinogenic potential of bacterial isolates from the microbiomes of deep-sea fish. A combined approach of *in vitro* and *in silico* methods was used to identify antimicrobial-producing isolates and to further establish the potential source of this antimicrobial activity, with a focus on finding bacteriocin gene clusters.

The initial repeated screening against *L. bulgaricus* demonstrated that 44 of these isolates retained their bioactive activity (Table 2). The loss of activity by the remaining 21 isolates could be explained by the fact that antimicrobial production is often a tightly regulated process, dependent on the induction by external factors, such as other microbial metabolites present in the natural environment. Purification and removal of these isolates from the microbiome community may have affected their activity. Analysis of the spectrum of activity of these deep-sea isolates revealed that most strains had a narrow spectrum of activity. The only exceptions were *Pseudoalteromonas* strains APC 4017 and APC 3893 which were able to inhibit nine and eight of the indicator strains, respectively, with APC 4017 being particularly active against *E. faecium* 25644 and *Arthrobacter* sp. APC 3897 (Figure 6). Overall, just over half (11 out of 20) of the indicator strains were inhibited under the conditions tested. None of the Gram-negative indicator strains were impacted. Some inhibition patterns were noted amongst the taxonomic groups. The *Planococcus* strains were most effective against the *Planococcus* indicator strain APC 3906 and against the Micrococcales strains: *Arthrobacter* sp. APC 3897 and *Microbacterium* sp. APC 3901. The Gram-negative *Pseudoalteromonas* and *Psychrobacter* strains were only active against *L. bulgaricus* (with the aforementioned exceptions). The *Arthrobacter* strain APC 3897 was almost exclusively active against the *Listeria* strains. Based on the *in vitro* screening and taxonomic results, a number of isolates were chosen for whole-genome sequencing and *in silico* screening. Interestingly, the whole-genome sequence data from the isolates was not found within the MAGs. This is a reflection of the difference in sample sizes between the two studies but also indicates that the isolates in this study may only represent a very small sample of the deep-sea fish microbiome, many of which may, in fact, be unculturable.

Through *in silico* screening using BAGEL4, we identified 19 AOs, including lanthipeptide PBGCs and a cyclic bacteriocin PBGC. Though eleven of these AOs lacked an identifiable core biosynthetic bacteriocin gene, the predictions by BAGEL4 may be based on similarities of the AOs with bacteriocin-associated proteins. For example, in four of the predicted sactipeptide clusters, a gene was identified which encoded a putative GTP 3',8cyclase that shared similarity with MoaA, a member of the radical S-adenosylmethionine (rSAM) enzyme superfamily (Balo et al., 2022). Specific rSAM enzymes (sactisynthases) are essential for forming the characteristic sactonine linkages of sactipeptides (Grell et al., 2018). It is also possible that the core peptide genes are located elsewhere on these genomes, especially given that the *in silico* screening was carried out on contigs rather than full genomes. The lack of identifiable core peptides could also be due to the presence of unique motifs that are not found in the database.

The *Planococcus* strains APC 4015 and APC 4016, as mentioned above, demonstrated potent antimicrobial activity against several of the Gram-positive indicator strains, particularly against *Planococcus* sp. APC 3906, *Microbacterium* sp. APC 3901 and *Arthrobacter* sp. APC 3897. The *Planococcus* genomes were subsequently found to have identical gene clusters that encoded two putative bacteriocins. Of the known bacteriocin/RiPP sequences, the *Planococcus* peptides were most similar to the cerecidin peptides of *B.cereus*. Cerecidins are class II lanthipeptides that occur as two variants, CerA1 (of which there are six identical encoding genes, CerA1–6) and CerA7 (Wang et al., 2014). The sequences also shared some similarities with the haemolytic enterotoxin of *E. faecalis*, cytolysin, which is composed of two structural subunits, CyLL and CyLLs (Booth et al., 1996), and the two-component lantibiotic carnolysin (*C. maltaromaticum*; Lohans et al., 2014). Cerecidin and cytolysin both contain two cleavage sites. The first occurs at a GA motif site after the first leader sequence. The second cleavage site occurs after a secondary leader sequence of six amino acids, leaving behind the core peptide peptide (Wang et al., 2014, Booth et al., 1996). These processing sites are also potentially present in the *Planococcus* peptides, as indicated by the area of homology seen in the region between residues 50 and 60 (Figure 9B). This suggests that these peptides undergo multistep processing.

Arthrobacter isolate APC 3897, which was observed to have anti-listerial activity *in vitro*, was found to contain genes encoding a class IIc cyclic bacteriocin. The predicted bacteriocin peptide precursor was found to be most closely related to enterocin AS-48, a circular bacteriocin from *Enterococcus faecalis*. AS-48 is most active against Gram-positive bacteria (Gálvez et al., 1989), particularly *Listeria* species (Mendoza et al., 1999), which is consistent with the *in vitro* activity we observed for *Arthrobacter* sp. APC 3897. The structure of the PBGC was also consistent with that of enterocin AS-48, each consisting of ten genes, including a structural peptide precursor, a transporter protein, ATP-binding/transport proteins and an immunity protein gene (Grande Burgos et al., 2014). The *Arthrobacter* gene cluster was also found to encode a putative DUF95 family protein which is characteristic of circular bacteriocin gene clusters and is believed to be involved in the transport and secretion of these peptides (Mu et al., 2014). To date, the identification of an AS-48-like circular bacteriocin has not been reported in an *Arthrobacter* strain; in fact, some *Arthrobacter* strains have been used as sensitive indicators in studies of AS-48 activity (Cebrián et al., 2010, Cebrián et al., 2015).

Curtobacterium sp. APC 4022 displayed very little *in vitro* antimicrobial activity, having only shown inhibition of *L. bulgaricus* when cultured on mTSA (Table 2). The genome of this strain was found to contain two separate gene clusters encoding predicted bacteriocins with sequence similarities, albeit low, to the class IId bacteriocins lactococcin 972 and mutacin II. Lactococcin 972, first isolated from *Lactococcus lactis* subsp. *lactis* IPLA972, inhibits peptidoglycan synthesis by targeting lipid II and is almost exclusively active against lactococci (Martínez et al., 1996). Members of the lactococcin 972 protein family (Pfam IPR006540) are usually associated with transmembrane putative immunity proteins. Mutacin II, from *S. mutans* NTCC 10449, is active against a range of Gram-positive species, including streptococci, and targets cell membrane functions (Chikindas et al., 1995).

The observed *in vitro* antimicrobial activity of the *Planococcus*, *Arthrobacter* and *Curtobacterium* strains may be attributed to their respective putatively identified bacteriocins; however, we cannot rule out the possibility that these strains also are producing other bioactive secondary metabolites. The production of bacteriocins can also be a tightly regulated process, which we expect holds true for these deep-sea

strains. For example, the cerecidins of *B. cereus* are predicted to be under the control of competence quorum-sensing system genes (*comQXPA*), as well as the CerR regulatory protein, the stimulus of which is still unknown. Production of cecidinin in its native host has only been possible by constitutive expression of *cerR* using plasmids (Zhang et al., 2018). Further genomic and molecular analysis is needed to understand the regulatory systems behind the predicted bacteriocins of this study. Lastly, we screened all of the genomes for the presence of AMR genes using ABRicate. Only three AMR genes were detected, and all three were from a single genome, that of *Rhodococcus* sp. APC 3903. Firstly, a resistant variant of the RNA polymerase beta subunit

(RpoB2) was detected. Mutations to the RNA polymerase beta subunit (RpoB) can result in a gene duplication whereby one gene produces the rifamycin-sensitive RpoB and the second produces the resistant variant (Ishikawa et al., 2006). Possessing *rpoB* paralogs can be indicative of secondary metabolism and antibiotic production in Actinobacteria, though *in vitro* antimicrobial production was not detected in APC 4022 (Talà et al., 2009). Secondly, a (*rif*)/*iri* gene, which encodes a rifampin-inactivating monooxygenase, was identified. Production of rifamycin-inactivating or -modifying enzymes and mutations of RpoB are common resistance mechanisms and have previously been reported in *Rhodococcus* strains and related taxa; they are particularly well-studied in pathogenic bacteria such as *Mycobacteria* (Tanaka et al., 1996, Goldstein, 2014). The third AMR gene detected was a multidrug efflux pump gene. The identification of multiple AMR genes indicates a potential multiple resistance mechanism utilised by *Rhodococcus* sp. APC 3903. The overall relative lack of detectable AMR genes across all the genomes is consistent with other reports of the low incidence of AMR genes in deep-sea fish-associated microbial genomes (Collins et al., 2021, Yu et al., 2019). AMR genes have been detected in deep-sea water and sediment, but the incidence may be low compared to sample sites heavily impacted by anthropogenic activities (Chen et al., 2019, Zhang et al., 2022). There have been suggestions that AMR genes found in marine environments are largely unclassified (Hatosy and Martiny, 2015); therefore, the accurate identification of AMR genes in deep-sea samples may be even more challenging.

It is particularly notable that we were able to successfully isolate strains capable of targeting food-borne pathogens. Foodborne disease is a huge threat to public health and safety. The World Health Organisation (WHO) has estimated that globally up to 1 in 10 people fall ill to foodborne disease every year, leading to almost half a million premature deaths and huge economic losses (World Health, 2015). It is important, now more than ever, to be able to find novel compounds for combatting foodborne pathogens, particularly when treatments are being compounded by the increasing incidence of antimicrobial resistance. The aquaculture industry is another potential application of deep-sea bacteriocins, which we have not explored here in detail. For example, the overuse of antibiotics and the development of antimicrobial resistance is also a great concern in the fish farming industry, with disease outbreaks having significant economic and health impacts (Rather et al., 2017).

2.5. Conclusions

This study has shown that not only are bacteria from the skin and gut microbiomes of deep-sea fish cultivable, but they are a potentially rich source of antimicrobial agents. The majority of the antimicrobial producers in this study were Gram-negative *Pseudoalteromonas* strains of the phylum Pseudomonadota, though some of the most potent producers included *Planococcus* and *Arthrobacter* strains of the phyla Bacillota and Actinomycetota, respectively. Whilst initially selected based on their inhibitory activity against *Lactobacillus bulgaricus*, it was demonstrated that these deep-sea isolates could inhibit foodborne pathogens, such as *Listeria monocytogenes* and *Enterococcus* spp. Overall, the most susceptible indicators were other native marine strains, or more closely related organisms, a characteristic of the narrow spectrum of activity of many bacteriocins. *In silico* screening found that these isolates encode a wide variety of PBGCs encoding antimicrobial molecules from type I and II bacteriocins to LAPs and sactipeptides as well as PKSs, NRPSs, betalactones and terpenes. Included in these are potentially novel bacteriocins and variants of known bacteriocins. This provides great scope for future work in exploring applications of targeting pathogenic bacteria with antimicrobials that are produced, and potentially active at low temperatures. The dual *in vitro* and *in silico* approach allowed for a more comprehensive understanding of the potential targets of deep-sea bacteriocins and an insight into the antimicrobial molecules.

Data Availability Statement: The datasets generated in this study are available under the BioProject accession number PRJNA883941.

2.6. References

- ARNISON, P. G., BIBB, M. J., BIERBAUM, G., BOWERS, A. A., BUGNI, T. S., BULAJ, G., CAMARERO, J. A., CAMPOPIANO, D. J., CHALLIS, G. L., CLARDY, J., COTTER, P. D., CRAIK, D. J., DAWSON, M., DITTMANN, E., DONADIO, S., DORRESTEIN, P. C., ENTIAN, K.-D., FISCHBACH, M. A., GARAVELLI, J. S., GÖRANSSON, U., GRUBER, C. W., HAFT, D. H., HEMSCHIEDT, T. K., HERTWECK, C., HILL, C., HORSWILL, A. R., JASPARS, M., KELLY, W. L., KLINMAN, J. P., KUIPERS, O. P., LINK, A. J., LIU, W., MARAHIEL, M. A., MITCHELL, D. A., MOLL, G. N., MOORE, B. S., MÜLLER, R., NAIR, S. K., NES, I. F., NORRIS, G. E., OLIVERA, B. M., ONAKA, H., PATCHETT, M. L., PIEL, J., REANEY, M. J. T., REBUFFAT, S., ROSS, R. P., SAHL, H.-G., SCHMIDT, E. W., SELSTED, M. E., SEVERINOV, K., SHEN, B., SIVONEN, K., SMITH, L., STEIN, T., SÜSSMUTH, R. D., TAGG, J. R., TANG, G.-L., TRUMAN, A. W., VEDERAS, J. C., WALSH, C. T., WALTON, J. D., WENZEL, S. C., WILLEY, J. M. & VAN DER DONK, W. A. 2013. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Natural Product Reports*, 30, 108-160.
- BACK, C. R., STENNETT, H. L., WILLIAMS, S. E., WANG, L., OJEDA GOMEZ, J., ABDULLE, O. M., DUFFY, T., NEAL, C., MANTELL, J., JEPSON, M. A., HENDRY, K. R., POWELL, D., STACH, J. E. M., ESSEX-LOPRESTI, A. E., WILLIS, C. L., CURNOW, P. & RACE, P. R. 2021. A new *Micromonospora* strain with antibiotic activity isolated from the microbiome of a mid-atlantic deep-sea sponge. *Marine Drugs*, 19, 105.
- BALO, A. R., TAO, L. & BRITT, R. D. 2022. Characterizing SPASM/twitch domain-containing radical sam enzymes by EPR spectroscopy. *Applied Magnetic Resonance*, 53, 809-820.
- BANKEVICH, A., NURK, S., ANTIPOV, D., GUREVICH, A. A., DVORKIN, M., KULIKOV, A. S., LESIN, V. M., NIKOLENKO, S. I., PHAM, S., PRJIBELSKI, A. D., PYSHKIN, A. V., SIROTKIN, A. V., VYAHHI, N., TESLER, G., ALEKSEYEV, M. A. & PEVZNER, P. A. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology*, 19, 455-77.
- BINDIYA, E. S., TINA, K. J., RAGHUL, S. S. & BHAT, S. G. 2015. Characterization of deep sea fish gut bacteria with antagonistic potential, from *Centroscyllium fabricii* (Deep Sea Shark). *Probiotics and Antimicrobial Proteins*, 7, 157-63.
- BINDIYA, E. S., TINA, K. J., SASIDHARAN, R. S. & BHAT, S. G. 2019. BaCf3: highly thermostable bacteriocin from *Bacillus amyloliquefaciens* BTSS3 antagonistic on food-borne pathogens. *3 Biotech*, 9, 136.
- BLIN, K., SHAW, S., STEINKE, K., VILLEBRO, R., ZIEMERT, N., LEE, S. Y., MEDEMA, M. H. & WEBER, T. 2019. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Research*, 47, W81-W87.

- BOLGER, A. M., LOHSE, M. & USADEL, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*, 30, 2114-20.
- BOOTH, M. C., BOGIE, C. P., SAHL, H.-G., SIEZEN, R. J., HATTER, K. L. & GILMORE, M. S. 1996. Structural analysis and proteolytic activation of *Enterococcus faecalis* cytolysin, a novel lantibiotic. *Molecular Microbiology*, 21, 1175-1184.
- BRAÑA, A. F., SARMIENTO-VIZCAÍNO, A., OSSET, M., PÉREZ-VICTORIA, I., MARTÍN, J., DE PEDRO, N., DE LA CRUZ, M., DÍAZ, C., VICENTE, F., REYES, F., GARCÍA, L. A. & BLANCO, G. 2017a. Lobophorin K, a New natural product with cytotoxic activity produced by *Streptomyces* sp. M-207 associated with the deep-sea coral *Lophelia pertusa*. *Marine Drugs*, 15, 144..
- BRAÑA, A. F., SARMIENTO-VIZCAÍNO, A., PÉREZ-VICTORIA, I., OTERO, L., FERNÁNDEZ, J., PALACIOS, J. J., MARTÍN, J., DE LA CRUZ, M., DÍAZ, C., VICENTE, F., REYES, F., GARCÍA, L. A. & BLANCO, G. 2017b. Branimycins B and C, antibiotics produced by the abyssal Actinobacterium *Pseudonocardia carboxydivorans* M-227. *Journal of Natural Products*, 80, 569-573.
- CARATTOLI, A., ZANKARI, E., GARCÍA-FERNÁNDEZ, A., VOLDY LARSEN, M., LUND, O., VILLA, L., MØLLER AARESTRUP, F. & HASMAN, H. 2014. *In silico* detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrobial Agents and Chemotherapy*, 58, 3895-903.
- CDC 2019. Antibiotic resistance threats in the United States, 2019. U.S. Department of Health and Human Services, CDC.
- CEBRIÁN, R., MAQUEDA, M., NEIRA JOSÉ, L., VALDIVIA, E., MARTÍNEZ-BUENO, M. & MONTALBÁN-LÓPEZ, M. 2010. Insights into the functionality of the putative residues involved in Enterocin AS-48 maturation. *Applied and Environmental Microbiology*, 76, 7268-7276.
- CEBRIÁN, R., MARTÍNEZ-BUENO, M., VALDIVIA, E., ALBERT, A., MAQUEDA, M. & SÁNCHEZ-BARRENA, M. J. 2015. The bacteriocin AS-48 requires dimer dissociation followed by hydrophobic interactions with the membrane for antibacterial activity. *Journal of Structural Biology*, 190, 162-172.
- CHEN, H., BAI, X., JING, L., CHEN, R. & TENG, Y. 2019. Characterization of antibiotic resistance genes in the sediments of an urban river revealed by comparative metagenomics analysis. *Science of The Total Environment*, 653, 1513-1521.
- CHEN, L., ZHENG, D., LIU, B., YANG, J. & JIN, Q. 2016. VFDB 2016: hierarchical and refined dataset for big data analysis--10 years on. *Nucleic Acids Res*, 44, D694-7.
- CHIKINDAS, M. L., NOVÁK, J., DRIESSEN, A. J., KONINGS, W. N., SCHILLING, K. M. & CAUFIELD, P. W. 1995. Mutacin II, a bactericidal antibiotic from *Streptococcus mutans*. *Antimicrobial Agents and Chemotherapy*, 39, 2656-60.

- COLLINS, F. W. J., WALSH, C. J., GOMEZ-SALA, B., GUIJARRO-GARCÍA, E., STOKES, D., JAKOBSDÓTTIR, K. B., KRISTJÁNSSON, K., BURNS, F., COTTER, P. D., REA, M. C., HILL, C. & ROSS, R. P. 2021. The microbiome of deep-sea fish reveals new microbial species and a sparsity of antibiotic resistance genes. *Gut Microbes*, 13, 1-13.
- COTTER, P. D., HILL, C. & ROSS, R. P. 2005. Bacteriocins: developing innate immunity for food. *Nature Reviews Microbiology*, 3, 777-788.
- COTTER, P. D., ROSS, R. P. & HILL, C. 2013. Bacteriocins — a viable alternative to antibiotics? *Nature Reviews Microbiology*, 11, 95-105.
- DANOVARO, R., CORINALDESI, C., DELL'ANNO, A. & SNELGROVE, P. V. R. 2017. The deep-sea under global change. *Current Biology*, 27, R461-R465.
- DOSTER, E., LAKIN, S. M., DEAN, C. J., WOLFE, C., YOUNG, J. G., BOUCHER, C., BELK, K. E., NOYES, N. R. & MORLEY, P. S. 2020. MEGARes 2.0: a database for classification of antimicrobial drug, biocide and metal resistance determinants in metagenomic sequence data. *Nucleic Acids Research*, 48, D561-D569.
- FELDGARDEN, M., BROVER, V., HAFT, D. H., PRASAD, A. B., SLOTTA, D. J., TOLSTOY, I., TYSON, G. H., ZHAO, S., HSU, C. H., MCDERMOTT, P. F., TADESSE, D. A., MORALES, C., SIMMONS, M., TILLMAN, G., WASILENKO, J., FOLSTER, J. P. & KLIMKE, W. 2019. Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. *Antimicrobial Agents and Chemotherapy*, 63.
- GOLDSTEIN, B. P. 2014. Resistance to rifampicin: a review. *The Journal of Antibiotics*, 67, 625-630.
- GRANDE BURGOS, M. J., PULIDO, R. P., DEL CARMEN LÓPEZ AGUAYO, M., GÁLVEZ, A. & LUCAS, R. 2014. The cyclic antibacterial peptide Enterocin AS-48: Isolation, mode of action, and possible food applications. *International Journal of Molecular Sciences*. Switzerland.
- GREENAWAY, S. F., SULLIVAN, K. D., UMFRESS, S. H., BEITTEL, A. B. & WAGNER, K. D. 2021. Revised depth of the Challenger Deep from submersible transects; including a general method for precise, pressure-derived depths in the ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, 178, 103644.
- GRELL, T. A. J., KINCANNON, W. M., BRUENDER, N. A., BLAESI, E. J., KREBS, C., BANDARIAN, V. & DRENNAN, C. L. 2018. Structural and spectroscopic analyses of the sporulation killing factor biosynthetic enzyme SkfB, a bacterial AdoMet radical sactisynthase. *Journal of Biological Chemistry*, 293, 17349-17361.
- GUPTA, S. K., PADMANABHAN, B. R., DIENE, S. M., LOPEZ-ROJAS, R., KEMPF, M., LANDRAUD, L. & ROLAIN, J. M. 2014. ARG-ANNOT, a new bioinformatic tool

- to discover antibiotic resistance genes in bacterial genomes. *Antimicrobial Agents and Chemotherapy*, 58, 212-20.
- GUREVICH, A., SAVELIEV, V., VYAHHI, N. & TESLER, G. 2013. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*, 29, 1072-5.
- GÁLVEZ, A., MAQUEDA, M., MARTÍNEZ-BUENO, M. & VALDIVIA, E. 1989. Bactericidal and bacteriolytic action of peptide antibiotic AS-48 against gram-positive and gram-negative bacteria and other organisms. *Research in Microbiology*, 140, 57-68.
- HARRISON, K. J., CRÉCY-LAGARD, V. & ZALLOT, R. 2018. Gene Graphics: a genomic neighborhood data visualization web application. *Bioinformatics*, 34, 1406-1408.
- HATOSY, S. M. & MARTINY, A. C. 2015. The ocean as a global reservoir of antibiotic resistance genes. *Applied and Environmental Microbiology*, 81, 7593-7599.
- INGLE, D. J., VALCANIS, M., KUZEVSKI, A., TAUSCHEK, M., INOUE, M., STINEAR, T., LEVINE, M. M., ROBINS-BROWNE, R. M. & HOLT, K. E. 2016. *In silico* serotyping of *E. coli* from short read data identifies limited novel O-loci but extensive diversity of O:H serotype combinations within and between pathogenic lineages. *Microbial Genomics*, 2, e000064.
- ISHIKAWA, J., CHIBA, K., KURITA, H. & SATOH, H. 2006. Contribution of rpoB2 RNA polymerase beta subunit gene to rifampin resistance in *Nocardia* species. *Antimicrobial Agents and Chemotherapy*, 50, 1342-6.
- JANG, K. H., NAM, S.-J., LOCKE, J. B., KAUFFMAN, C. A., BEATTY, D. S., PAUL, L. A. & FENICAL, W. 2013. Anthracimycin, a potent anthrax antibiotic from a marine-derived Actinomycete. *Angewandte Chemie International Edition*, 52, 7822-7824.
- JIA, B., RAPHENYA, A. R., ALCOCK, B., WAGLECHNER, N., GUO, P., TSANG, K. K., LAGO, B. A., DAVE, B. M., PEREIRA, S., SHARMA, A. N., DOSHI, S., COURTOT, M., LO, R., WILLIAMS, L. E., FRYE, J. G., ELSAYEGH, T., SARDAR, D., WESTMAN, E. L., PAWLOWSKI, A. C., JOHNSON, T. A., BRINKMAN, F. S., WRIGHT, G. D. & MCARTHUR, A. G. 2017. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Research*, 45, D566-d573.
- LOHANS, C. T., LI, J. L. & VEDERAS, J. C. 2014. Structure and Biosynthesis of Carnolysin, a Homologue of Enterococcal Cytolysin with d-Amino Acids. *Journal of the American Chemical Society*, 136, 13150-13153.
- MADEIRA, F., PEARCE, M., TIVEY, A. R. N., BASUTKAR, P., LEE, J., EDBALI, O., MADHUSOODANAN, N., KOLESNIKOV, A. & LOPEZ, R. 2022. Search and sequence analysis tools services from EMBL-EBI in 2022. *Nucleic acids research*, 50, W276-W279.

- MARTÍNEZ, B., SUÁREZ, J. E. & RODRÍGUEZ, A. 1996. Lactococcin 972 : a homodimeric lactococcal bacteriocin whose primary target is not the plasma membrane. *Microbiology (Reading)*, 142 (Pt 9), 2393-8.
- MENDOZA, F., MAQUEDA, M., GÁLVEZ, A., MARTÍNEZ-BUENO, M. & VALDIVIA, E. 1999. Antilisterial activity of peptide AS-48 and study of changes induced in the cell envelope properties of an AS-48-adapted strain of *Listeria monocytogenes*. *Applied and Environmental Microbiology*, 65, 618-25.
- MONTALBÁN-LÓPEZ, M., SCOTT, T. A., RAMESH, S., RAHMAN, I. R., VAN HEEL, A. J., VIEL, J. H., BANDARIAN, V., DITTMANN, E., GENILLOUD, O., GOTO, Y., GRANDE BURGOS, M. J., HILL, C., KIM, S., KOEHNKE, J., LATHAM, J. A., LINK, A. J., MARTÍNEZ, B., NAIR, S. K., NICOLET, Y., REBUFFAT, S., SAHL, H.-G., SAREEN, D., SCHMIDT, E. W., SCHMITT, L., SEVERINOV, K., SÜSSMUTH, R. D., TRUMAN, A. W., WANG, H., WENG, J.-K., VAN WEZEL, G. P., ZHANG, Q., ZHONG, J., PIEL, J., MITCHELL, D. A., KUIPERS, O. P. & VAN DER DONK, W. A. 2021. New developments in RiPP discovery, enzymology and engineering. *Natural Product Reports*, 38, 130-239.
- MU, F., MASUDA, Y., ZENDO, T., ONO, H., KITAGAWA, H., ITO, H., NAKAYAMA, J. & SONOMOTO, K. 2014. Biological function of a DUF95 superfamily protein involved in the biosynthesis of a circular bacteriocin, leucocyclicin Q. *Journal of Bioscience and Bioengineering*, 117, 158-164.
- NISSEN-MEYER, J., ROGNE, P., OPPEGARD, C., HAUGEN, S. H. & KRISTIANSEN, E. P. 2009. Structure-function relationships of the non-lanthionine-containing peptide (class ii) bacteriocins produced by Gram-positive bacteria. *Current Pharmaceutical Biotechnology*, 10, 19-37.
- NIU, S., LI, S., CHEN, Y., TIAN, X., ZHANG, H., ZHANG, G., ZHANG, W., YANG, X., ZHANG, S., JU, J. & ZHANG, C. 2011. Lobophorins E and F, new spirotetronate antibiotics from a South China Sea-derived *Streptomyces* sp. SCSIO 01127. *The Journal of Antibiotics*, 64, 711-716.
- O'SHEA, E. F., GARDINER, G. E., O'CONNOR, P. M., MILLS, S., ROSS, R. P. & HILL, C. 2009. Characterization of enterocin- and salivaricin-producing lactic acid bacteria from the mammalian gastrointestinal tract. *FEMS Microbiology Letters*, 291, 24-34.
- O'SULLIVAN, J. N., O'CONNOR, P. M., REA, M. C., O'SULLIVAN, O., WALSH, C. J., HEALY, B., MATHUR, H., FIELD, D., HILL, C. & ROSS, R. P. 2020. Nisin J, a novel natural nisin variant, is produced by *Staphylococcus capitis* sourced from the human skin microbiota. *Journal of Bacteriology*, 202, e00639-19.
- PARKS, D. H., IMELFORT, M., SKENNERTON, C. T., HUGENHOLTZ, P. & TYSON, G. W. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome research*, 25, 1043-1055.

- PENHALLURICK, R. W. & ICHIYE, T. 2021. Pressure adaptations in deep-sea *Moritella* dihydrofolate reductases: Compressibility versus stability. *Biology (Basel)*, 10.
- PERNICE, M. C., GINER, C. R., LOGARES, R., PERERA-BEL, J., ACINAS, S. G., DUARTE, C. M., GASOL, J. M. & MASSANA, R. 2016. Large variability of bathypelagic microbial eukaryotic communities across the world's oceans. *The ISME journal*, 10, 945-958.
- QUAST, C., PRUESSE, E., YILMAZ, P., GERKEN, J., SCHWEER, T., YARZA, P., PEPLIES, J. & GLÖCKNER, F. O. 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research*, 41, D590-D596.
- RATHER, I. A., GALOPE, R., BAJPAI, V. K., LIM, J., PAEK, W. K. & PARK, Y. H. 2017. Diversity of marine bacteria and their bacteriocins: applications in aquaculture. *Reviews in Fisheries Science & Aquaculture*, 25, 257-269.
- SIEVERS, F., WILM, A., DINEEN, D., GIBSON, T. J., KARPLUS, K., LI, W., LOPEZ, R., MCWILLIAM, H., REMMERT, M., SÖDING, J., THOMPSON, J. D. & HIGGINS, D. G. 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, 7, 539.
- SOMERO, G. N. 1992. Biochemical ecology of deep-sea animals. *Experientia*, 48, 537-543.
- SUGRUE, I., O'CONNOR, P. M., HILL, C., STANTON, C. & ROSS, R. P. 2020. *Actinomyces* produces defensin-like bacteriocins (actifensins) with a highly degenerate structure and broad antimicrobial activity. *Journal of Bacteriology*, 202, e00529-19.
- TALÀ, A., WANG, G., ZEMANOVA, M., OKAMOTO, S., OCHI, K. & ALIFANO, P. 2009. Activation of dormant bacterial genes by *Nonomuraea* sp. strain ATCC 39727 mutant-type RNA polymerase. *Journal of Bacteriology*, 191, 805-814.
- TANAKA, Y., YAZAWA, K., DABBS, E. R., NISHIKAWA, K., KOMAKI, H., MIKAMI, Y., MIYAJI, M., MORISAKI, N. & IWASAKI, S. 1996. Different Rifampicin Inactivation Mechanisms in *Nocardia* and Related Taxa. *Microbiology and Immunology*, 40, 1-4.
- TATUSOVA, T., DICUCCIO, M., BADRETDIN, A., CHETVERNIN, V., NAWROCKI, E. P., ZASLAVSKY, L., LOMSADZE, A., PRUITT, K. D., BORODOVSKY, M. & OSTELL, J. 2016. NCBI prokaryotic genome annotation pipeline. *Nucleic Acids Research*, 44, 6614-24.
- VAN HEEL, A. J., DE JONG, A., SONG, C., VIEL, J. H., KOK, J. & KUIPERS, O. P. 2018. BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. *Nucleic Acids Research*, 46, W278-W281.
- WANG, J., ZHANG, L., TENG, K., SUN, S., SUN, Z. & ZHONG, J. 2014. Cerecidins, novel lantibiotics from *Bacillus cereus* with potent antimicrobial activity. *Applied and Environmental Microbiology*, 80, 2633-43.

- WATERHOUSE, A. M., PROCTER, J. B., MARTIN, D. M. A., CLAMP, M. & BARTON, G. J. 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 25, 1189-1191.
- WORLD HEALTH, O. 2015. *WHO estimates of the global burden of foodborne diseases: foodborne disease burden epidemiology reference group 2007-2015*, Geneva, World Health Organization.
- YAN, B.-Q., CHEN, X.-L., HOU, X.-Y., HE, H., ZHOU, B.-C. & ZHANG, Y.-Z. 2009. Molecular analysis of the gene encoding a cold-adapted halophilic subtilase from deep-sea psychrotolerant bacterium *Pseudoalteromonas* sp. SM9913: cloning, expression, characterization and function analysis of the C-terminal PPC domains. *Extremophiles*, 13, 725-733.
- YU, Y., ZHANG, Z., WANG, Y., LIAO, M., RONG, X., LI, B., WANG, K., CHEN, J. & ZHANG, H. 2019. Complete genome sequence of *Photobacterium damsela* subsp. *damsela* strain SSPD1601 isolated from deep-sea cage-cultured *Sebastes schlegelii* with septic skin ulcer. *International Journal of Genomics*, 2019, 4242653.
- ZANKARI, E., HASMAN, H., COSENTINO, S., VESTERGAARD, M., RASMUSSEN, S., LUND, O., AARESTRUP, F. M. & LARSEN, M. V. 2012. Identification of acquired antimicrobial resistance genes. *Journal of Antimicrobial Chemotherapy*, 67, 2640-4.
- ZHANG, H., WANG, Y., LIU, P., SUN, Y., DONG, X. & HU, X. 2022. Unveiling the occurrence, hosts and mobility potential of antibiotic resistance genes in the deep ocean. *Science of The Total Environment*, 816, 151539.
- ZHANG, J.-W. & ZENG, R.-Y. 2008. Molecular cloning and expression of a cold-adapted lipase gene from an antarctic deep sea psychrotrophic bacterium *Pseudomonas* sp. 7323. *Marine Biotechnology*, 10, 612-621.
- ZHANG, L., TENG, K., WANG, J., ZHANG, Z., ZHANG, J., SUN, S., LI, L., YANG, X. & ZHONG, J. 2018. CerR, a single-domain regulatory protein of the LuxR family, promotes Cerecidin production and immunity in *Bacillus cereus*. *Applied and Environmental Microbiology*, 84, e02245-17.
- ZHOU, S., LIU, G., ZHENG, R., SUN, C. & WU, S. 2020. Structural and functional insights into iturin w, a novel lipopeptide produced by the deep-sea bacterium *Bacillus* sp. strain wsm-1. *Applied and Environmental Microbiology*, 86, e01597-20.
- ZHOU, X., HUANG, H., CHEN, Y., TAN, J., SONG, Y., ZOU, J., TIAN, X., HUA, Y. & JU, J. 2012. Marthiapeptide A, an anti-infective and cytotoxic polythiazole cyclopeptide from a 60 l scale fermentation of the deep sea-derived *Marinactinospira thermotolerans* SCSIO 00652. *Journal of Natural Products*, 75, 2251-2255.

2.7. Chapter 2 Figures and Tables

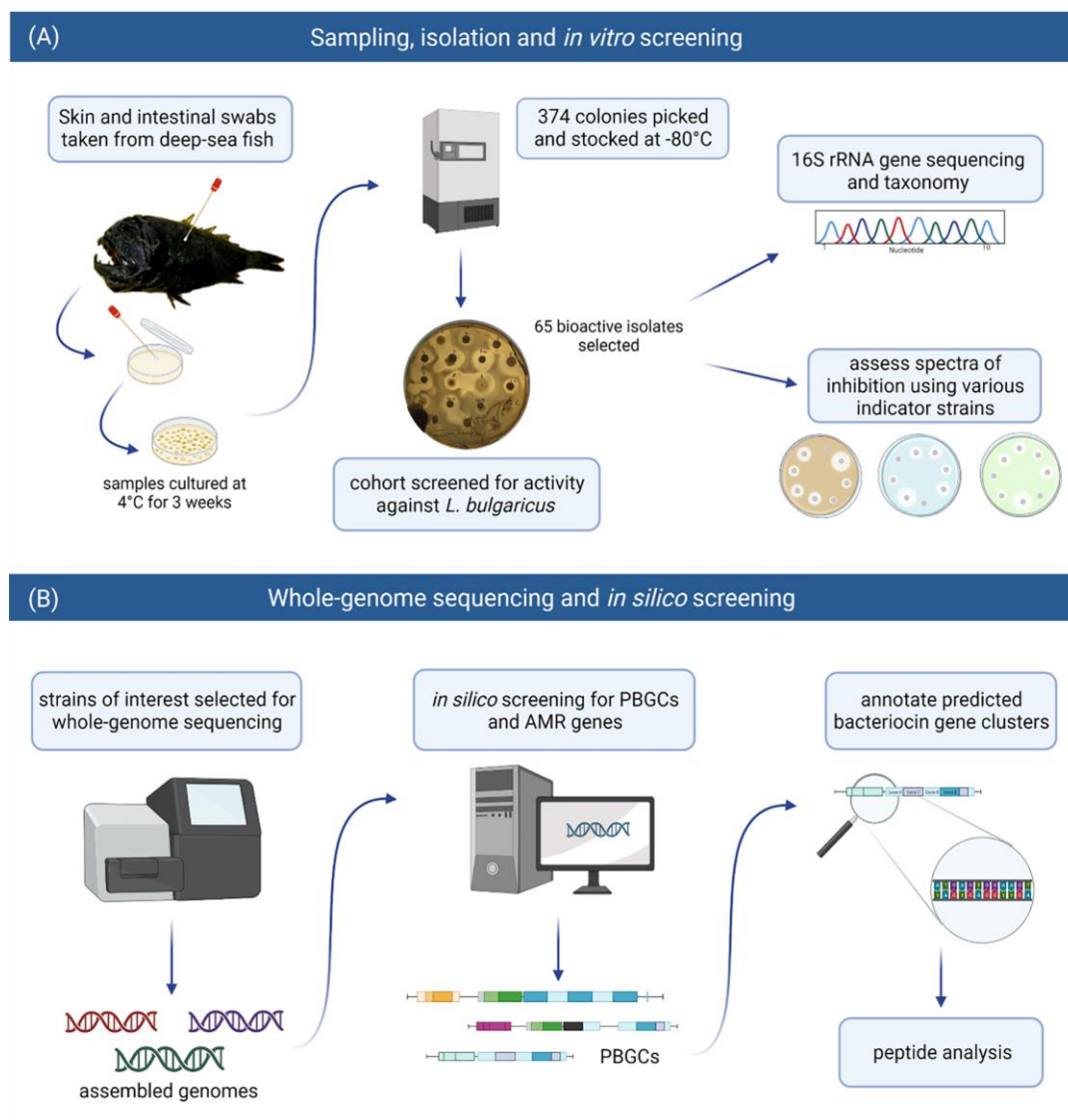


Figure 1. The *in vitro* and *in silico* -based approaches used in this study to assess the bacteriocinogenic potential in the deep-sea fish microbiome. (A) Swabs were taken from the skin and intestines of deep-sea fish. Samples were cultured at 4 °C for three weeks, and 374 isolates were recovered. A cohort of these isolates were screened for activity against *Lactobacillus delbrueckii* subsp. *bulgaricus*. Sixty-five bioactive isolates were identified and selected for further *in vitro* screening and identification through 16S rRNA gene analysis. (B) Strains of interest were selected for whole-genome sequencing and subsequent *in silico* screening for bacteriocin/RiPP genes, putative biosynthetic gene clusters (PBGCs) and antimicrobial resistance (AMR) genes. Predicted bacteriocin gene clusters were annotated, and the bacteriocin peptide precursors were analysed (Figure created with BioRender.com).

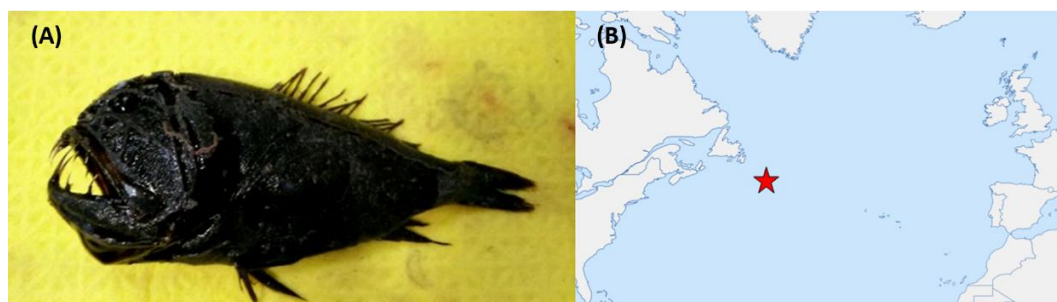


Figure 2. (A) *Anoplogaster cornuta*—an example of one of the fish species from which bioactive bacteria were isolated. (B) Approximate sampling location (indicated by the red star icon) in the Atlantic Ocean. Deep-sea fish specimens were collected from depths of approximately 1000 m.

Table 1: The target strains and respective growth conditions used in this study.

species	strain ID	temperature	atmosphere	growth media	notes
<i>Arthrobacter</i> sp.	APC 3897	20 °C	aerobic	MA	
<i>Bacillus cereus</i>	DPC 6087	37 °C	aerobic	BHI	
<i>Enterococcus faecalis</i>	OG1RF	37 °C	aerobic	BHI	
<i>Enterococcus faecium</i>	DSM 25644	37 °C	aerobic	BHI	
<i>Escherichia coli</i>	MG1655	37 °C	aerobic	LB	
<i>Kocuria</i> sp.	APC 4018	20 °C	aerobic	MA	
<i>Lactobacillus acidophilus</i>	EM066-BC-T3-3	37 °C	anaerobic	MRS	+ 0.5g/L cysteine
<i>Lactobacillus delbrueckii</i> subsp. <i>bulgaricus</i>	LMG 6901	37 °C	aerobic	MRS	+ 0.5g/L cysteine
<i>Lactococcus lactis</i> subsp. <i>cremoris</i>	HP	37 °C	aerobic	GM17	
<i>Listeria innocua</i>	ATCC 33090	30 °C	aerobic	BHI	
<i>Listeria monocytogenes</i>	EDG-e	37 °C	aerobic	BHI	
<i>Microbacterium</i> sp.	APC 3901	20 °C	aerobic	MA	
<i>Micrococcus luteus</i>	DSM 1790	30 °C	aerobic	BHI	
<i>Planococcus</i> sp.	APC 3906	20 °C	aerobic	MA	
<i>Pseudomonas aeruginosa</i>	PA01	37 °C	aerobic	BHI	
<i>Psychrobacter</i> sp.	APC 3276	20 °C	aerobic	MA	
<i>Rhodococcus</i> sp.	APC 3903	20 °C	aerobic	MA	
<i>Salmonella enterica</i> ser. Typhimurium	DPC 6046	37 °C	aerobic	BHI	
<i>Staphylococcus aureus</i>	RN4220	37 °C	aerobic	BHI	
<i>Staphylococcus intermedius</i>	DSM 20373	37 °C	aerobic	BHI	
<i>Streptococcus pyogenes</i>	DPC 6992	37 °C	aerobic	BHI	
<i>Vibrio fischeri</i>	n/a	20 °C	aerobic	MA	

ATCC, American Type Culture Collection; APC, APC Microbiome Ireland Culture Collection; DSM, Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures; DPC, Teagasc Culture Collection; LMG, Belgian Coordinated Collections of Microorganisms (BCCM).

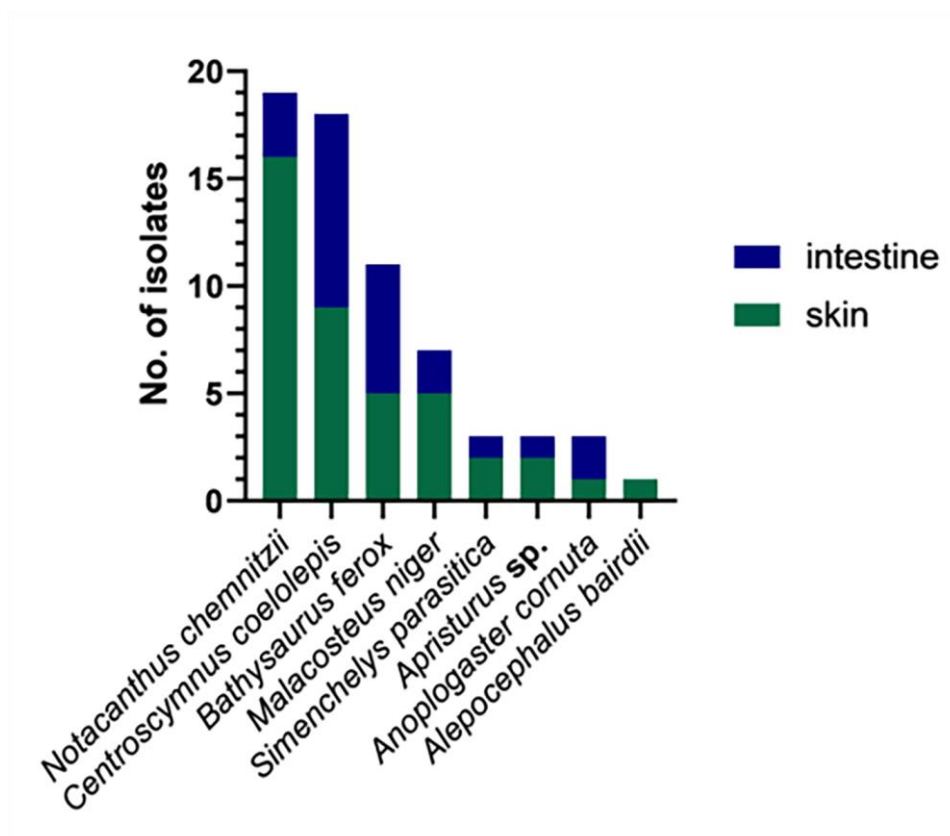


Figure 3. Total counts of isolates recovered per sample type (skin or intestine swab) and per fish host species. The majority of isolates were from skin swabs of *Notacanthus chemnitzii* and *Centroscymnus coelolepis*.

Table 2. Results of inhibitory activity observed against the target organism *Lactobacillus delbrueckii* subsp. *bulgaricus* (*L. bulgaricus*) on selected media. The isolate sample source and fish host are also given and numbered as follows: 1, *Anoplogaster cornuta*; 2, *Apristurus* sp.; 3, *Bathysaurus ferox*; 4, *Centroscymnus coelolepis*; 5, *Malacosteus niger*; 6, *Notacanthus chemnitzii*; and 7, *Alepocephalus bairdii*. Inhibition is indicated by + (0.5–2.5 mm), ++ (2.6–5 mm), +++ (5.1–10 mm), or – (no inhibition).

Isolate	Activity vs <i>L. bulgaricus</i> LMG 6901			Source (Sample, Host)
	MA	BHI	mTSA	
<i>Arthrobacter</i> sp. APC 3897	-	+	+	skin, 6
<i>Curtobacterium</i> sp. APC 4022	-	-	+	intestine, 4
<i>Photobacterium</i> sp. APC 3280	-	+	+	intestine, 4
<i>Planococcus</i> sp. 26D.a_F	-	nd	+++	skin, 6
<i>Pseudoalteromonas</i> spp.				
APC 3495	+	+	++	intestine, 1
APC 3213	++	ng	+	intestine, 2
APC 3238	+	ng	ng	intestine, 3
APC 3419	+	+	+	intestine, 3
APC 4023	-	+++	+	intestine, 3
APC 4024	+	ng	ng	intestine, 3
APC 3284	++	+++	ng	intestine, 4
APC 3355	++	nd	-	intestine, 4
APC 3356	++	+++	++	intestine, 4
APC 3358	+	++	+	intestine, 4
APC 3412	+	+	+	intestine, 5
APC 3224	++	ng	ng	intestine, 6
APC 3391	+	ng	+	skin, 7
APC 3502	+	+	+	skin, 1
APC 4025	+	ng	+	skin, 2
APC 4026	+	ng	ng	skin, 2
APC 3893	-	+	+++	skin, 3
APC 3895	++	ng	ng	skin, 3
APC 3896	+	ng	-	skin, 3
APC 4017	-	+++	+	skin, 3
APC 3274	-	+	+	skin, 4
APC 3691	+	ng	ng	skin, 4
APC 3407	+	nd	+	skin, 5
APC 3904	+	nd	++	skin, 5
APC 3221	+	+	+	skin, 6
APC 3250	+	+	+++	skin, 6
APC 4019	+	+	+	skin, 6
<i>Psychrobacter</i> spp.				
5A.1	-	++	+	skin, 4
5A.2	-	+++	+	skin, 4
APC 3272	-	++	-	skin, 6
APC 3275	-	+	+	intestine, 6
APC 3276	-	+	+	skin, 4
APC 3277	+	+	+	skin, 4
APC 3281	-	+	+	intestine, 4
APC 3350	+	++	-	skin, 4
APC 3426	-	++	+++	skin, 3
APC 3692	++	nd	+	skin, 6
APC 4028	-	++	+	skin, 4

MA = BD Difco 2216 marine agar; BHI = brain heart infusion; mTSA = tryptic soy agar +0.5 g/L NaCl; ng = no growth of producing strain; nd = no measurement data.

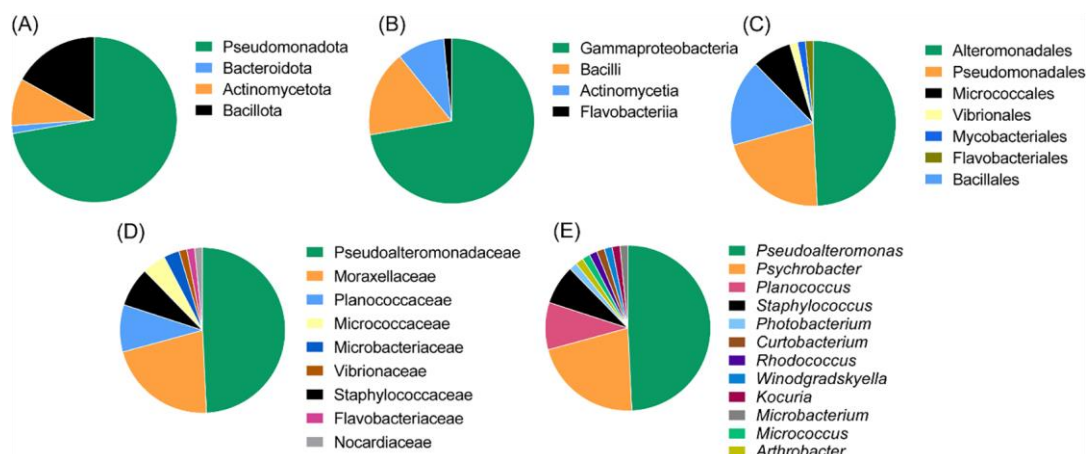


Figure 4. Diversity of the deep-sea fish microbiome isolates at the phylum level (A), class level (B), order level (C), family level (D) and genus level (E). The most abundant isolates in this study were *Pseudoalteromonas* and *Psychrobacter* species of the phylum Pseudomonadota (Proteobacteria).

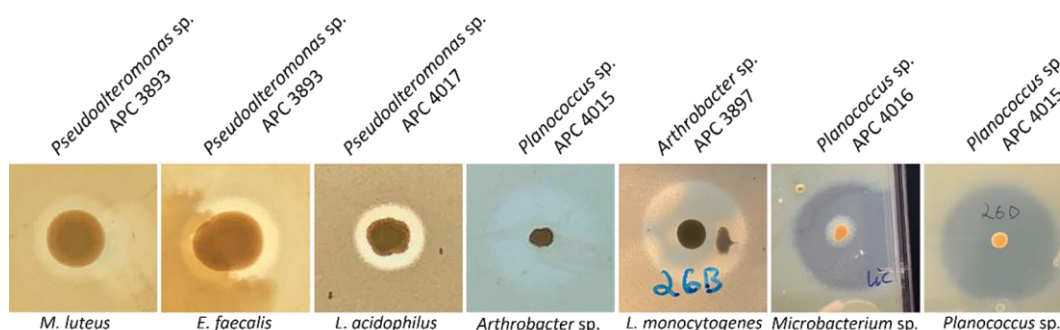


Figure 5. Representative zones of inhibition of the bioactive marine isolates. The strongest inhibition was demonstrated against other marine isolates, *Planococcus* sp. and *Microbacterium* sp. and *Listeria*, including *L. monocytogenes*. The producer IDs are given above the images, and the indicator strains are below.

isolate	<i>L. bulgaricus</i> LMG6901	<i>L. acidophilus</i> EM06-BC-T3-3	<i>L. innocua</i> ATCC33090	<i>L. monocytogenes</i> EDGe	<i>Planococcus</i> sp. APC-3906	<i>S. aureus</i> RN4220	<i>E. faecalis</i> OG1RE	<i>E. faecium</i> 25644	<i>S. pyogenes</i> 6992	<i>L. lactis</i> HP	<i>Microbacterium</i> sp. APC-3901	<i>Arthrobacter</i> sp. APC-3897	<i>M. luteus</i> DSM1790	source (sample, host)
<i>Planococcus</i> spp.														
APC 3900	0	0	0	0	0	0	0	0	0	0	+	0	0	skin, 6
APC 4015	0	0	0	0	++++	0	+	0	+	0	0	++++	0	skin, 6
APC 4016	0	0	0	0	++++	0	0	0	0	0	++++	++	+	skin, 6
26D.a_f	0	0	0	0	++++	0	0	0	0	0	0	+	0	skin, 6
26D.b_f	0	0	0	0	0	0	0	0	0	0	0	+	0	skin, 6
<i>Arthrobacter</i> sp.														
APC 3897	0	0	+++	++	+	0	0	0	0	0	+	0	0	skin, 6
<i>Psychrobacter</i> spp.														
APC 3350	+	0	0	0	0	0	0	0	0	0	0	0	0	skin, 4
APC 3277	+	0	0	0	0	0	0	0	0	0	0	0	0	skin, 4
APC 3692	++	0	0	0	0	0	0	0	0	0	0	0	0	skin, 6
<i>Pseudoalteromonas</i> spp.														
APC 3495	++	0	0	0	0	0	0	0	0	0	0	0	0	intestine, 1
APC 3213	+	0	0	0	0	0	0	0	0	0	0	0	0	intestine, 2
APC 3238	+	0	0	0	0	0	0	0	0	0	0	0	0	intestine, 3
APC 3419	++	0	0	0	0	0	0	0	0	0	0	0	0	intestine, 3
APC 4023	+	0	0	0	0	0	0	0	0	0	0	0	0	intestine, 3
APC 4024	+	0	0	0	0	0	0	0	0	0	0	0	0	intestine, 3
APC 3284	+	0	0	0	0	0	0	0	0	0	0	0	0	intestine, 4
APC 3355	++	0	0	0	0	0	0	0	0	0	0	0	0	intestine, 4
APC 3356	++	0	0	0	0	0	0	0	0	0	0	0	0	intestine, 4
APC 3358	++	0	0	0	0	0	0	0	0	0	0	0	0	intestine, 4
APC 3412	++	0	0	0	0	0	0	0	0	0	0	0	0	intestine, 5
APC 3224	+	0	0	0	0	0	0	0	0	0	0	0	0	intestine, 6
APC 3502	+	0	0	0	0	0	0	0	0	0	0	0	0	skin, 1
APC 4025	+	0	0	0	0	0	0	0	0	0	0	0	0	skin, 2
APC 4026	+	0	0	0	0	0	0	0	0	0	0	0	0	skin, 2
APC 3893	++	+	0	0	+	+	+	0	+	+	0	0	+	skin, 3
APC 3895	+	0	0	0	0	0	0	0	0	0	0	0	0	skin, 3
APC 4017	+	+	0	0	+	+	+	++	0	+	0	++	+	skin, 3
APC 3691	+	0	0	0	0	0	0	0	0	0	0	0	0	skin, 4
APC 3407	+	0	0	0	0	0	0	0	0	0	0	0	0	skin, 5
APC 3221	++	0	0	0	0	0	0	0	0	0	0	0	0	skin, 6
APC 3250	+	0	0	0	0	0	0	0	0	0	0	0	0	skin, 6
APC 4019	+	0	0	0	0	0	0	0	0	0	0	0	0	skin, 6
APC 3391	+	0	0	0	0	0	0	0	0	0	0	0	0	skin, 7

Figure 6. The spectrum of inhibition of the bioactive deep-sea isolates against the target strains by colony overlay assays on marine agar. The overall broadest range of activity was exhibited by *Pseudoalteromonas* strains APC 4017 and APC 3893. Inhibition is indicated by + (0.5–2.5 mm), ++ (2.6–5 mm) and +++ (5.1–10 mm) and is coloured to highlight inhibition patterns. Only the results of strains that demonstrated inhibitory activity under these conditions are reported in the Table.

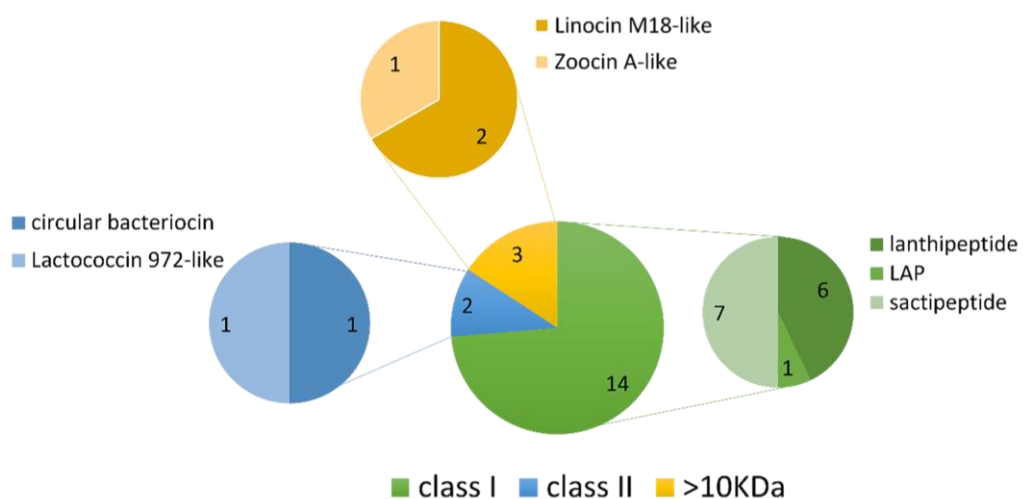


Figure 7. Summary of the bacteriocin classes and their subtypes identified using BAGEL4. Nineteen AOs were predicted in total across the 36 deep-sea fish gut isolate genomes. Sactipeptides and lanthipeptides (class I bacteriocins) were the most frequently predicted bacteriocin subtypes.

Table 3. Summary of the potential bacteriocin/RiPP gene clusters and their annotations predicted by BAGEL4 in the deep-sea bacterial genomes. The predicted bacteriocins included subclass II lanthipeptides (with similarity to mersacidin/cerecidin), subclass III lanthipeptides and a circularin A-like class II bacteriocin. Other predicted bacteriocin subtypes included sactipeptides, a LAP and a lactococcin 972-like peptide. Predicted AOIs which lacked an identifiable core peptide gene and could not be annotated by BAGEL4 are given as “undefined”. ×2 = two clusters found; %sim = % similarity.

Strain	BAGEL4		
	Prediction	Annotation	% Sim.
<i>Arthrobacter</i> sp. APC 3897	circularin_A	bacteriocin class cyclical uberolysin-like	67
	sactipeptide	GTP 3',8-cyclase (moaA)	73.70
<i>Curtobacterium</i> sp. APC 4022	lactococcin_972	bacteriocin (lactococcin_972)	39.70
	sactipeptide	GTP 3',8-cyclase (moaA)	59.50
<i>Kocuria</i> sp. APC 4018	linocin-M18	encapsulating protein for peroxidase	58.90
<i>Planococcus</i> sp. APC 3900	bacteriocin: mersacidin/cerecidin (x2)	bacteriocin class II with double-glycine leader (x2)	50–55.6
<i>Planococcus</i> sp. APC 4015	bacteriocin: mersacidin/cerecidin (x2)	bacteriocin class II with double-glycine leader peptide (x2)	50–55.6
	sactipeptide	undefined	-
<i>Planococcus</i> sp. APC 4016	bacteriocin: mersacidin/cerecidin (x2)	bacteriocin class II with double-glycine leader (x2)	50–55.6
<i>Pseudoalteromonas</i> sp. APC 3224	zoocin_A	undefined	36.40
<i>Pseudoalteromonas</i> sp. APC 3426	sactipeptide	undefined	-
<i>Pseudoalteromonas</i> sp. APC 3893	lanthipeptide class III	undefined	-
<i>Pseudoalteromonas</i> sp. APC 4017	lanthipeptide class III	undefined	-
<i>Psychrobacter</i> sp. APC 3350	sactipeptide	probable GTP 3',8-cyclase	-
<i>Rhodococcus</i> sp. APC 3903	lanthipeptide class III	undefined	-
	LAPs	undefined	-
	linocin_M18/putative bacteriocin family protein	encapsulating protein for peroxidase	66.70
	sactipeptide	undefined	-
<i>Winogradskyella</i> sp. APC 3343	sactipeptide	GTP 3',8-cyclase (moaA)	41.40

Table 4. *In silico* predictions by AntiSMASH of bacteriocin/RiPP-like PBGCs and other secondary metabolite gene clusters. Where no core peptide or no similar cluster was found, the result is given as “none.” ×2 = two clusters found; % sim = similarity; nd = no data.

Strain	AntiSMASH			
	Bacteriocin Prediction	Similar Cluster	% Sim.	Other PBGCs Present
<i>Arthrobacter</i> sp. APC 3897	RiPP-like	head-to-tail cyclized peptide	71	betalactone (x2), siderophore, T3PKS, terpene
<i>Curtobacterium</i> sp. APC 4022	RiPP-like (x2)	lactococcin_97 2 (x2)	nd	betalactone, NRPS-like, siderophore, T3PKS (x2), terpene
<i>Kocuria</i> sp. APC 4018	RiPP-like	linocin_M18	nd	betalactone, NRPS-like, siderophore, T3PKS, terpene
<i>Microbacterium</i> sp. APC 3898	none	-	-	betalactone, NAPAA, T3PKS, terpene
<i>Microbacterium</i> sp. APC 3901	none	-	-	NAPAA, T3PKS, betalactone, terpene
<i>Micrococcus</i> sp. APC 4021	none	-	-	betalactone, ectoine, NRPS-like, siderophore, terpene
<i>Planococcus</i> sp. APC 3900	lanthipeptide class II	cerecidin	70	terpene (x2)
<i>Planococcus</i> sp. APC 3906	none	-	-	terpene (x2)
<i>Planococcus</i> sp. APC 4015	lanthipeptide	cerecidin	70	terpene (x2)
<i>Planococcus</i> sp. APC 4016	lanthipeptide class II	cerecidin	70	terpene (x2)
<i>Pseudoalteromonas</i> sp. APC 3213	RiPP-like	burkholderic acid	15	siderophore
<i>Pseudoalteromonas</i> sp. APC 3215	RiPP-like	none	-	siderophore
<i>Pseudoalteromonas</i> sp. APC 3218	RiPP-like	none (DUF692-family protein)	nd	siderophore
<i>Pseudoalteromonas</i> sp. APC 3224	RiPP-like (x2)	none (DUF692-family protein, x2)	nd	NRPS, siderophore
<i>Pseudoalteromonas</i> sp. APC 3227	RiPP-like	none	-	siderophore
<i>Pseudoalteromonas</i> sp. APC 3250	RiPP-like	none	-	siderophore
<i>Pseudoalteromonas</i> sp. APC 3350	RiPP-like	none (DUF692-family protein)	nd	betalactone, redox-cofactor, siderophore
<i>Pseudoalteromonas</i> sp. APC 3356	RiPP-like	none	-	siderophore
<i>Pseudoalteromonas</i> sp. APC 3358	RiPP-like	nucleocidin	17	arylpolyene
<i>Pseudoalteromonas</i> sp. APC 3426	none	-	-	betalactone
<i>Pseudoalteromonas</i> sp. APC 3495	RiPP-like	none (DUF692-family protein)	nd	siderophore
<i>Pseudoalteromonas</i> sp. APC 3691	RiPP-like	none (DUF692-family protein)	nd	arylpolyene, betalactone
<i>Pseudoalteromonas</i> sp. APC 3694	RiPP-like	none (DUF692-family protein)	nd	none

continued

Table 4 continued:

<i>Pseudoalteromonas</i> sp. APC 3893	RiPP-like	none	-	arylpolyene, resorcinol, NRPS (x2), PKS, T3PKS siderophore
<i>Pseudoalteromonas</i> sp. APC 3894	RiPP-like	none (DUF692-family protein)	nd	
<i>Pseudoalteromonas</i> sp. APC 3895	RiPP-like	none	-	siderophore
<i>Pseudoalteromonas</i> sp. APC 3896	RiPP-like	none	-	siderophore
<i>Pseudoalteromonas</i> sp. APC 3907	RiPP-like	none	-	siderophore, betalactone
<i>Pseudoalteromonas</i> sp. APC 4017	RiPP-like	none	-	arylpolyene, resorcinol(x2), PKS, NRPS, T3PKS siderophore
<i>Pseudoalteromonas</i> sp. APC 4026	RiPP-like	none (DUF692-family protein)	nd	
<i>Psychrobacter</i> sp. 5A.1	RiPP-like	none (DUF692-family protein)	nd	betalactone, redox-cofactor, siderophore
<i>Psychrobacter</i> sp. APC 3279	none	-	-	betalactone
<i>Psychrobacter</i> sp. APC 3281	none	-	-	betalactone, redox-cofactor
<i>Psychrobacter</i> sp. APC 3355	RiPP-like (x2)	none (DUF692-family protein, x2)	nd	siderophore
<i>Rhodococcus</i> sp. APC 3903	LAP	dissonitrile antibiotic SF2768	11	butyrolactone, ectoine, NRPS (x15), NRPS-like (x3), terpene (x2), PKS-like, T1PKS (x2),
	RiPP-like	branched-chain fatty acids	75	
	lanthipeptide class III	none	-	
<i>Winogradskyella</i> sp. APC 3343	none	-	-	NRPS, T1pks, T3PKS, Terpene (x2)

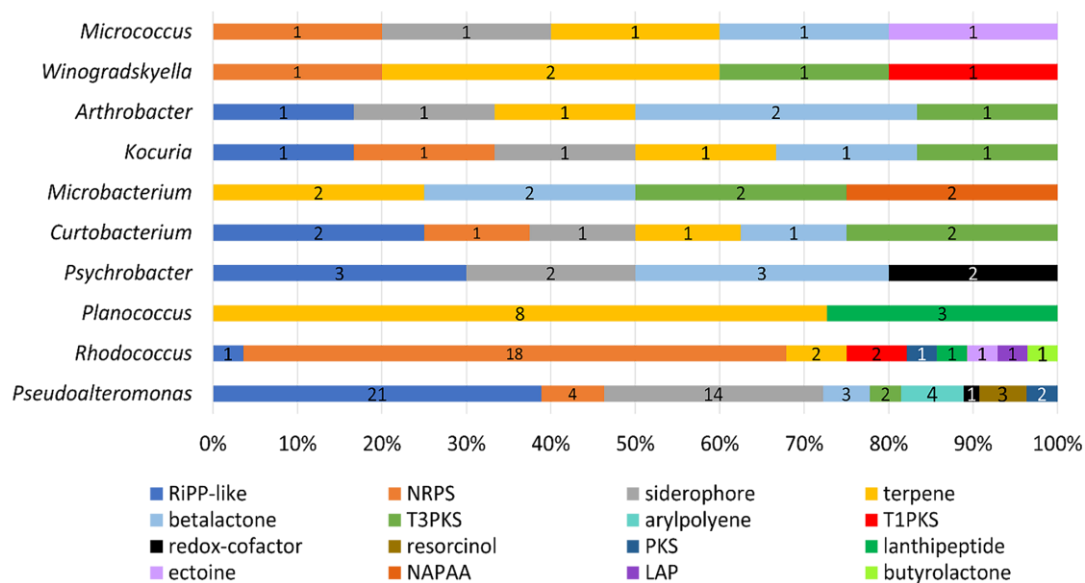


Figure 8. PBGC hits predicted by the AntiSMASH database of secondary metabolites and antibiotics. Results are given as a percentage of the total per genus and counts of each PBGC type are given. Default search strictness parameters were used. The most frequent hits included RiPPs and lanthipeptides, which suggests these isolates encode bioactive compounds, such as bacteriocins. Other gene cluster hits include polyketide synthases, siderophores, terpenes and non-ribosomal peptide synthases (NRPS).

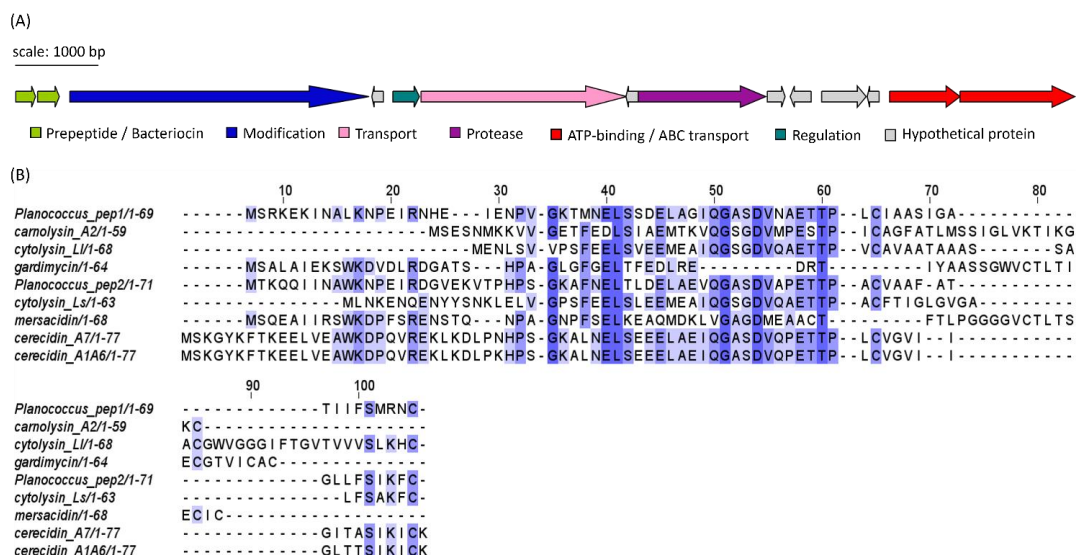


Figure 9. (A) Putative lanthipeptide (subclass II) gene cluster identified in *Planococcus* spp. isolates (*Planococcus* peptides 1 and 2). Identical clusters were found in APC 3900, APC 4015 and APC 4017. Genes are coloured according to predicted function. **(B)** MSA of the predicted *Planococcus* peptide precursor/bacteriocin sequence with most similar bacteriocin sequences identified using RiPPMiner: cytolysin LI (*E. faecalis*), carnolysin A2 (*Carnobacterium maltaromaticum*), mersacidin (*B. amyloliquefaciens*), cytolysin Ls (*E. faecalis*), cerecidin A1-6 and cerecidin A7 (*Bacillus cereus*), gardimycin (*Actinoplanes liguriae*). The MSA is coloured according to shared identity of the aligned amino acid residues at a given position, with the colour shade deepening as the percentage of sequences sharing the same residue increases. No colour indicates <40% of sequences share the residue.

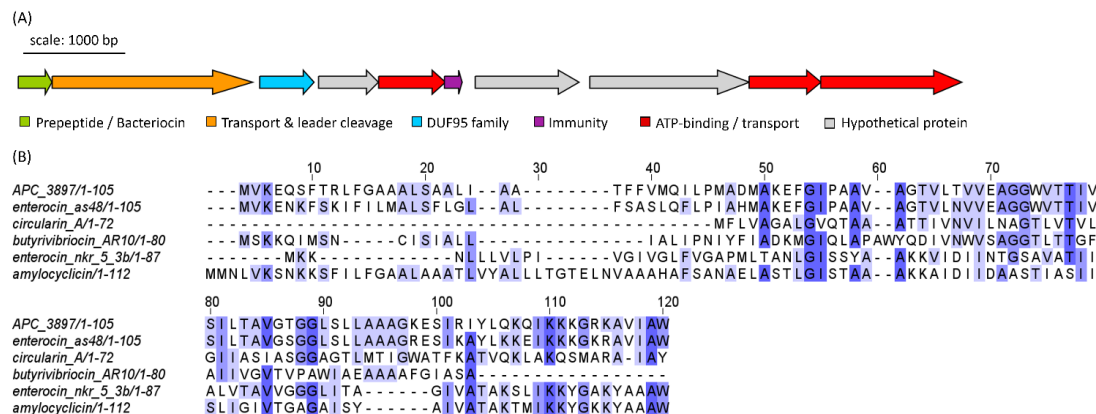


Figure 10. (A) Biosynthetic gene cluster (BGC) of a putative class IIc bacteriocin from *Arthrobacter* sp. isolate APC 3897. Genes are coloured according to predicted function. **(B)** Multiple sequence alignment (MSA) of APC 3897 BGC with similar bacteriocin sequences identified using RiPPMiner: butyrivibriocin (*Butyrivibrio fibrisolvens*), circularin A (*Clostridium beijerinckii*), enterocin NKR-5-3b (*Enterococcus faecium*), enterocin AS-48 (*E. faecalis*) and amylocyclicin (*Bacillus amyloliquefaciens*). The MSA is coloured according to shared identity of the aligned amino acid residues at a given position, with the colour shade deepening as the percentage of sequences sharing the same residue increases. No colour indicates <40% of sequences share the residue.

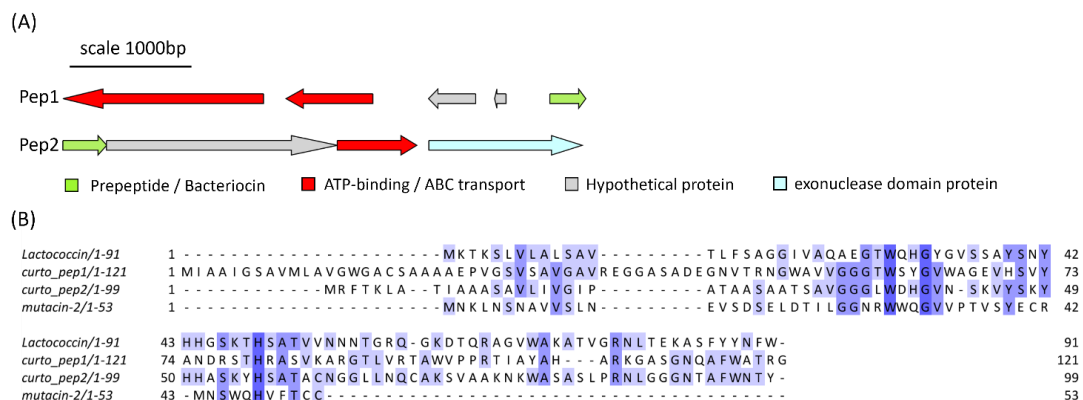


Figure 11. (A) The predicted gene clusters of the two lactococcin 972-like class IId bacteriocins from *Curtobacterium* sp. APC 4022, labelled pep1 and pep2. (B) The MSA of the peptide precursor sequences with mutacin II (*Streptococcus mutans*) and lactococcin 972 (*L. lactis* subsp. *lactis*) shows some conserved residues between the sequences. The MSA is coloured according to shared identity of the aligned amino acid residues at a given position, with the colour shade deepening as the percentage of sequences sharing the same residue increases. No colour indicates <40% of sequences share the residue.

Table 5. The output of ABRicate genome screening for antimicrobial resistance (AMR) genes. All hits came from the genome sequence of *Rhodococcus* sp. APC 3903.

contig	gene	resistance	%ID	database	product	accession
6	<i>rpoB2</i>	rifamycin/rifampin	80	card	beta-subunit of RNA polymerase (RpoB2)	AP006618.1:483519-4838688
8	<i>(rif)iri</i>	rifamycin/rifampin	97.81	argannot	(Rif)iri	U56415:280-1719
8	<i>iri</i>	rifamycin/rifampin	97.81	ncbi	rifampin monooxygenase Iri	NG_047911.1
8	<i>iri</i>	rifamycin/rifampin	97.81	card	rifampin monooxygenase Iri	U56415:279-1719
8	<i>iri</i>	rifamycin/rifampin	97.81	megares	rifampin monooxygenase Iri	MEG_3434
9	<i>mtrA</i>	macrolide;penam	80.41	card	transcriptional activator of multidrug efflux pump (MtrCDE)	AL123456.3:362734-3626662
9	<i>mtrA D</i>	multidrug	80.41	megares	multi-drug RND efflux regulator MTRAD	MEG_4078

A)					
plano_Pep2	1	-----MTKQQIINAWKNPEIRDGVEKVT	PHPSGKAFNELTLDELAEVQG	44	
cerecidin_A7	1	MSKGYKFTKEELVEAWKDPQVREKLDLP	KNHPSGKALNELSEELAEIQG	50	
plano_Pep2	45	ASDVAPETTPACVAAFATGLLFSIKFC-	71		
cerecidin_A7	51	ASDVQPETTPCLCVGVI-IGITASIKICK	77		
#=====					
B)					
plano_Pep2	1	-----MTKQQIINAWKNPEIRDGVEKVT	PHPSGKAFNELTLDELAEVQG	44	
cerecidin_A1A	1	MSKGYKFTKEELVEAWKDPQVREKLDLP	KHPSGKALNELSEELAEIQG	50	
plano_Pep2	45	ASDVAPETTPACVAAFATGLLFSIKFC-	71		
cerecidin_A1A	51	ASDVQPETTPCLCVGVI-IGLTTSIKICK	77		
#=====					
C)					
plano_Pep2	1	-MTKQQIINAWKNPEIRDGVEKVT	PHPSGKAFNELTLDELAEVQGASDVA	49	
cytolysin_Ls	1	MLNKENQENYYSNK-----LELVGP----	SFEELSLEEMAIQSGSDVQ	40	
plano_Pep2	50	PETTPACVAAFATGL----LFSIKFC	71		
cytolysin_Ls	41	AETTPAC---FTIGLVGALFSAKFC	63		
#=====					
D)					
plano_Pep2	1	MTKQQIINAWKNPEIRDGVEKVT	PHPSGKAFNELTLDELAEVQGASDVAP	50	
carnolysin_A1	1	-----MSEL--SMEKVV---GETFEDLS	IAEMTMVQSGDING	33	
plano_Pep2	51	E--TTPACV-----AAATGLLFSIKFC	71		
carnolysin_A1	34	EFTTSPACVYSVMVSKASSAKCAAGASAV	SGAILSAIRC	73	
#=====					
E)					
plano_Pep2	1	MTKQQIINAWKNPEIRDGVEKVT	PHPSGKAFNELTLDELAEVQGASDVAP	50	
carnolysin_A2	1	-----MSESNMKKVV---GETFEDLS	IAEMTKVQSGDVMP	33	
plano_Pep2	51	ETTPACVAAFAT----GLLFSIK-FC	71		
carnolysin_A2	34	ESTPIC-AGFATLMSSIGLVKTIKGC	59		
#=====					
F)					
plano_Pep2	1	MTKQQIINAWKNPEIRDGVEKVT	PHPSGKAFNELTLDELAEVQGASDVAP	50	
cytolysin_L1	1	-----MENLSVVP-----FEELSVEEM	AIQSGDVQA	29	
plano_Pep2	51	ETTPAC-VAAFA-----TG--LLFSIKFC	71		
cytolysin_L1	30	ETTPCVAATAAASSAACGWGGGIFTGVT	VVSLKHC	68	
#=====					
G)					
plano_Pep2	1	MTKQQIINAWKNPEIRDGVEKVT	PHPSGKAFNELTLDELAE---VQGASD	47	
gardimycin	1	MSALAIKSWKVDLRDG---ATSHPAGLGF	GELTFEDLREDRTIYAAS-	46	
plano_Pep2	48	VAPETTPACVAAFATGLLFSIKFC	71		
gardimycin	47	---SGWVCTLTIECGTV--ICAC	64		
#=====					
H)					
plano_Pep2	1	MTKQQIINAWKNPEIRDGVEKVT	PHPSGKAFNELTLDELAEVQGASDVAP	50	
mersacidin	1	MSQEAIIRSWKDPFSR---ENSTQNPAGN	PFSELKEAQMCKLVGAGDM--	45	
plano_Pep2	51	ETTPACVAAFATG----LLFSIKFC	71		
mersacidin	46	EA--ACTFTLPGGGGVCTLTSECIC	68		
#-----					

Figure S2: Pairwise alignment of the *Planococcus* precursor peptide, Pep2, with A) cerecidin A7 (*B. cereus*); B) cerecidin A1A6 (*B. cereus*); C) cytolysin Ls (*E. faecalis*); D) carnolysin A1 (*C. maltaromaticum*); E) carnolysin A2 (*C. maltaromaticum*); F) cytolysin L1 (*E. faecalis*); G) gardimycin (*A. liguriae*); H) mersacidin (*B. amyloliquifaciens*).

A)				
APC_3897	1	MVKEQSFTRLFGAAALS---AALIAATFFVMQILPMADMAKEFGIPAAVA	47	
enterocin_as4	1	MVKENKFSKIFILMALSFLGLALFSAS---LQFLPIAHMAKEFGIPAAVA	47	
APC_3897	48	GTVLTVVEAGGWTTIVSILTAVGTGGLSLLAAAGKESIRIYLQKQIKKK	97	
enterocin_as4	48	GTVLNVVEAGGWTTIVSILTAVGSGGLSLLAAAGRESIKAYLKKEIKKK	97	
APC_3897	98	GRKAVIAW 105		
enterocin_as4	98	GKRAVIAW 105		
#=====				
B)				
APC_3897	1	---MVK--EQSFTRLFGAAALSAAALIAATFFVMQILPM-----ADM	36	
amylocyclicin	1	MMNLVKSNKKSFI-LFGAALAAATLVYALLTGTELVAAAHAFSANAEL	49	
APC_3897	37	AKEFGIPAAGVAGTVLTVVEAGGWTTIVSILTAV-GTGGLSLLAAAGKES	85	
amylocyclicin	50	ASTLGISTAAAKAIDIIDAASTIASIISLIGIVTGAGAISYAIVATAKT	99	
APC_3897	86	IRIYLQKQIKKKGRKAVIAW 105		
amylocyclicin	100	M-----IKKYGKKYAAAW 112		
#=====				
C)				
APC_3897	1	MVKEQSFTRLFGAAALSAAALIAATFFVMQILPMADMAKEFGIPAAGVAGTV	50	
circularin_A	1	-----MFLVAGALGVQTAAATTI	18	
APC_3897	51	LTVVEAGGWTTIVSILTAVGTGGLSLLAAAGKESIRIYLQKQIKKKGRK	100	
circularin_A	19	VNVILNAGTLTVLGTIASIASGGAGTLMTIGWATFKATVQKLAKQSMAR	68	
APC_3897	101	AVIAW 105		
circularin_A	69	A-IAY 72		
#=====				
D)				
APC_3897	1	MVKEQSFTRLFGAAALSAAALIAATFFVMQIL-----PMADMAKEFG	41	
enterocin_nkr	1	-----MKKNLLVLPIVGIVGLFVGAPM--LTANLG	29	
APC_3897	42	IPAAVAGTVLTVVEAGGWTTIVSILTAVGTGGL---SLLAAAGKESIRI	88	
enterocin_nkr	30	ISSYAAKKVIDIINTGSAVATIIALVTAVVGGGLITAGIVATA-----	72	
APC_3897	89	YLQKQIKKKGRKAVIAW 105		
enterocin_nkr	73	--KSLIKKYGAKYAAAW 87		
#=====				
E)				
APC_3897	1	MVKEQSFTRLFGAAALSAAALIAATFFVMQILPMADMAKEFGIPAAVA--G	48	
butyrivibrioc	1	MSKKQIMSNCI-SIALIALIPNIYFI-----ADKMGIQLAPAWYQ	40	
APC_3897	49	TVLTVVEAGGWTTIVSILTAVGTGGLSLLAAA--GKESIRIYLQKQIKK	96	
butyrivibrioc	41	DIVNWVSAGGTLTTGFAIIVGVTPAWIAEAAAFGIASA-----	80	
APC_3897	97	KGRKAVIAW 105		
butyrivibrioc	81	----- 80		
#-----				
#-----				

Figure S3: Pairwise alignment of the *Arthrobacter* sp. APC 3897 precursor peptide with A) enterocin AS48 (*E. faecalis*); B) amylocyclicin (*B. amyloliquifaciens*); C) circularin A (*C. beierinckii*); D) enterocin nkr_5_3b (*E. faecium*); E) butyrivibriocin AR10 (*B. fibriosolvens*). Pairwise alignments were created in EMBOSS Needle.

Figure S4: Pairwise alignment of the *Curtobacterium* sp. APC 4022 precursor peptide curto_Lct1 with A) curto_Lct2); B) lactococcin 972 (*L. lactis*); C) mutacin 2 (*S. mutans*). Pairwise alignments were created in EMBOSS Needle.

A)			
curto_Lct2	1	MRFTKLATIAAASAVLIVGIPATAASAATSAVGGGLWDHGVN-SKVYSKY	49
		:.. :. . :..... . . . : .. .	
Lactococcin_9	1	MKTKSL--VLALSAVTLFSAGGIVAQAE-----GTWQHGYGVSSAYSNY	42
curto_Lct2	50	HHASKYHSATACNGGLLNQCAKSVAANKWASASLPRNLGGGNTAFWNTY	99
		 :.. . :..... . :.. :	
Lactococcin_9	43	HHGSKTHSATVNNNTGRQ-GKDTQRAGVWAKATVGRNLTEKASFYYN--	89
#=====			
B)			
curto_Lct2	1	MRFTKLATIAAASAVLIVGIPATAASAATSAVGGGLWDHGVNSKVYSKYH	50
		.. :.. : :.....:	
mutacin-2	1	--MNKLNSNA-----VVSLENEVSDSELDITLGGNRWWQGVVPTVSYECR	42
curto_Lct2	51	HASKYHSATACNGGLLNQCAKSVAANKWASASLPRNLGGGNTAFWNTY	99
		:.	
mutacin-2	43	MNSWQHVFTECC-----	53
#-----			

Figure S5: Pairwise alignment of the *Curtobacterium* sp. APC 4022 precursor peptide curto_Lct2 with A) lactococcin 972 (*L. lactis*) and B) mutacin 2 (*S. mutans*). Pairwise alignments were created in EMBOSS Needle.

Table S1a: Overview of whole genome sequencing and assembly data.

Sequencing and Assembly						
sample	genome accession	sequencing company	sequencing platform	reference	assembly type	assembly method (version)
5A.1	JASDDJ000000000	GenproBio	Illumina MiSeq (2 x 250)	*	<i>de novo</i>	spades (v. 3.15.3)
APC_3213	JASDCG000000000	MicrobesNG	Illumina MiSeq (2 x 250)	†	<i>de novo</i>	spades (v.13)
APC_3215	JASDCH000000000	MicrobesNG	Illumina MiSeq (2 x 250)	†	<i>de novo</i>	spades (v.13)
APC_3218	JASDCI000000000	MicrobesNG	Illumina MiSeq (2 x 250)	†	<i>de novo</i>	spades (v.13)
APC_3224	JASDDF000000000	GenproBio	Illumina MiSeq (2 x 250)	*	<i>de novo</i>	spades (v. 3.15.3)
APC_3227	JASDCJ000000000	MicrobesNG	Illumina MiSeq (2 x 250)	†	<i>de novo</i>	spades (v.13)
APC_3250	JASDCM000000000	MicrobesNG	Illumina MiSeq (2 x 250)	†	<i>de novo</i>	spades (v.13)
APC_3279	JASDCR000000000	MicrobesNG	Illumina MiSeq (2 x 250)	†	<i>de novo</i>	spades (v.13)
APC_3281	JASDCS000000000	GenproBio	Illumina MiSeq (2 x 250)	*	<i>de novo</i>	spades (v. 3.15.3)
APC_3343	JASDDK000000000	MicrobesNG	Illumina MiSeq (2 x 250)	†	<i>de novo</i>	spades (v.13)
APC_3350	JASDCX000000000	GenproBio	Illumina MiSeq (2 x 250)	*	<i>de novo</i>	spades (v. 3.15.3)
APC_3355	JASDDB000000000	GenproBio	Illumina MiSeq (2 x 250)	*	<i>de novo</i>	spades (v. 3.15.3)
APC_3356	JASDCN000000000	MicrobesNG	Illumina MiSeq (2 x 250)	†	<i>de novo</i>	spades (v.13)
APC_3358	JASDCE000000000	MicrobesNG	Illumina MiSeq (2 x 250)	†	<i>de novo</i>	spades (v.13)
APC_3426	JASDCF000000000	MicrobesNG	Illumina MiSeq (2 x 250)	†	<i>de novo</i>	spades (v.13)
APC_3495	JASDDA000000000	GenproBio	Illumina MiSeq (2 x 250)	*	<i>de novo</i>	spades (v. 3.15.3)
APC_3691	JASDCB000000000	MicrobesNG	Illumina MiSeq (2 x 250)	†	<i>de novo</i>	spades (v.13)
APC_3694	JASDDE000000000	GenproBio	Illumina MiSeq (2 x 250)	*	<i>de novo</i>	spades (v. 3.15.3)
APC_3893	JASDCD000000000	MicrobesNG	Illumina MiSeq (2 x 250)	†	<i>de novo</i>	spades (v.13)
APC_3894	JASDCK000000000	MicrobesNG	Illumina MiSeq (2 x 250)	†	<i>de novo</i>	spades (v.13)
APC_3895	JASDCL000000000	MicrobesNG	Illumina MiSeq (2 x 250)	†	<i>de novo</i>	spades (v.13)
APC_3896	JASDCP000000000	MicrobesNG	Illumina MiSeq (2 x 250)	†	<i>de novo</i>	spades (v.13)
APC_3897	JASDDG000000000	GenproBio	Illumina MiSeq (2 x 250)	*	<i>de novo</i>	spades (v. 3.15.3)
APC_3898	JASDDI000000000	GenproBio	Illumina MiSeq (2 x 250)	*	<i>de novo</i>	spades (v. 3.15.3)
APC_3900	JASDCT000000000	GenproBio	Illumina MiSeq (2 x 250)	*	<i>de novo</i>	spades (v. 3.15.3)

continued

Table S1a continued:

APC_3901	JASDCU000000000	GenproBio	Illumina MiSeq (2 x 250)	*	<i>de novo</i>	spades (v. 3.15.3)
APC_3903	JASDCV000000000	GenproBio	Illumina MiSeq (2 x 250)	*	<i>de novo</i>	spades (v. 3.15.3)
APC_3906	JASDCW000000000	GenproBio	Illumina MiSeq (2 x 250)	*	<i>de novo</i>	spades (v. 3.15.3)
APC_3907	JASDCO00000000	MicrobesNG	Illumina MiSeq (2 x 250)	†	<i>de novo</i>	spades (v.13)
APC_4015	JASDDH000000000	GenproBio	Illumina MiSeq (2 x 250)	*	<i>de novo</i>	spades (v. 3.15.3)
APC_4016	JASDCQ00000000	Genewiz	Illumina NovaSeq (2x150)	**	<i>de novo</i>	spades (v.13)
APC_4017	JASDCC000000000	MicrobesNG	Illumina MiSeq (2 x 250)	†	<i>de novo</i>	spades (v.13)
APC_4018	JASDCY000000000	GenproBio	Illumina MiSeq (2 x 250)	*	<i>de novo</i>	spades (v. 3.15.3)
APC_4021	JASDCZ000000000	GenproBio	Illumina MiSeq (2 x 250)	*	<i>de novo</i>	spades (v. 3.15.3)
APC_4022	JASDDD000000000	GenproBio	Illumina MiSeq (2 x 250)	*	<i>de novo</i>	spades (v. 3.15.3)
APC_4026	JASDDC000000000	GenproBio	Illumina MiSeq (2 x 250)	*	<i>de novo</i>	spades (v. 3.15.3)

* = <http://www.genprobio.com>, Parma, Italy; † = <http://www.microbesng.com>, University of Birmingham, UK;

** = Azenta Life Sciences, Leipzig, Germany

Table S1b. Overview of genome assembly quality metrics calculated with QUAST.

sample	mean coverage	no. reads	no. contigs	# contigs (≥ 1000 bp)	largest contig	QUAST metrics					# Ns per 100 kbp	%GC	# predicted rRNA genes
						total length	N50	L50	N90	L90			
5A.1	65	1168828	151	151	128044	3433409	49487	23	14349	75	11.8	42.97	4 + 1 part
APC_3213	119	1180329	45	86	720861	4177316	194983	7	67986	22	0	40.11	2 + 5 part
APC_3215	150	1424512	55	109	474907	4185738	137902	8	47610	26	0	40.18	1 + 7 part
APC_3218	117	1149754	49	96	500270	4195343	153850	9	69701	24	0	40.13	3 + 8 part
APC_3224	52	1078392	80	80	215426	4221042	97136	15	33310	41	25.6	39.88	1 + 2 part
APC_3227	52	482048	58	88	463852	4197794	199789	7	59970	25	0	40.12	3 + 8 part
APC_3250	139	1276369	47	111	542569	4110034	243731	7	59346	21	0	40.2	1 + 8 part
APC_3279	205	1838004	62	128	268467	3288828	148455	10	56162	24	0	42.91	6 + 0 part
APC_3281	53	826134	24	24	502341	3114096	347654	4	120480	10	6.7	42.8	3 + 0 part
APC_3343	32	310241	35	53	1191771	3832957	1042040	2	88982	7	0.03	33.43	6 + 0 part
APC_3350	45	797896	277	277	90824	3424077	24625	40	6520	140	20.4	42.95	2 + 3 part
APC_3355	66	1414420	30	30	737095	4648031	359143	5	98799	15	29.6	39.09	1 + 0 part
APC_3356	104	971677	62	91	507526	4133448	120175	9	59966	27	0	40.15	1 + 10 part
APC_3358	44	538885	56	73	606347	4818863	375331	5	94792	16	0	39.06	2 + 8 part
APC_3426	38	285509	69	90	505034	3366063	97269	10	33266	32	0	43.15	3 + 4 part
APC_3495	66	1345440	106	106	194214	4070828	66243	18	22421	56	9.7	40.18	1 + 3 part
APC_3691	36	377423	161	211	200380	4383601	55278	26	17890	81	0	39.17	2 + 4 part
APC_3694	58	1192024	139	139	200706	4052885	59830	21	17126	68	29.3	40.33	2 + 1 part
APC_3893	39	475668	93	131	907474	4981760	179630	8	47957	27	0	40.27	9 + 0 part
APC_3894	162	1546768	59	142	463852	4198257	165631	8	68025	25	0	40.12	3 + 8 part
APC_3895	108	1028782	57	98	463852	4198012	199796	7	67319	25	0	40.17	3 + 8 part
APC_3896	200	1850022	59	165	463852	4200073	165631	8	68018	25	0	40.12	3 + 8 part
APC_3897	19	397496	60	60	248651	3785755	89215	15	37485	36	6.4	64.79	3 + 0 part
APC_3898	25	527282	64	64	270340	3546622	99416	11	28065	35	5.9	69.42	3 + 0 part
APC_3900	54	882340	27	27	821418	3653572	454295	3	110074	9	0.0	45.25	3 + 1 part
APC_3901	24	484604	151	29	327488	3490274	49487	23	63145	18	8.5	42.97	3 + 0 part
APC_3903	48	1883716	39	106	510782	6972545	208333	8	45372	33	8.9	46.47	3 + 0 part
APC_3906	54	805188	76	76	268454	3481361	97663	11	26514	36	0.0	47.54	5 + 1 part
APC_3907	133	1267239	45	91	733003	4093022	185492	7	68202	20	0	40.17	2 + 6 part
APC_4015	42	673554	31	31	592665	3611903	409097	4	65066	13	0	45.33	5 + 1 part
APC_4016	4577	123927864	21	21	1522897	3640984	727438	2	185382	5	0	45.26	4 + 1 part
APC_4017	117	1467530	102	165	638108	4980535	191628	8	43060	29	0	40.27	8 + 0 part
APC_4018	34	529504	124	124	113067	2719444	36183	23	12547	71	8.4	70.57	1 + 1 part

continued

Table S1b continued:

APC_4021	51	705412	241	241	70532	2424984	16866	44	5323	144	8.4	73.10	1 + 1 part
APC_4022	25	516718	107	107	185296	3513354	51335	18	18227	62	19.1	71.40	4 + 0 part
APC_4026	67	1373130	106	106	274307	4146085	92837	14	21224	46	17.3	40.20	1 + 4 part

Table S1c. Overview of genome assembly checkM metrics.

sample	marker lineage	checkM	
		completeness (%)	contamination (%)
5A.1	o__Pseudomonadales (UID4679)	99.53	0.62
APC_3213	c__Gammaproteobacteria (UID4787)	100	0.34
APC_3215	c__Gammaproteobacteria (UID4787)	100	0.83
APC_3218	c__Gammaproteobacteria (UID4787)	100	0.68
APC_3224	c__Gammaproteobacteria (UID4787)	99.93	0.34
APC_3227	c__Gammaproteobacteria (UID4787)	100	0.68
APC_3250	c__Gammaproteobacteria (UID4787)	100	0.49
APC_3279	o__Pseudomonadales (UID4679)	99.95	1.41
APC_3281	o__Pseudomonadales (UID4679)	99.95	0.55
APC_3343	s__algicola (UID2846)	99.67	0.33
APC_3350	o__Pseudomonadales (UID4679)	99.4	0.67
APC_3355	c__Gammaproteobacteria (UID4787)	100	0.17
APC_3356	c__Gammaproteobacteria (UID4787)	100	0.34
APC_3358	c__Gammaproteobacteria (UID4787)	100	1.37
APC_3426	o__Pseudomonadales (UID4679)	99.63	0.27
APC_3495	c__Gammaproteobacteria (UID4787)	100	0.34
APC_3691	c__Gammaproteobacteria (UID4787)	100	0.56
APC_3694	c__Gammaproteobacteria (UID4787)	100	0.76
APC_3893	c__Gammaproteobacteria (UID4787)	100	0.42
APC_3894	c__Gammaproteobacteria (UID4787)	100	1.01
APC_3895	c__Gammaproteobacteria (UID4787)	100	1.16
APC_3896	c__Gammaproteobacteria (UID4787)	100	0.34
APC_3897	f__Micrococcaceae (UID1623)	99.31	1.97
APC_3898	o__Actinomycetales (UID1593)	100	1.26
APC_3900	c__Bacilli (UID259)	99.34	0
APC_3901	o__Actinomycetales (UID1593)	99.49	0
APC_3903	o__Actinomycetales (UID1815)	99.94	1.29
APC_3906	c__Bacilli (UID259)	99.34	0.66
APC_3907	c__Gammaproteobacteria (UID4787)	100	0.87
APC_4015	c__Bacilli (UID259)	99.34	0
APC_4016	c__Bacilli (UID259)	99.34	0
APC_4017	c__Gammaproteobacteria (UID4787)	100	0.39
APC_4018	o__Actinomycetales (UID1530)	98.68	0
APC_4021	f__Micrococcaceae (UID1623)	98	0.23
APC_4022	o__Actinomycetales (UID1593)	98.99	0.76
APC_4026	c__Gammaproteobacteria (UID4787)	100	0.34

Table S2: Taxonomic descriptions, based on the 16S rRNA gene sequence, of the strains selected for whole-genome sequencing. Sequences were searched against the SILVA rRNA gene project database.

strain ID	phylum	class	order	family	genus	closest reference hit(s) according to SILVA database (% identity), in order of score where multiple hits are given
APC 3279	Pseudomonadota	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	<i>Psychrobacter</i>	<i>Psychrobacter adeliensis</i> (99.6), <i>P. nivimaris</i> (99.9)
APC 3691	Pseudomonadota	Gammaproteobacteria	Alteromonadales	<i>Pseudoalteromonadaceae</i>	<i>Pseudoalteromonas</i>	<i>Pseudoalteromonas haloplanktis</i> TB13 (99.87)
APC 4017	Pseudomonadota	Gammaproteobacteria	Alteromonadales	<i>Pseudoalteromonadaceae</i>	<i>Pseudoalteromonas</i>	<i>Pseudoalteromonas flavipulchra</i> NCIMB 2033 (97.27) , <i>Pseudoalteromonas agarivorans</i> S816 (97.33)
APC 3893	Pseudomonadota	Gammaproteobacteria	Alteromonadales	<i>Pseudoalteromonadaceae</i>	<i>Pseudoalteromonas</i>	<i>Pseudoalteromonas agarivorans</i> S816 (98.76) , <i>Pseudoalteromonas flavipulchra</i> NCIMB 2033 (98.69) , <i>Pseudoalteromonas haloplanktis</i> ANT/505 (98.69)
APC 3343	Bacteroidota	Flavobacteriia	Flavobacteriales	<i>Flavobacteriaceae</i>	<i>Winogradskyella</i>	<i>W. thalassocola</i> (97.8)
APC 3358	Pseudomonadota	Gammaproteobacteria	Alteromonadales	<i>Pseudoalteromonadaceae</i>	<i>Pseudoalteromonas</i>	<i>Pseudoalteromonas citrea</i> (99.73), <i>Pseudoalteromonas haloplanktis</i> TAC125 (99.59)
APC 3426	Pseudomonadota	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	<i>Psychrobacter</i>	<i>Psychrobacter arcticus</i> 273-4 (98.83), <i>Psychrobacter maritimus</i> (99.46)
APC 3213	Pseudomonadota	Gammaproteobacteria	Alteromonadales	<i>Pseudoalteromonadaceae</i>	<i>Pseudoalteromonas</i>	<i>Pseudoalteromonas agarivorans</i> S816 (99.67)
APC 3215	Pseudomonadota	Gammaproteobacteria	Alteromonadales	<i>Pseudoalteromonadaceae</i>	<i>Pseudoalteromonas</i>	<i>Pseudoalteromonas atlantica</i> TB41 (100)
APC 3218	Pseudomonadota	Gammaproteobacteria	Alteromonadales	<i>Pseudoalteromonadaceae</i>	<i>Pseudoalteromonas</i>	<i>Pseudoalteromonas flavipulchra</i> NCIMB 2033 (99.8), <i>Pseudoalteromonas issachenkonii</i> (99.8)
APC 3227	Pseudomonadota	Gammaproteobacteria	Alteromonadales	<i>Pseudoalteromonadaceae</i>	<i>Pseudoalteromonas</i>	<i>Pseudoalteromonas flavipulchra</i> NCIMB 2033 (99.7), <i>Pseudoalteromonas issachenkonii</i> (99.7)
APC 3894	Pseudomonadota	Gammaproteobacteria	Alteromonadales	<i>Pseudoalteromonadaceae</i>	<i>Pseudoalteromonas</i>	<i>Pseudoalteromonas agarivorans</i> S816 (99.41), <i>Pseudoalteromonas haloplanktis</i> ANT/505 (99.48)
APC 3895	Pseudomonadota	Gammaproteobacteria	Alteromonadales	<i>Pseudoalteromonadaceae</i>	<i>Pseudoalteromonas</i>	<i>Pseudoalteromonas flavipulchra</i> NCIMB 2033 (99.72), <i>Pseudoalteromonas issachenkonii</i> (99.72), <i>Pseudoalteromonas haloplanktis</i> ANT/505 (99.71)
APC 3250	Pseudomonadota	Gammaproteobacteria	Alteromonadales	<i>Pseudoalteromonadaceae</i>	<i>Pseudoalteromonas</i>	<i>Pseudoalteromonas atlantica</i> TB41 (100)

continued

Table S2 continued

APC 3356	Pseudomonadota	Gammaproteobacteria	Alteromonadales	<i>Pseudoalteromonadaceae</i>	<i>Pseudoalteromonas</i>	<i>Pseudoalteromonas atlantica</i> TB41 (99.93)
APC 3907	Pseudomonadota	Gammaproteobacteria	Alteromonadales	<i>Pseudoalteromonadaceae</i>	<i>Pseudoalteromonas</i>	<i>Pseudoalteromonas atlantica</i> TB41 (99.93)
APC 3896	Pseudomonadota	Gammaproteobacteria	Alteromonadales	<i>Pseudoalteromonadaceae</i>	<i>Pseudoalteromonas</i>	<i>Pseudoalteromonas flavipulchra</i> NCIMB 2033 (99.72), <i>Pseudoalteromonas issachenkonii</i> (99.72)
APC 4016	Bacillota	Bacilli	Bacillales	<i>Planococcaceae</i>	<i>Planococcus</i>	<i>Planococcus okeanokoites</i> (99.34)
APC 3281	Pseudomonadota	Gammaproteobacteria	Pseudomonadales	<i>Moraxellaceae</i>	<i>Psychrobacter</i>	<i>Psychrobacter adeliensis</i> (99.6), <i>Psychrobacter nivimaris</i> (99.93)
APC 3900	Bacillota	Bacilli	Bacillales	<i>Planococcaceae</i>	<i>Planococcus</i>	<i>Planomicrobium glaciei</i> CHR43 (99.16)
APC 3901	Actinomycetota	Actinomycetia	Micrococcales	<i>Microbacteriaceae</i>	<i>Microbacterium</i>	<i>Microbacterium foliorum</i> (99.6), <i>Microbacterium oxydans</i> (99.14)
APC 3903	Actinomycetota	Actinomycetia	Mycobacteriales	<i>Nocardiaceae</i>	<i>Rhodococcus</i>	<i>Rhodococcus erythropolis</i> CCM2595 (100)
APC 3906	Bacillota	Bacilli	Bacillales	<i>Planococcaceae</i>	<i>Planococcus</i>	<i>Planomicrobium koreense</i> (100)
5A.1	Pseudomonadota	Gammaproteobacteria	Pseudomonadales	<i>Moraxellaceae</i>	<i>Psychrobacter</i>	<i>Psychrobacter arcticus</i> 273-4 (99.15)
APC 3350	Pseudomonadota	Gammaproteobacteria	Pseudomonadales	<i>Moraxellaceae</i>	<i>Psychrobacter</i>	<i>Psychrobacter arcticus</i> 273-4 (99.86)
APC 4018	Actinomycetota	Actinomycetia	Micrococcales	<i>Micrococcaceae</i>	<i>Kocuria</i>	<i>Kocuria rhizophila</i> P7-4 (100)
APC 4021	Actinomycetota	Actinomycetia	Micrococcales	<i>Micrococcaceae</i>	<i>Micrococcus</i>	<i>Micrococcus luteus</i> NCTC 2665 (99.6), <i>Micrococcus yunnanensis</i> (99.34)
APC 3495	Pseudomonadota	Gammaproteobacteria	Alteromonadales	<i>Pseudoalteromonadaceae</i>	<i>Pseudoalteromonas</i>	<i>Pseudoalteromonas atlantica</i> TB41 (99.93), <i>Pseudoalteromonas haloplanktis</i> ANT/505 (99.15)
APC 3355	Pseudomonadota	Gammaproteobacteria	Alteromonadales	<i>Pseudoalteromonadaceae</i>	<i>Pseudoalteromonas</i>	<i>Pseudoalteromonas haloplanktis</i> ANT/505 (100)
APC 4026	Pseudomonadota	Gammaproteobacteria	Alteromonadales	<i>Pseudoalteromonadaceae</i>	<i>Pseudoalteromonas</i>	<i>Pseudoalteromonas atlantica</i> TB41 (99.93)
APC 4022	Actinomycetota	Actinomycetia	Micrococcales	<i>Microbacteriaceae</i>	<i>Curtobacterium</i>	<i>Curtobacterium herbarum</i> (99.74)
APC 3694	Bacillota	Bacilli	Bacillales	<i>Staphylococcaceae</i>	<i>Staphylococcus</i>	<i>Pseudoalteromonas agarivorans</i> S816 (99.8), <i>Pseudoalteromonas haloplanktis</i> ANT/505 (99.74)

continued

Table S2 continued

APC 3224	Pseudomonadota	Gammaproteobacteria	Alteromonadales	<i>Pseudoalteromonadaceae</i>	<i>Pseudoalteromonas</i>	<i>Pseudoalteromonas atlantica</i> TB41 (99.9), <i>Pseudoalteromonas agarivorans</i> S816 (99.7)
APC 3897	Actinomycetota	Actinomycetia	Micrococcales	<i>Micrococcaceae</i>	<i>Arthrobacter</i>	<i>Arthrobacter luteolus</i> (99.01)
APC 3898	Actinomycetota	Actinomycetia	Micrococcales	<i>Microbacteriaceae</i>	<i>Microbacterium</i>	<i>Microbacterium lacus</i> (99.46)
APC 4015	Bacillota	Bacilli	Bacillales	<i>Planococcaceae</i>	<i>Planococcus</i>	<i>Planomicrobium glaciei</i> CHR43 (99.16)

Table S3: Resulting statistics from the pairwise alignments of precursor peptide sequences using EMBOSS Needle.

query	subject	matrix	length	identity		similarity		gaps	score
				matches	%	matches	%		
Planococcus Pep1	planococcus_Pep2	BLOSSUM50	73	37/73	50.7	51/73	59.9	6/73	228.5
	cerecidin A1A6 [<i>B. cereus</i>]	BLOSSUM50	79	33/79	41.8	47/79	59.5	12/79	206.5
	cerecidin A7 [<i>B. cereus</i>]	BLOSSUM50	83	35/83	42.2	47/83	56.6	20/83	214
	cytolysin Ls [<i>E. faecalis</i>]	BLOSSUM50	72	24/72	33.3	34/72	47.2	12/72	147
	carnolysin A2 [<i>C. maltaromaticum</i>]	BLOSSUM50	75	18/75	24	35/75	46.7	22/75	110.5
	cytolysin LI [<i>E. faecalis</i>]	BLOSSUM50	87	24/87	27.6	37/87	42.5	37/87	141.5
	carnolysin A1 [<i>C. maltaromaticum</i>]	BLOSSUM50	88	19/88	21.6	35/88	39.8	34/88	110.5
	gardimycin [<i>A. liguriae</i>]	BLOSSUM50	82	15/82	18.3	29/82	37.8	31/82	47
	mersacidin [<i>B. amyloliquifaciens</i>]	BLOSSUM50	69	19/69	27.5	33/69	47.8	1/69	98
Planococcus Pep2	cerecidin A7 [<i>B. cereus</i>]	BLOSSUM50	78	38/78	48.7	54/78	69.2	8/78	274
	cerecidin A1A6 [<i>B. cereus</i>]	BLOSSUM50	78	39/78	50	54/78	69.2	8/78	279
	cytolysin Ls [<i>E. faecalis</i>]	BLOSSUM50	76	30/76	39.5	40/76	52.6	18/76	162.5
	carnolysin A2 [<i>C. maltaromaticum</i>]	BLOSSUM50	77	25/77	32.5	37/77	48.1	24/77	133.5
	cytolysin LI [<i>E. faecalis</i>]	BLOSSUM50	89	25/89	41.6	37/89	41.6	39/89	127.5
	carnolysin A1 [<i>C. maltaromaticum</i>]	BLOSSUM50	90	22/90	24.4	36.9	40	36/90	106.5
	gardimycin [<i>A. liguriae</i>]	BLOSSUM50	74	23/74	31.1	35/74	47.3	13/74	115.5
	mersacidin [<i>B. amyloliquifaciens</i>]	BLOSSUM40	75	24/75	32	40/75	53.3	11/75	150.5
Arthrobacter APC_3897 peptide	enterocin AS48 [<i>E. faecalis</i>]	BLOSSUM60	108	77/108	91	108	84.3	6/108	351
	amylocyclicin [<i>B. amyloliquifaciens</i>]	BLOSSUM60	120	34/120	56	120	46.7	23/120	104.5
	circularin_A [<i>C. beierinckii</i>]	BLOSSUM60	105	20/105	19	38/105	36.2	33/105	81
	enterocin_nkr_5_3b [<i>E. faecium</i>]	BLOSSUM50	117	27/117	23.1	45/117	38.5	42/117	124.5
	butyrivibriocin_AR10 [<i>B. fibriosolvens</i>]	BLOSSUM50	109	27/109	24.8	38/109	34.9	33/109	75
Curtobacterium curto_Pep1	curto_Pep2	BLOSSUM40	139	39/139	28.1	51/139	35.7	58/139	174
	lactococcin_972 [<i>L. lactis</i>]	BLOSSUM40	129	28/129	21.7	39/129	30.2	46/129	118.5

continued

Table S3 continued:

	mutacin_2 [<i>S. mutans</i>]	BLOSSUM35	124	12/124	9.7	24/124	19.4	74/124	52
<i>Curtobacterium</i> curto_Pep2	lactococcin_972 [<i>L. lactis</i>]	BLOSSUM35	100	34/100	34	49/100	49	12/100	188
	mutacin_2 [<i>S. mutans</i>]	BLOSSUM35	99	15/99	15.2	26/99	26.3	46/99	77.5

Chapter 3

***Winogradskyella bathintestinalis* sp.
nov., Isolated from the Intestine of the
Deep-sea Loosejaw Dragonfish,
*Malacosteus niger***

Abstract

A novel bacterial strain, APC 3343^T, was isolated from the intestine of a deep-sea loosejaw dragon fish, *Malacosteus niger*, caught at a depth of 1000m in the northern Atlantic Ocean. Cells were aerobic, rod-shaped, yellow/orange-pigmented, non-motile and Gram-negative. The temperature, pH and NaCl growth ranges were 4–30 °C (optimum 25 °C), 5.5–10 (pH 7.5) and 0.5–8 % (2–4 %), respectively. Phylogenetic analysis based on 16S rRNA gene sequences showed that strain APC 3343^T was most closely related to members of the genus *Winogradskyella*, with the most closely related type strains being *Winogradskyella algae* Kr9-9^T (98.46 % identity), *Winogradskyella damuponensis* F081-2^T (98.07 %), *Winogradskyella eximia* KMM 3944^T (97.93 %), *Winogradskyella litoriviva* KMM 6491^T (97.79 %) and *Winogradskyella endarachnes* HL2-2^T (97.79 %). Major fatty acids (>10 % of total) were iso-C_{16:0} 3-OH, iso-C_{15:0}, anteiso-C_{15:0} and iso-C_{17:0} 3-OH. The predominant respiratory quinone was menaquinone-6 (MK-6). Polar lipids were phosphatidylethanolamine, three unknown aminolipids and eight unknown lipids. The draft genome sequence was 3.8 Mb in length with a G+C content of 33.43 mol%. Based on the phenotypic characteristics and phylogenetic analysis, strain APC 3343^T is deemed to be a novel species of the genus *Winogradskyella*, and for which the name *Winogradskyella bathintestinalis* sp. nov. is proposed. The type strain of this species is APC 3343^T (= DSM 115832 = NCIMB 15464).

Author Notes

The GenBank accession numbers for the 16S rRNA gene sequence and the genome sequence of strain APC 3343^T are OP920974 and JASDDK000000000, respectively.

Abbreviations

ANI, average nucleotide identity; BGC, biosynthetic gene cluster; dDDH, digital DNA-DNA hybridisation; MA, marine agar; MB, marine broth; gDNA, genomic DNA; OGRI, overall genome related index; ONPG, orthonitrophenyl-β-D-galactopyranoside; PNPG, paranitrophenyl-β-D-galactopyranoside

3.1. Introduction

The genus *Winogradskyella* was first proposed by Nedashkovskaya et al. (2005) as a novel member of the family *Flavobacteriaceae* and has subsequently been amended several times (Ivanova et al., 2010, Yoon et al., 2011, Nedashkovskaya et al., 2012). Members of the genus are Gram-stain-negative and produce non-diffusible yellow or orange pigments. They are usually aerobic rod-shaped cells, but facultative anaerobic and coccoid cells have also been found (Begum et al., 2013, Alexandre-Colomo et al., 2020). To date, there are 46 type strains within the genus *Winogradskyella* (<https://lpsn.dsmz.de/genus/winogradskyella>).

Winogradskyella species have been isolated from a wide variety of marine habitats such as seawater (Nedashkovskaya et al., 2015) and sediment (Wang et al., 2018), mostly in association with algae (Nedashkovskaya et al., 2005, Nedashkovskaya et al., 2012, Kurilenko et al., 2019, Xu et al., 2020) and marine invertebrates (Ivanova et al., 2010, Park et al., 2015, Franco et al., 2018). Culturable strains have also been recovered from the gills of marine fish (Valdenegro-Vega et al., 2013). Most recently, 16S rRNA gene amplicons belonging to members of the genus *Winogradskyella* have been detected in human faecal samples, yet no culturable representatives have been isolated (Magzal et al., 2022). Furthermore, *Winogradskyella* are of increasing biotechnological interest due to their antibiofilm activity (Rizzo et al., 2021, Dash et al., 2009), exopolysaccharide (EPS) production (Caruso et al., 2018), heavy metal resistance (Caruso et al., 2018) and bioremediation potential (Wang et al., 2014, Pira et al., 2022).

There are suggestions that some *Winogradskyella* species are extremophilic: many characterised isolates are psychrotrophic, capable of growth below 10 °C (Lee et al., 2013, Nedashkovskaya et al., 2015) and some strains have been shown to be resistant to freeze-thaw cycles (Caruso et al., 2018). *Winogradskyella* gene sequences have been detected in habitats where sub-zero temperatures were recorded (Rapp et al., 2018), however, growth of *Winogradskyella* isolates at sub-zero temperatures (under laboratory conditions) has not yet been reported. To date, *Winogradskyella ouciana* represents the only type strain of this genus to be isolated from a habitat of extreme hydrostatic pressure (seawater from 7500m deep; He et al., 2021). This is significant

in that hydrostatic pressure increases by one atmosphere for every 10m increase in water depth in the ocean (Somero, 1992).

A *Winogradskyella* strain, designated strain APC 3343, was previously isolated from the intestine of a deep-sea dragon fish, *Malacosteus niger*, by our lab group during a study of the antimicrobial potential of deep-sea fish microbiome isolates (Uniacke-Lowe et al., 2023). *Malacosteus niger* (a.k.a. Stoplight loosejaw fish) are true deep-sea dwellers and are rarely found above the mesopelagic zone (500m deep; Harold, 2015). They have specific adaptations to survive in their natural habitat such as unusual visual morphology for detecting specific bioluminescent wavelengths in dark waters (Somiya and Marshall, 1982), and a diet that consists predominantly of zooplankton (specifically copepods) but they are also capable of preying on large fish (Sutton, 2005). It has been reported that Flavobacteriales form part of the core gut microbiome of copepods from subtropical Atlantic and Antarctic waters (Shoemaker and Moisander, 2017, Datta et al., 2018) and may accumulate in the copepod gut through feeding on phytoplankton, on which *Flavobacteriaceae* are abundant (Datta et al., 2018). This perhaps also suggests a dietary route for strain APC 3343^T to the deep-sea fish gut. Furthermore, *Winogradskyella* genomic DNA has been identified in faecal metagenome samples from Antarctic copepods (Oh et al., 2021). However, it is yet to be determined if *Winogradskyella* species are core members of the gut microbiome of deep-sea fish.

In this study, we report the taxonomic characterization of strain APC 3343^T, a novel member of the genus *Winogradskyella*.

3.2. Materials and Methods

3.2.1. Isolate information

Strain APC 3343^T was previously isolated from the intestinal tract of a *Malacosteus niger* fish specimen. The fish was collected by research vessels in surveys in international waters near the Grand Banks of Newfoundland in the north-western Atlantic Ocean (43.282 N 49.121 W) from a depth of approximately 1000m (Uniacke-Lowe et al., 2023). Bacterial isolates were recovered from fish intestinal samples, as previously reported (Collins et al., 2021). In brief, swabs of the intestinal tract were streaked onto Difco™ Marine Agar 2216 (MA) and incubated aerobically for 3 weeks at 4 °C. A yellow/orange-pigmented colony was selected and purified through sub-culturing on MA. The strain was deposited into the APC Culture Collection (APC Microbiome Ireland, Teagasc Food Research Centre, Moorepark, Fermoy, Co. Cork, Ireland) and designated strain APC 3343. Cells were preserved in 35 % (v/v) glycerol suspensions at –80 °C. Strain APC 3343^T was routinely cultured on MA at 20–25 °C. *Winogradskyella* sp. APC 3343^T has been deposited into the National Collection of Industrial, Food and Marine Bacteria (= NCIMB 15464) and the DSMZ-German Collection of Microorganisms and Cell Cultures GmbH (= DSM 115832).

3.2.2. Genome sequencing, assembly and annotation

Strain APC 3343^T was cultured in Difco™ Marine Broth 2216 (MB) for 72hrs at 20 °C. Genomic DNA (gDNA) extraction, whole-genome sequencing (WGS) and annotation was carried out as previously described (Uniacke-Lowe et al., 2023). In brief, the bacterial gDNA was extracted using the GeneJET Genomic DNA Purification Kit (Thermo Scientific, Massachusetts, USA). The gDNA was sequenced and assembled by MicrobesNG (<https://microbesng.uk/>, University of Birmingham) using the Illumina platform and SPAdes *de novo* assembly method, and assembly quality was checked using QUAST. The assembled contigs were submitted to GenBank and annotated upon submission using NCBI's Prokaryotic Annotation Pipeline (PGAP). Additionally, gene annotations were made with RAST and subsystem features determined using SEED (Overbeek et al., 2014, Aziz et al., 2008, Brettin et al., 2015). The presence of secondary metabolite and bacteriocin biosynthetic gene clusters

(BGCs) was queried using antiSMASH and BAGEL4, respectively, as previously reported (Uniacke-Lowe et al., 2023).

3.2.3. 16S rRNA gene phylogeny

The complete 16S rRNA gene sequence (1522 base pairs) was extracted from the whole genome sequence data using barrnap (v.0.9, <https://github.com/tseemann/barrnap>). The 16S sequence was queried using the EzTaxon-e server in the EzBioCloud service (Yoon et al., 2017). FASTA files of the hit sequences were downloaded and imported into MEGA X (Kumar et al., 2018). Multiple alignments with the query sequence were created using Muscle (Edgar, 2004) within MEGA X. Gaps from the 5' and 3' ends were removed manually, and the partial gap deletion (95 % site coverage) parameter was applied. Phylogenetic trees: neighbor-joining, maximum-likelihood and maximum-parsimony were created in MEGA X and inferred using the Kimura-2-parameter method (Kimura, 1980). Tree robustness was guaranteed by bootstrap analyses based on 1000 replicates (Felsenstein, 1985). The tree was rooted by including a sequence from the family *Flavobacteriaceae* as an outlier; namely, the 16S rRNA gene sequence from *Marixanthomonas spongiae* Hn-E44 (NR 179850.1).

3.2.4. Genome phylogeny

For the following pairwise average nucleotide identity (ANI) and ROARY analyses, all available *Winogradskyella* reference genome assemblies (including type strains, n =31) were downloaded from the GenBank database (<https://www.ncbi.nlm.nih.gov/data-hub/genome/>). Prior to analysis using ROARY, the assemblies were annotated using PROKKA (Seemann, 2014) to obtain the GFF3 files. The genomes of *W. algae* and *W. damuponensis* have not yet been included in the TYGS database nor in GenBank, and we were not able to acquire these type strains for this study.

ANI values were calculated between APC 3343^T and the *Winogradskyella* reference assemblies. Pairwise %ANI values were calculated with Pyani (v 0.2.12; Pritchard et al., 2016) using the ANIm method (Richter and Rosselló-Móra, 2009). The DSMZ type (strain) genome server (TYGS) annotation platform (<https://tygs.dsmz.de/>; Meier-Kolthoff and Göker, 2019) was used to calculate digital DNA–DNA hybridization

(dDDH) values between APC 3343^T and related type strains from the database. ROARY (Page et al., 2015) was used to determine and create alignments of the core genes of APC 3343^T and the *Winogradskyella* reference genomes. The genomic assembly from *Marixanthomonas spongiae* Hn-E44 (GCA003095375.1) was also annotated and included as an outlier (ROARY parameters: 90 % BLAST id, 99 % of isolates that must have gene, 100000 cluster limit). RaxML (Stamatakis, 2014) was used to generate a maximum likelihood phylogenetic tree from the core gene alignment, which was then visualised in MEGA X.

3.2.5. Biochemical and phenotypic characterisation

The following tests were performed on strain APC 3343^T and, unless stated otherwise, were carried out at 25 °C under aerobic conditions. Growth at different temperatures was tested at 4, 9, 21, 25, 30, 37, 40 and 44 °C on MA. NaCl range for growth was tested by measuring the optical density at OD_{600nm} in MB with the NaCl concentration adjusted within the range 0 – 16 % w/v (in increments of 1 % from 0–10 % NaCl, and in increments of 2 % from 12–16 % NaCl). The pH range for growth was tested in MB from pH 5 to 11 (in increments of 0.5 from pH 5–9 and increments of 1 from pH 9–11). The pH was adjusted by the addition of 1M HCl and/or 1M NaOH.

Anaerobic growth was assessed on MA incubated in an AnaeroGen anaerobic system (Thermo Scientific Oxoid) for up to 14 days at 21 °C. Colony morphology was assessed after incubating on MA at 25 °C for 3 days. Cells were examined using phase contrast microscopy and scanning electron microscopy (SEM). SEM imaging was carried out by UCD Conway Imaging Core (UCD Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Ireland). Gram staining was carried out according to standard methods. Production of oxidase was tested using oxidase strips (Millipore) according to manufacturer's instructions. Production of catalase was tested by transferring a mass of colonies onto a glass slide and exposing them to 1–2 drops of 3 % hydrogen peroxide and observing for effervescence. Observation for motility was carried out using the hanging drop method according to the method outlined by Bernardet et al. (2002) except that a cavity microscope slide was used. The presence of flexirubin-type pigments was tested using the KOH method as described by Bernardet et al. (2002).

Artificial seawater (ASW) salts base used in the subsequent tests was prepared, as per Kurilenko et al. (2019) and Bruns et al. (2001), as follows: 30 gL⁻¹ NaCl, 5.94 gL⁻¹ MgSO₄·7H₂O, 4.53 gL⁻¹ MgCl₂·6H₂O, 0.64 gL⁻¹ KCl and 1.3 gL⁻¹ CaCl₂·2H₂O. Hydrolysis of DNA was determined by plating on DNase Test Agar (Thermo Scientific) supplemented with ASW salts; following incubation, the DNase plates were flooded with 1M HCl and observed for zones of clearance around colonies. Hydrolysis of casein was tested on agar-medium (1.5 % agar) containing ASW salts supplemented with 10 % skimmed milk powder. Hydrolysis of cellulose was tested on agar-medium containing ASW salts supplemented with 2 gL⁻¹ cellulose and 5 gL⁻¹ peptone; following incubation, the plate was then stained with 1 % Congo Red, washed with 1M NaCl and observed for zones of clearance around colonies. Assimilation of Tween 80 and starch was tested by plating on ASW-based medium containing ASW salts, 5 gL⁻¹ peptone, 1 gL⁻¹ yeast extract, 0.1 gL⁻¹ K₂HPO₄, 15 gL⁻¹ agar and 1 % v/v Tween 80 or 0.2 % w/v starch. Production of hydrogen sulfide was determined using Watman® lead acetate strips (Sigma-Aldrich) suspended above an inoculum of the strain in MB for up to 7 days.

Remaining biochemical tests were carried out using the API 20E and API 20NE kits (bioMérieux SA, Marcy-l'Étoile, France) incubated at 25 °C for up to 5 days, the API ZYM kits, incubated at 25 °C for 18h; and the GenIII Microplate (Biolog) incubated at 25 °C for up to 7 days. Results were recorded every day. Bacterial suspensions were prepared in ASW (3 % NaCl) for the API test kits. For the Gen III microplate, protocol B was used with inoculating fluid B (IFB) supplemented with 2 % NaCl. Due to the mucoid nature of strain APC 3343^T, the inoculum for the IFB was prepared from broth culture. In brief, an overnight culture of strain 3343^T was resuspended and washed in PBS. The PBS cell suspension was then used to inoculate the IFB to the correct cell density.

Analyses of cellular fatty acids, polar lipids and respiratory quinones were carried out by DSMZ (Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH, Braunschweig Germany) For fatty acid and polar lipid analyses, strain APC 3343^T was cultivated in MA at 25 °C. Cellular fatty acids were determined using the Sherlock MIS (MIDI) system (v6.1, TSBA40 method, TSBA6 calculation).

Susceptibility to antibiotics was determined using the disc-diffusion method on MA with commercial antibiotics discs (Oxoid) with the following antibiotics: ampicillin (10 µg), chloramphenicol (30 µg) erythromycin (15 µg), gentamicin (10 µg), kanamycin (30 µg), lincomycin (15 µg), neomycin (30 µg), novobiocin (5 µg), oleandomycin (15 µg), penicillin G (10 U), polymyxin B (300 U), rifampicin (30 µg), streptomycin (10 µg) and tetracycline (30 µg). Zone diameters were measured after 36hrs at 25 °C (aerobic). Due to the lack of CLSI guidelines for interpretation criteria for the genus *Winogradskyella*, antibiotic susceptibility and resistance was interpreted as no growth or growth, respectively, in accordance with previous species descriptions of taxonomic neighbours (Kurilenko et al., 2019, Lee et al., 2013).

3.3. Results

3.3.1. Genomic features and characterisation

The draft genome assembly of strain APC 3343^T was 3,832,957 bp in length with a G+C content of 33.43 %. The draft assembly consisted of 53 contigs. An overview of the general characteristics of the draft genome of strain APC 3343^T is given in Table 1. An overview of the RAST annotation of subsystem features is given in Fig. 1. In total, 978 subsystem features were identified across 25 categories. The largest subsystem category was ‘amino acids and derivatives’ for which 197 features were identified, representing 20 % of the overall features. The second largest category was ‘cofactors, vitamins, prosthetic groups and pigments’ for which 130 (13 %) features were identified; this category included genes associated with riboflavin metabolism and Vitamin B biosynthesis. Some other interesting subsystems in this strain identified by RAST included heavy metal resistance (copper homeostasis x5 features, cobalt-zinc-cadmium resistance x4, resistance to fluoroquinolones x2) and stress response (including cold shock x1, oxidative stress x9). Among the secondary metabolism features were genes associated with plant alkaloid synthesis and auxin biosynthesis. No features associated with motility were identified, which is in agreement with the phenotypic motility test results. It was also noted that more features were dedicated to protein metabolism (103) than carbohydrate metabolism (92).

3.3.2. Secondary metabolite BGC and bacteriocin gene screening

Four secondary metabolite BGCs were identified using antiSMASH: two terpenes, one type-3 polyketide synthase (T3PKS) and one hybrid non-ribosomal peptide synthetase (NRPS) – type-1 polyketide synthase (T1PKS). One of the terpene BGCs was most similar to that of carotenoid, at 28 % similarity. The T3PKS shared 8 % similarity to a beta-lactam BGC in the database from an uncultured organism. The remaining BGCs of APC 3343^T had no identifiable cluster hit within the antiSMASH database. BAGEL4 identified one sactipeptide BGC, though it lacked a core sactipeptide gene. One of the genes in the cluster shared 41.4 % similarity to *moaA*, a gene encoding the radical S-adenosyl-L-methionine (SAM) family protein GTP-3',8-cyclase (uniprot B7ULX8). An overview of the antiSMASH and BAGEL screening results is given in Table S1.

3.3.3. 16S rRNA phylogenetic analysis

The 16S rRNA gene sequence similarity analysis indicated that APC 3343^T was most closely related to members of the genus *Winogradskyella* within the family *Flavobacteriaceae*. The APC 3343^T sequence shared the highest similarity with that of *Winogradskyella algae* Kr9-9^T (98.46 % similarity; Kurilenko et al., 2019) followed by *Winogradskyella damuponensis* F081-2^T (98.07 %; Lee et al., 2013), *Winogradskyella eximia* CECT7946^T (97.33 %; Nedashkovskaya et al., 2005), *Winogradskyella litoriviva* KMM 6491^T (97.79 %; Nedashkovskaya et al., 2015) and *Winogradskyella endarachnes* HL2-2^T (97.79 %; Xu et al., 2020). These closely related strains were selected to be used as reference strains for the comparative phenotypic, biochemical and chemotaxonomic analyses. The 16S rRNA sequence % similarity values between APC 3343^T and the reference strains are below the currently accepted threshold of 98.7 % for species (Chun et al., 2018), indicating that strain 3343^T is a separate species of the genus *Winogradskyella*. In the maximum-likelihood (Fig. 2) and neighbor-joining (Fig. S1) phylogenetic trees, generated from the 16S rRNA gene sequence alignments, APC 3343^T placed closely to *Winogradskyella algae* Kr9-9^T. In the maximum parsimony phylogenetic tree APC 3343^T formed its own clade (Fig. S2).

3.3.4. Genome phylogenetic analysis

The %ANI values between APC 3343^T and the available *Winogradskyella* type strain reference genomes ranged from 83.69 % to 85.03 % (Fig. 3). The ANI alignment coverage ranged from 1 % to 28 % (data not shown). An overview of the dDDH values and G+C content % difference between APC 3343^T and the top TYGS database hits are given in Table S2. The dDDH (d4) values between APC 3343^T and *W. eximia*, *W. litoriviva* and *W. endarachnes* were 23.4 %, 22.6 % and 21.5 %, respectively. ANI and dDDH values of 95-96 % and 70 %, respectively, are considered the general cut-off for species and should be assessed as part of the overall genome related index (OGRI) when the 16S rRNA gene sequence %identity values are ≥ 98.7 % (Chun et al., 2018). The ANI and dDDH values in this study indicate that strain APC 3343^T is a separate species among these *Winogradskyella* reference genomes. ROARY was used also to assess phylogeny based on the alignment of 33 core genes from strain APC 3343^T, the

reference *Winogradskyella* genomes and the outlier, *Marixanthomonas spongiae* Hn-E44. The maximum-likelihood phylogenetic tree generated from this core gene alignment is shown in Fig. S3.

3.3.5. Biochemical and phenotypic characterisation

Strain APC 3343^T was found to be aerobic, Gram-stain-negative and non-motile. Colonies were circular, smooth and convex, measuring 0.8 - 1.2mm after incubating on MA at 25 °C for 3 days (Fig. 4). Colonies produced a pale yellow, non-flexirubin-type pigment and intensified in colour with prolonged incubation time, resulting in colonies turning progressively. Cells were rod-shaped (1–1.5µm x 0.3µm) without flagella (Fig. 5). SEM images of clustering cells revealed the presence of network-like structures on the cell surface of strain APC 3343^T (Fig. 5B). Similar structures have been described for *W. thalassocola* KMM 3907^T, *W. epiphytica* KMM 3906^T and *W. eximia* KMM 3944^T but are not present on other *Winogradskyella* species such as *W. haliclona* (Schellenberg et al., 2017). Such structures are believed to be an adaptation for cellular attachment and aggregate formation (Nedashkovskaya et al., 2005).

Differential characteristics between APC 3343^T and the reference *Winogradskyella* strains (*Winogradskyella algae* Kr9-9T [Kurilenko et al., 2019], *Winogradskyella damuponensis* F081-2T [Lee et al., 2013], *Winogradskyella eximia* CECT7946T [Nedashkovskaya et al., 2005, Nedashkovskaya et al., 2012], *Winogradskyella litoriviva* KMM 6491T [Nedashkovskaya et al., 2015] and *Winogradskyella endarachnes* HL2-2T [Xu et al., 2020]) are provided in Table 2. All data for strain APC 3343^T is from this study. The phenotypic and biochemical data for the reference strains was acquired from the respective literature. In the phenotypic assays, APC 3343^T shared a number of characteristics that are in accordance with *Winogradskyella* species (Nedashkovskaya et al., 2005), such as a requirement for NaCl, alkaline phosphatase, oxidase and catalase and gelatin degradation activities, except strain APC 3343^T did not degrade DNA. The temperature range for growth of strain APC 3343^T, of 4–30 °C (optimum 21–25 °C), was lower than those of the reference *Winogradskyella* strains (Nedashkovskaya et al., 2005, Nedashkovskaya et al., 2012, Nedashkovskaya et al., 2015, Wang et al., 2018, Park et al., 2015, Lee et al.,

2013). The NaCl range for growth of strain APC 3343^T, at 0.5-8 % (optimum 2-4 %), was wider than those of the reference strains, though almost identical to that of *W. litoriviva* KMM 6491^T (0.5-7 %NaCl; Nedashkovskaya et al., 2015). The pH range for growth of strain APC 3343^T was 5.5-10 (optimum 7-8), equal to that of *W. algae* Kr9-9^T (Kurilenko et al., 2019) and *W. litoriviva* KMM 6491^T (Nedashkovskaya et al., 2015). The main differences observed between strain APC 3343^T and the reference strains in the phenotypic assays were in its ability to produce, albeit weakly, lipase and β -glucosidase. Strain APC 3343^T also uniquely showed some β -galactosidase activity with paranitrophenyl- β -D-galactopyranoside as the substrate (API 20NE PNPG assay). It is to be noted that the Biolog Gen III plate results were difficult to interpret due to apparent weak metabolic activity by strain APC 3343^T. This is possibly an indication of the fastidious nature of this strain or an inability to grow on a single carbon source. A second assay using the Gen III microplates and protocol C2 with inoculating fluid C (for fastidious organisms) was also carried out, however, some colour production (oxidation) was observed in the negative control well and therefore these results were not used. From the API ZYM assay results, APC 3343^T was positive for alkaline phosphatase, esterase, esterase lipase, leucine arylamidase, valine arylamidase, cystine arylamidase, acid phosphatase and naphthol-as-bi-phosphohydrolase activity; weakly positive for lipase (C14), α -glucosidase and β -glucosidase activity and negative for trypsin, α -chymotrypsin, α -galactosidase, β -galactosidase (2-naphthyl- β -D-galactopyranoside as substrate), β -glucuronidase, n-acetyl- β -glucosaminidase, mannosidase and fucosidase activity.

The predominant (>10 % of total) cellular fatty acids of strain APC 3343^T were iso-C_{16:0} 3-OH (13.3 %), iso-C_{15:0} (13.1 %), anteiso-C_{15:0} (11.9 %) and iso-C_{17:0} 3-OH (10.6 %). The full list of identified cellular fatty acids in this strain and those of the reference strains are provided in Table 3. APC 3343^T contained a unique fatty acid profile compared to the reference strains, particularly with regard to the hydroxy and saturated fatty acids and summed-in features (Nedashkovskaya et al., 2015, Nedashkovskaya et al., 2012, Wang et al., 2018, Park et al., 2015, Lee et al., 2013). Unusually, a trace amount (<1 %) of an unknown fatty acid of an estimated carbon length (ECL) of 13.565 was identified in strain APC 3343^T. The sole respiratory quinone was menaquinone 6 (MK6), in accordance with members of the genus

Winogradskyella (Nedashkovskaya et al., 2005). The cellular polar lipid profile of APC 3343^T consisted of phosphatidylethanolamine, three unknown aminolipids and eight unknown lipids (Fig. S4) and, when compared to the reference strains, was most similar to that of *W. algae* Kr9-9^T (Kurilenko et al., 2019) except that APC 3343^T contained a higher number of unknown/unidentified lipids (Table S3).

Strain APC 3343^T was susceptible to ampicillin (10 µg), chloramphenicol (30 µg), erythromycin (15 µg), lincomycin (15 µg), novobiocin (5 µg), oleandomycin (15 µg), penicillin G (10 U), rifampicin (30 µg) and tetracycline (30 µg); and resistant to gentamicin (10 µg), kanamycin (30 µg), neomycin (30 µg), polymyxin B (300 U) and streptomycin (10 µg). The antibiotic susceptibility profile of strain 3343^T, including the zone diameter measurements, is given in Table S4. Antibiotic susceptibility and resistance were interpreted as no growth or growth, respectively, in accordance with previous species descriptions of taxonomic neighbours (Kurilenko et al., 2019, Lee et al., 2013) and due to the lack of CLSI guidelines for interpretation criteria for the genus *Winogradskyella*. It was also noted that members of the *Winogradskyella* genus lack standardised protocols for growth conditions from which phenotypic information has been described, such as the use of MA rather than Mueller-Hinton, and variable growth temperatures among species representatives ranging from 28 °C for *W. algae* and 25 °C for *W. damuponensis*, rather than the recommended incubation temperature of 18 °C for psychrophilic *Flavobacteriaceae* (Lee et al., 2013). Only three species within this genus (*W. lutea* [Yoon et al., 2011], *W. flava* [Lee et al., 2017], and *W. damuponensis* [Lee et al., 2013]) reported to have followed CLSI guidelines, with none detailing the break-point values used to assign resistance/susceptibility, and only one (*W. maritima*) stating interpretations were performed according to the CLSI criteria for the family *Enterobacteriaceae* (Kang et al., 2017). This divergence in antimicrobial susceptibility testing highlights a deficiency in the standardisation within this group, however, we have reported our results in line with the existing literature.

Strain APC 3343^T clearly differed from the most closely related strain, *W. algae* Kr9-9^T (Kurilenko et al., 2019), in its inability to reduce nitrate or hydrolyse cellulose, lack of α-chymotrypsin activity and its sensitivity to ampicillin, erythromycin, lincomycin

and tetracycline (as well as the above-mentioned range for growth parameters). These assays may be useful in differentiating these two strains.

Based on the combination of phylogenetic distance, phenotypic and biochemical characteristics, strain APC 3343^T represents a novel species for which we propose the name *Winogradskyella bathintestinalis* sp. nov..

3.4. Description of *Winogradskyella bathintestinalis* sp. nov.

Winogradskyella bathintestinalis (bath.in.tes.ti.na'lis. Gr. masc. adj. *bathys*, deep; N.L. masc. adj. *intestinalis*, pertaining to the intestine; N.L. fem. adj. *bathintestinalis*, of the deep intestine). Cells are Gram-stain-negative, rod-shaped (1–1.5µm x 0.3µm), aerobic, non-motile and non-flagellated. Colonies are pale yellow/orange-pigmented, raised, circular, shiny/mucoid with smooth margins after incubation on MA at 25 °C for 3 days. Cell-associated filaments are produced. Pigments are non-flexirubin-type. Growth occurs at 4–30 °C (optimum 21–25 °C), 0.5–8 % NaCl (optimum 2–4 %) and pH 5.5–10 (optimum pH 7–8). Positive for oxidase and catalase activity. H₂S is weakly produced under aerobic conditions. Positive for hydrolysis of Tween 80, weakly positive for hydrolysis of starch. Agar, casein, cellulose and DNA are not hydrolysed. From the API 20E assay results: positive for gelatinase and weakly positive for tryptophan deaminase activity; negative for activity of β-galactosidase (OPNG assay), arginine dihydrolase (anaerobic), lysine decarboxylase (anaerobic), ornithine decarboxylase (anaerobic) and urease (anaerobic); negative for citrate utilisation and H₂S production (anaerobic) and negative for acid production from D-glucose, D-mannitol, inositol, D-sorbitol, L-rhamnose, D-sucrose, D-melibiose amygdalin, L-arabinose. From the API 20NE assay results: positive for hydrolysis of gelatin and esculin; weakly positive for activity of β-galactosidase (PNPG assay); negative for nitrate reduction, indole production, fermentation of glucose, arginine dihydrolase activity (anaerobic), urease production (anaerobic), and negative for assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, n-acetylglucosamine, D-maltose, potassium gluconate, capric acid, adipic acid, malic acid, trisodium citrate and phenylacetic acid. From the Biolog GenIII assay: positive for oxidation of D-trehalose, sucrose, stachyose, D-melibiose, N-acetyl-β-D-mannosamine, N-acetyl-D-galactosamine, N-acetyl neuraminic acid, L-fucose, L-glutamic acid, acetoacetic acid and acetic acid; weakly positive for oxidation of D-maltose, D-cellobiose, gentobiose, D-turanose, α-D-lactose, α-D-glucose, D-

mannose, D-fructose, D-galactose, D-fucose, L-rhamnose, sorbitol, D-mannitol, D-arabitol, myo-inositol, glycyl-L-proline, glucuronamide, mucic acid and α -keto-glutaric acid; negative for the remaining 40 substrates. The sole cellular respiratory quinone is menaquinone-6 (MK6). The cellular polar lipids are phosphatidylethanolamine, three unknown aminolipids and eight unknown lipids. The major cellular fatty acids are iso-C_{16:0} 3-OH, iso-C_{15:0}, anteiso-C_{15:0} and iso-C_{17:0} 3-OH.

The type strain APC 3343^T (= DSM 115832 = NCIMB 15464) was isolated from the intestinal tract of a deep-sea fish, *Malacosteus niger*, from a depth of approximately 1000m in the Northern Atlantic Ocean (43.282 N 49.121 W). The DNA G+C content of the type strain is 33.43 mol%.

The GenBank accession numbers for the 16S rRNA gene sequence and the genome sequence of strain APC 3343^T are OP920974 and JASDDK000000000, respectively.

3.5. References

- ALEJANDRE-COLOMO, C., VIVER, T., URDIAIN, M., FRANCIS, B., HARDER, J., KÄMPFER, P., AMANN, R. & ROSSELLÓ-MÓRA, R. 2020. Taxonomic study of nine new *Winogradskyella* species occurring in the shallow waters of Helgoland Roads, North Sea. Proposal of *Winogradskyella schleiferi* sp. nov., *Winogradskyella costae* sp. nov., *Winogradskyella helgolandensis* sp. nov., *Winogradskyella vidalii* sp. nov., *Winogradskyella forsetii* sp. nov., *Winogradskyella ludwigii* sp. nov., *Winogradskyella ursingii* sp. nov., *Winogradskyella wichelsiae* sp. nov., and *Candidatus "Winogradskyella atlantica"* sp. nov. *Systematic and Applied Microbiology*, 43, 126128.
- AZIZ, R. K., BARTELS, D., BEST, A. A., DEJONGH, M., DISZ, T., EDWARDS, R. A., FORMSMA, K., GERDES, S., GLASS, E. M., KUBAL, M., MEYER, F., OLSEN, G. J., OLSON, R., OSTERMAN, A. L., OVERBEEK, R. A., MCNEIL, L. K., PAARMANN, D., PACZIAN, T., PARRELLO, B., PUSCH, G. D., REICH, C., STEVENS, R., VASSIEVA, O., VONSTEIN, V., WILKE, A. & ZAGNITKO, O. 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*, 9, 75.
- BEGUM, Z., SRINIVAS, T. N. R., MANASA, P., SAILAJA, B., SUNIL, B., PRASAD, S. & SHIVAJI, S. 2013. *Winogradskyella psychrotolerans* sp. nov., a marine bacterium of the family *Flavobacteriaceae* isolated from Arctic sediment. *International Journal of Systematic and Evolutionary Microbiology*, 63, 1646-1652.
- BERNARDET, J.-F., NAKAGAWA, Y., HOLMES, B. & PROKARYOTES, S. O. T. T. O. F. A. C.-L. B. O. T. I. C. O. S. O. 2002. Proposed minimal standards for describing new taxa of the family *Flavobacteriaceae* and emended description of the family. *International Journal of Systematic and Evolutionary Microbiology*, 52, 1049-1070.
- BRETTIN, T., DAVIS, J. J., DISZ, T., EDWARDS, R. A., GERDES, S., OLSEN, G. J., OLSON, R., OVERBEEK, R., PARRELLO, B., PUSCH, G. D., SHUKLA, M., THOMASON, J. A., 3RD, STEVENS, R., VONSTEIN, V., WATTAM, A. R. & XIA, F. 2015. RASTtk: a modular and extensible implementation of the RAST algorithm for building custom annotation pipelines and annotating batches of genomes. *Scientific Reports*, 5, 8365.
- BRUNS, A., ROHDE, M. & BERTHE-CORTI, L. 2001. *Muricauda ruestringensis* gen. nov., sp. nov., a facultatively anaerobic, appendaged bacterium from German North Sea intertidal sediment. *International journal of systematic and evolutionary microbiology*, 51, 1997-2006.
- CARUSO, C., RIZZO, C., MANGANO, S., POLI, A., DI DONATO, P., FINORE, I., NICOLAUS, B., DI MARCO, G., MICHAUD, L. & LO GIUDICE, A. 2018. Production and Biotechnological Potential of Extracellular Polymeric

- Substances from Sponge-Associated Antarctic Bacteria. *Applied and Environmental Microbiology*, 84.
- CHUN, J., OREN, A., VENTOSA, A., CHRISTENSEN, H., ARAHAL, D. R., DA COSTA, M. S., ROONEY, A. P., YI, H., XU, X.-W., DE MEYER, S. & TRUJILLO, M. E. 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*, 68, 461-466.
- COLLINS, F. W. J., WALSH, C. J., GOMEZ-SALA, B., GUIJARRO-GARCÍA, E., STOKES, D., JAKOBSDÓTTIR, K. B., KRISTJÁNSSON, K., BURNS, F., COTTER, P. D., REA, M. C., HILL, C. & ROSS, R. P. 2021. The microbiome of deep-sea fish reveals new microbial species and a sparsity of antibiotic resistance genes. *Gut Microbes*, 13, 1-13.
- DASH, S., JIN, C., LEE, O. O., XU, Y. & QIAN, P.-Y. 2009. Antibacterial and antilarval-settlement potential and metabolite profiles of novel sponge-associated marine bacteria. *Journal of Industrial Microbiology & Biotechnology*, 36, 1047-1056.
- DATTA, M. S., ALMADA, A. A., BAUMGARTNER, M. F., MINCER, T. J., TARRANT, A. M. & POLZ, M. F. 2018. Inter-individual variability in copepod microbiomes reveals bacterial networks linked to host physiology. *The ISME Journal*, 12, 2103-2113.
- EDGAR, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32, 1792-1797.
- FELSENSTEIN, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, 39, 783-791.
- FRANCO, A., BUSSE, H.-J., SCHUBERT, P., WILKE, T., KÄMPFER, P. & GLAESER, S. P. 2018. *Winogradskyella pocilloporae* sp. nov. isolated from healthy tissue of the coral *Pocillopora damicornis*. *International Journal of Systematic and Evolutionary Microbiology*, 68, 1689-1696.
- HAROLD, A. 2015. *Malacosteus niger*. *The IUCN Red List of Threatened Species*.
- HE, X., LIU, R., LIANG, J., LI, Y., ZHAO, X., RAN, L., AHMAD, W. & ZHANG, X.-H. 2021. *Winogradskyella ouciana* sp. nov., isolated from the hadal seawater of the Mariana Trench. *International Journal of Systematic and Evolutionary Microbiology*, 71.
- IVANOVA, E. P., CHRISTEN, R., GORSHKOVA, N. M., ZHUKOVA, N. V., KURILENKO, V. V., CRAWFORD, R. J. & MIKHAILOV, V. V. 2010. *Winogradskyella exilis* sp. nov., isolated from the starfish *Stellaster equestris*, and emended description of the genus *Winogradskyella*. *International Journal of Systematic and Evolutionary Microbiology*, 60, 1577-1580.
- KANG, H., KIM, H., JOUNG, Y. & JOH, K. 2017. *Winogradskyella maritima* sp. nov., isolated from seawater. *International Journal of Systematic and Evolutionary Microbiology*, 67, 3840-3845.

- KIMURA, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16, 111-120.
- KUMAR, S., STECHER, G., LI, M., KNYAZ, C. & TAMURA, K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, 35, 1547-1549.
- KURILENKO, V. V., ROMANENKO, L. A., ISAEVA, M. P., SVETASHEV, V. I. & MIKHAILOV, V. V. 2019. *Winogradskyella algae* sp. nov., a marine bacterium isolated from the brown alga. *Antonie Van Leeuwenhoek*, 112, 731-739.
- LEE, D.-H., CHO, S. J., KIM, S. M. & LEE, S. B. 2013. *Winogradskyella damuponensis* sp. nov., isolated from seawater. *International Journal of Systematic and Evolutionary Microbiology*, 63, 321-326.
- LEE, J. H., KANG, J. W., SHIN, S. B. & SEONG, C. N. 2017. *Winogradskyella flava* sp. nov., isolated from the brown alga, *Sargassum fulvellum*. *International Journal of Systematic and Evolutionary Microbiology*, 67, 3540-3546.
- MAGZAL, F., SHOCHAT, T., HAIMOV, I., TAMIR, S., ASRAF, K., TUCHNER-ARIELI, M., EVEN, C. & AGMON, M. 2022. Increased physical activity improves gut microbiota composition and reduces short-chain fatty acid concentrations in older adults with insomnia. *Scientific Reports*, 12, 2265.
- MEIER-KOLTHOFF, J. P. & GÖKER, M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nature Communications*, 10, 2182.
- NEDASHKOVSKAYA, O. I., KIM, S. B., HAN, S. K., SNAUWAERT, C., VANCANNEYT, M., SWINGS, J., KIM, K.-O., LYSENKO, A. M., ROHDE, M., FROLOVA, G. M., MIKHAILOV, V. V. & BAE, K. S. 2005. *Winogradskyella thalassocola* gen. nov., sp. nov., *Winogradskyella epiphytica* sp. nov. and *Winogradskyella eximia* sp. nov., marine bacteria of the family *Flavobacteriaceae*. *International Journal of Systematic and Evolutionary Microbiology*, 55, 49-55.
- NEDASHKOVSKAYA, O. I., KUKHLEVSKIY, A. D. & ZHUKOVA, N. V. 2012. *Winogradskyella ulvae* sp. nov., an epiphyte of a Pacific seaweed, and emended descriptions of the genus *Winogradskyella* and *Winogradskyella thalassocola*, *Winogradskyella echinorum*, *Winogradskyella exilis* and *Winogradskyella eximia*. *International Journal of Systematic and Evolutionary Microbiology*, 62, 1450-1456.
- NEDASHKOVSKAYA, O. I., KUKHLEVSKIY, A. D., ZHUKOVA, N. V., KIM, S. J., RHEE, S. K. & MIKHAILOV, V. V. 2015. *Winogradskyella litoriviva* sp. nov., isolated from coastal seawater. *International Journal of Systematic and Evolutionary Microbiology*, 65, 3652-7.
- OH, H. N., RI, M. N., KIM, T., MIN, G.-S., KIM, S. & SUL, W. J. 2021. Gut microbiome alteration of cold-adapted Antarctic copepod *Tigriopus kingsejongensis* with temperature changes and developmental stages.

- OVERBEEK, R., OLSON, R., PUSCH, G. D., OLSEN, G. J., DAVIS, J. J., DISZ, T., EDWARDS, R. A., GERDES, S., PARRELLO, B., SHUKLA, M., VONSTEIN, V., WATTAM, A. R., XIA, F. & STEVENS, R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Res*, 42, D206-14.
- PAGE, A. J., CUMMINS, C. A., HUNT, M., WONG, V. K., REUTER, S., HOLDEN, M. T. G., FOOKES, M., FALUSH, D., KEANE, J. A. & PARKHILL, J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*, 31, 3691-3693.
- PARK, S., PARK, J. M., WON, S. M. & YOON, J. H. 2015. *Winogradskyella crassostreae* sp. nov., isolated from an oyster (*Crassostrea gigas*). *International Journal of Systematic and Evolutionary Microbiology*, 65, 2890-2895.
- PIRA, H., RISDIAN, C., MÜSKEN, M., SCHUPP, P. J. & WINK, J. 2022. *Winogradskyella luteola* sp. nov., *Erythrobacter ani* sp. nov., and *Erythrobacter crassostrea* sp. nov., isolated from the hemolymph of the Pacific Oyster *Crassostrea gigas*. *Archives of Microbiology*, 204, 488.
- PRITCHARD, L., GLOVER, R. H., HUMPHRIS, S., ELPHINSTONE, J. G. & TOTH, I. K. 2016. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Analytical Methods*, 8, 12-24.
- RAPP, J. Z., FERNÁNDEZ-MÉNDEZ, M., BIENHOLD, C. & BOETIUS, A. 2018. Effects of ice-algal aggregate export on the connectivity of bacterial communities in the central Arctic Ocean. *Frontiers in Microbiology*, 9.
- RICHTER, M. & ROSSELLÓ-MÓRA, R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences*, 106, 19126-19131.
- RIZZO, C., ZAMMUTO, V., LO GIUDICE, A., RIZZO, M. G., SPANÒ, A., LAGANÀ, P., MARTINEZ, M., GUGLIELMINO, S. & GUGLIANDOLO, C. 2021. Antibiofilm Activity of Antarctic Sponge-Associated Bacteria against *Pseudomonas aeruginosa* and *Staphylococcus aureus*. *Journal of Marine Science and Engineering*, 9, 243.
- SCHELLENBERG, J., BUSSE, H. J., HARDT, M., SCHUBERT, P., WILKE, T., KÄMPFER, P. & GLAESER, S. P. 2017. *Winogradskyella haliclona* sp. nov., isolated from a marine sponge of the genus *Haliclona*. *International Journal of Systematic and Evolutionary Microbiology*, 67, 4902-4910.
- SEEMANN, T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 30, 2068-9.
- SHOEMAKER, K. M. & MOISANDER, P. H. 2017. Seasonal variation in the copepod gut microbiome in the subtropical North Atlantic Ocean. *Environmental Microbiology*, 19, 3087-3097.
- SOMERO, G. N. 1992. Biochemical ecology of deep-sea animals. *Experientia*, 48, 537-543.

- SOMIYA, H. & MARSHALL, N. B. 1982. 'Yellow lens' eyes of a stomiatoid deep-sea fish, *Malacosteus niger*. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, 215, 481-489.
- STAMATAKIS, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30, 1312-1313.
- SUTTON, T. T. 2005. Trophic ecology of the deep-sea fish *Malacosteus niger* (Pisces: Stomiidae): An enigmatic feeding ecology to facilitate a unique visual system? *Deep Sea Research Part I: Oceanographic Research Papers*, 52, 2065-2076.
- UNIACKE-LOWE, S., COLLINS, F. W. J., HILL, C. & ROSS, R. P. 2023. Bioactivity screening and genomic analysis reveals deep-sea fish microbiome isolates as sources of novel antimicrobials. *Marine Drugs*, 21, 444.
- VALDENEGRO-VEGA, V., NAEEM, S., CARSON, J., BOWMAN, J., TEJEDOR DEL REAL, J. & NOWAK, B. 2013. Culturable microbiota of ranched southern bluefin tuna (*Thunnus maccoyii* Castelnau). *Journal of Applied Microbiology*, 115, 923-932.
- WANG, C., HAN, J.-R., LIU, C.-L. & DU, Z.-J. 2018. *Winogradskyella tangerina* sp. nov., a member of the *Flavobacteriaceae* isolated from coastal sediment. *International Journal of Systematic and Evolutionary Microbiology*, 68, 2832-2837.
- WANG, W., ZHONG, R., SHAN, D. & SHAO, Z. 2014. Indigenous oil-degrading bacteria in crude oil-contaminated seawater of the Yellow sea, China. *Applied Microbiology and Biotechnology*, 98, 7253-7269.
- XU, Y., LI, J., HU, Y., LI, H., PENG, T., ZHONG, M. & HU, Z. 2020. *Winogradskyella endarachnes* sp. nov., a marine bacterium isolated from the brown alga *Endarachne binghamiae*. *International Journal of Systematic and Evolutionary Microbiology*, 71.
- YOON, B.-J., BYUN, H.-D., KIM, J.-Y., LEE, D.-H., KAHNG, H.-Y. & OH, D.-C. 2011. *Winogradskyella lutea* sp. nov., isolated from seawater, and emended description of the genus *Winogradskyella*. *International Journal of Systematic and Evolutionary Microbiology*, 61, 1539-1543.
- YOON, S.-H., HA, S.-M., KWON, S., LIM, J., KIM, Y., SEO, H. & CHUN, J. 2017. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *International Journal of Systematic and Evolutionary Microbiology*, 67, 1613-1617.

3.6. Chapter 3 Figures and Tables

Table 1: General characteristics of the draft genome of strain APC 3343^T.

Genomic feature	
Draft genome size (bp)	3832957
G+C content (mol%)	33.43
N50	1042040
L50	2
Number of contigs	53
Largest contig	1191771
CDS	3371
RNA genes (RAST)	37
rRNA genes (RAST)	1
rRNA genes (barrnap)	6
tRNA genes (RAST)	36
AMR genes (ABRicate)	0

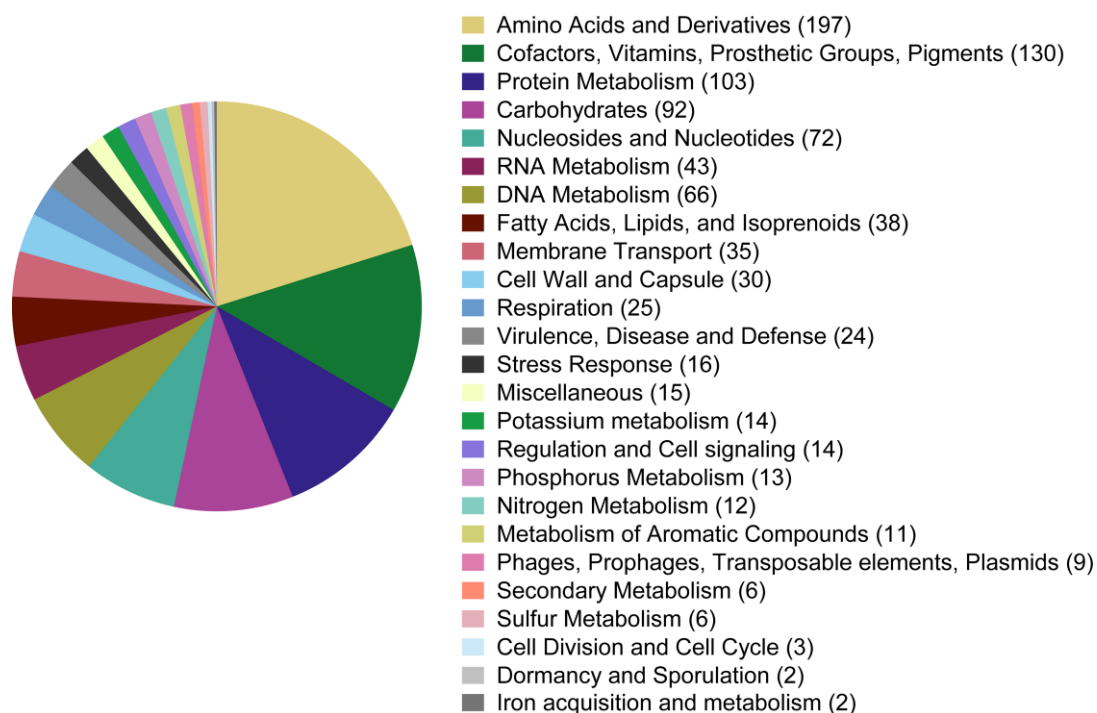


Fig. 1. Overview of RAST annotation of subsystem features from the draft genome of APC 3343^T. The counts of features per subsystem category are given. The total number of subsystem features was 978.

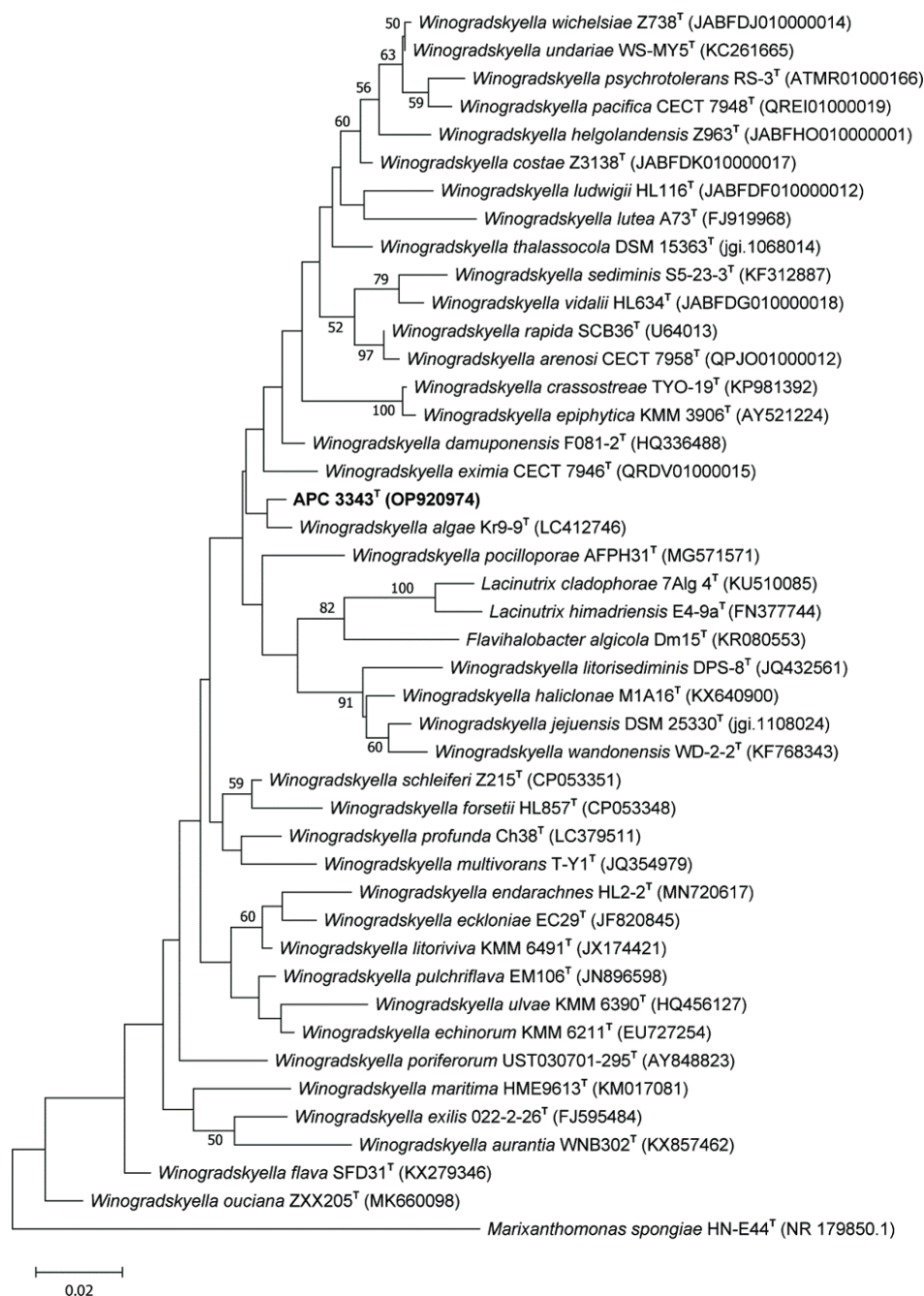


Fig. 2. Maximum-likelihood tree based on 16S rRNA gene sequences showing the relationship of strain APC 3343^T and other members of the genus *Winogradskyella* and related taxa. Distances were calculated based on the Kimura two-parameter model. Bootstrap values as % based on 1000 replications (>50 %) are shown beside each branch node. Bar, 0.02 substitution per nucleotide position.

Fig. 3. Pairwise average nucleotide identity (ANI) of APC 3343^T and all available *Winogradskyella* type strain reference genomes from the GenBank database. ANI values were calculated using Pyani using the ANIm method. Blue cells correspond to an ANI of <95 %, indicating separate species. Red cells correspond to an ANI of >95 %, indicating the same species. The colour intensities lighten as the ANI value approaches 95 %. The dendrograms are created by hierarchal clustering of the pairwise ANI values.

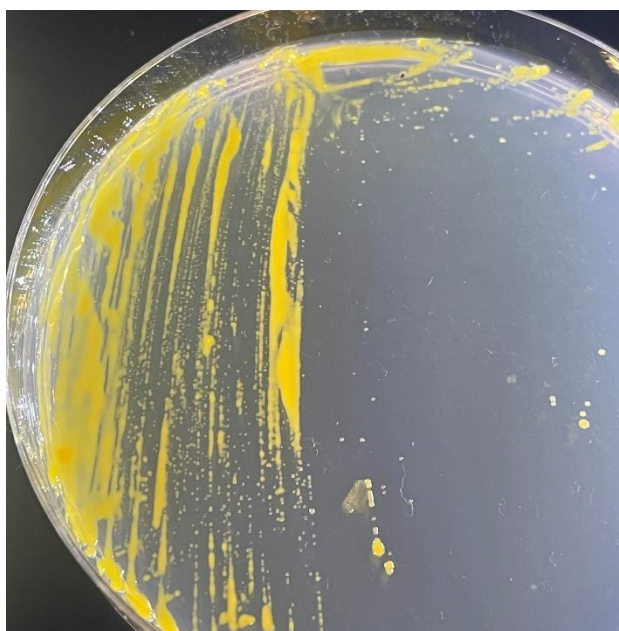


Fig. 4. Colonies of strain APC 3343^T (after 3 days on MA at 25 °C).

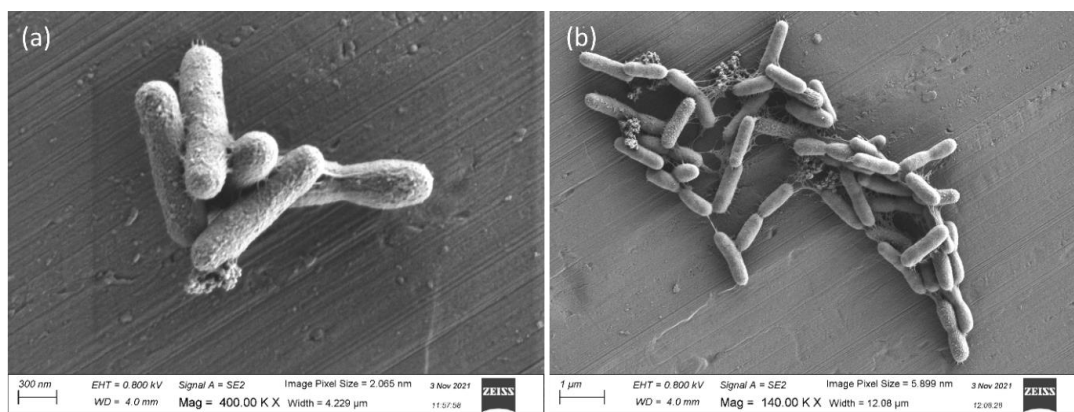


Fig. 5. Scanning electron microscopy image of APC 3343^T at ×400K (a) and ×140K (b).

Table 2. Differential characteristics of isolate APC 3343^T and closely related *Winogradskyella* type strains.

Strains: 1, APC 3343^T (data from this study); 2, *W. algae* Kr9-9^T (data from Kurilenko et al. [2019]); 3, *W. damuponensis* F081-2^T (data from Lee et al. [2013]); 4, *W. eximia* KMM 3944^T (data from Nedashkovskaya et al. [2005] and Nedashkovskaya et al. [2012]); 5, *W. litoriviva* KMM 6491^T (data from Nedashkovskaya et al. [2015]); 6, *W. endarachnes* HL2-2^T (data from Xu et al. [2020]). nd, No data available; +, positive; –, negative; w, weakly positive

Characteristic	1	2	3	4	5	6
Growth at/with:						
37 °C	-	-	-	-	-	+
Temperature range (°C)	4 - 30	7 - 36	4 - 35	4 - 35	4 - 34	20 - 40
Optimum (°C)	21 - 25	28 - 30	25 - 30	21 - 23	25 - 28	30
NaCl range (% w/v)	0.5 - 8	2 - 6	1 - 5	1 - 8	0.5 - 7	0 - 5
pH range	5.5 - 10	5.5 - 10.0	6.0 - 9.5	ND	5.5 - 10	5.5 - 8
Nitrate reduction	-	+	-	-	-	ND
Glucose fermentation	-	-	-	-	+	+
H ₂ S production	W	W	-	-	W	ND
Hydrolysis:						
Aesculin	+	+	+	+	+	ND
Casein	-	-	-	+	-	ND
DNA	-	-	+	-	-	ND
Starch	W	+	-	+	+	ND
Assimilation of:						
Cellobiose	W	ND	+	-	+	+
D-galactose	+	ND	+	-	+	ND
D-glucose	-	-	+	+	+	+
D-mannose	-	-	+	+	+	+
Mannitol	-	-	+	+	-	+
Enzyme activity (API ZYM):						
Esterase	+	+	+	+	W	+
Esterase lipase	+	+	+	+	+	+
Lipase (C14)	W	-	-	-	-	+
Leucine arylamidase	+	+	+	-	+	+
Valine arylamidase	+	+	+	+	+	+
Cystine arylamidase	+	W	+	-	+	+
Trypsin	-	-	-	-	-	+
α-Chymotrypsin	-	+	-	-	+	+
α-Galactosidase	-	-	-	+	-	-

continued

Table 2. Continued:

Characteristic	1	2	3	4	5	6
α -Glucosidase	W	-	-	-	+	+
β -Glucosidase	W	-	-	-	-	-
Susceptibility to antibiotics:						
Ampicillin	+	-	+	-	+	+
Chloramphenicol	+	+	+	+	+	+
Erythromycin	+	-	+	+	+	+
Gentamicin	-	-	-	+	-	-
Kanamycin	-	-	-	+	-	-
Lincomycin	+	-	+	+	+	+
Neomycin	-	-	-	+	-	-
Oleandomycin	+	+	+	-	+	ND
Penicillin G	+	+	+	-	+	ND
Polymyxin B	-	-	-	+	-	-
Rifampcin	+	+	+	+	+	ND
Streptomycin	-	-	-	+	-	-
Tetracycline	+	-	+	ND	+	+

Table 3. Cellular fatty acid compositions (%) of APC 3343^T and closely related type strain members of the genus *Winogradskyella*

Strains: 1, APC 3343^T (data from this study); 2, *W. algae* Kr9-9^T (Kurilenko et al., 2019); 3, *W. damuponensis* F081-2^T (Lee et al., 2013); 4, *W. eximia* KMM 3944^T (Nedashkovskaya et al., 2005); 5, *W. litoriviva* KMM 6491^T (Nedashkovskaya et al., 2015); 6, *W. endarachnes* HL2-2^T (Xu et al., 2020). Major fatty acids (>10 %) are highlighted in bold. Fatty acids amounting <1.0 % for all reference strains are not shown. tr, Trace amount (<1 %); –, not detected. ECL, estimated carbon length.

Fatty acid	1*	2†	3*	4†	5**	6‡
Saturated:						
C _{15:0}	TR	8.3	5.3	7.9	6.4	-
C _{16:0}	-	1.8	TR	-	TR	19.6
C _{17:0} cyclo	-	-	-	2.3	-	-
C _{18:0}	-	-	-	-	-	2.1
Unsaturated:						
C _{15:1}	-	2.7	-	-	-	-
C _{15:1} ω6c	2.8	-	1.9	-	1.4	-
C _{15:1} ω8c	-	-	-	-	1.7	-
C _{16:1}	-	3.2	-	-	-	-
C _{17:1} ω6c	2.8	-	1.1	-	-	-
C _{18:1} ω5c	-	-	2.5	-	-	-
Branched chain:						
iso-C _{12:0}	-	-	-	-	-	1.3
iso-C _{13:0}	-	-	-	-	-	1.3
iso-C _{14:0}	1.4	1.3	1.3	1.4	-	TR
iso-C _{15:0}	13.1	14.4	25.3	25.6	19.4	37.0
iso-C _{15:1}	-	9.9	-	11.4	15.5	-
iso-C _{15:1} G	9.0	-	14.6	-	-	8.9
iso-C _{16:0}	2.2	2.7	1.9	5.7	TR	-
iso-C _{16:1}	-	1.8	-	4.7	-	-
iso-C _{16:1} H	4.3	-	2.2	-	-	-
anteiso-C _{15:0}	11.9	12.7	7.8	4.9	9.5	4.1
anteiso-C _{15:1}	-	-	-	1.6	4.8	-
anteiso-C _{15:1} A	4.4	3.4	3.9	-	-	-
anteiso-C _{17:1}	-	-	-	2.3	-	-
anteiso-C _{17:1} ω9c	2.6	-	-	-	-	-
Hydroxy:						
C _{15:0} 2-OH	3.0	-	2.0	1.0	-	TR
C _{15:0} 3-OH	-	-	1.2	-	TR	-
C _{16:0} 3-OH	-	1.5	TR	-	TR	-
C _{17:0} 2-OH	5.2	-	2.6	1.0	-	TR
C _{18:0} 2-OH	-	-	-	-	-	1.0
iso-C _{15:0} 2-OH	-	1.4	-	-	-	1.9

continued

Table 3. Continued

Fatty acid	1*	2†	3*	4†	5**	6‡
iso-C _{15:0} 3-OH	4.7	3.6	7.6	2.6	9.6	7.8
iso-C _{16:0} 2-OH	-	-	-	-	1.2	-
iso-C _{16:0} 3-OH	13.3	12.9	6.7	3.2	4.6	1.6
iso-C _{17:0} 3-OH	10.6	7.4	9.3	6.7	8.4	8.1
anteiso-C _{15:0} 2-OH	-	1.7	-	-	-	-
anteiso-C _{15:0} 3-OH	-	2.5	-	-	-	-
anteiso-C _{17:0} 3-OH	-	4.5	-	-	4.7	-
Methylated:						
C _{16:0} 10 methyl	-	-	-	6.3	-	-
Other:						
C _{16:1} ω7, iso-C _{15:0} 2-OH	-	-	-	6.1	5.8	-
Unknown	TR (ECL 13.6)	-	-	5.6	-	-
Summed feature 3	7.0 (C _{16:1} ω7c / C _{16:1} ω6c)	-	2.0 (C _{16:1} ω6c and/or C _{16:1} ω7c)	-	-	-
Summed feature 9	1.8 (iso- C _{17:1} ω9c)	-	2.4 (C ₁₀ - methyl and/or iso- C _{17:1} ω9c)	-	-	TR (iso- C _{17:1} ω9c/10- methyl- C _{16:0})

*Analysis carried out at 25 °C.

†Analysis carried out at 28 °C.

‡Analysis carried out at 30 °C.

3.7. Chapter 3 Supplementary Material

Table S1: Overview of *in silico* screening for secondary metabolite BGCs and bacteriocin gene clusters using antiSMASH and BAGEL, respectively, within the genome of strain APC 3343^T.

Query	Hit type	Node	Most similar known / database hit	% Similarity / match	Accession / motif	Database
APC 3343	Sactipeptides	1	GTP 3',8-cyclase (<i>moaA</i>)	41.4	A8MLW5	BAGEL4
	Nrps-T1pks	2	none	-	-	antiSMASH
	T3pks	1	NRP: Betalactam	8	BGC0000350	antiSMASH
	Terpene	4	Carotenoid	28	BGC0000650	antiSMASH
	Terpene	5	none	-	-	antiSMASH

BGC = biosynthetic gene cluster; NRP = non-ribosomal peptide; Nrps = Non-ribosomal peptide synthetase; T1pks = Type one polyketide synthase; T3pks = Type three polyketide synthase.

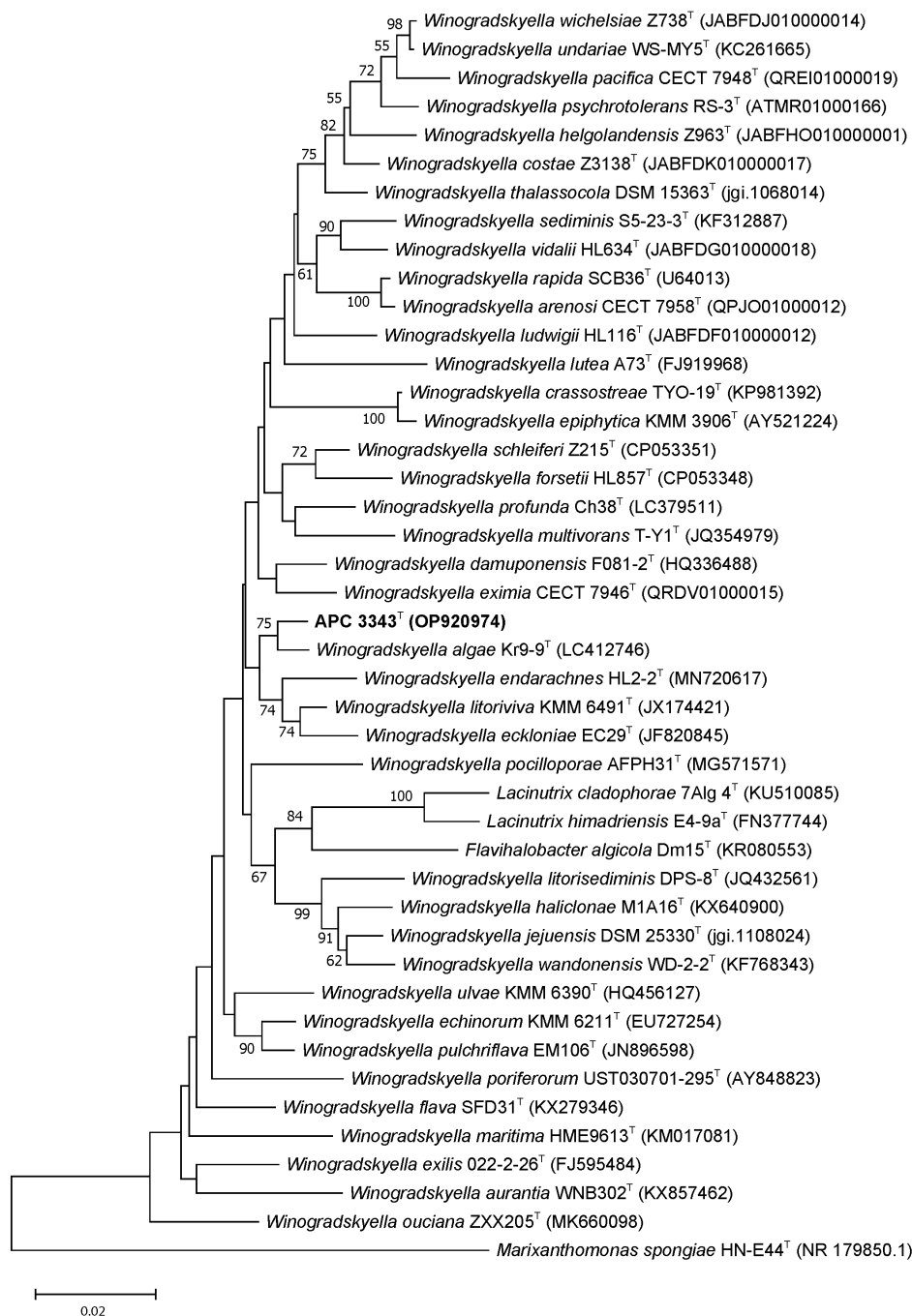


Fig. S1. Neighbor-joining phylogenetic tree based on the 16S rRNA gene sequences of APC 3343^T and closely related hits identified through EzBioCloud. Distances were calculated based on the Kimura-2 parameter model and bootstrap analysis of 1000 replicates, values >50% are given at the branching points.

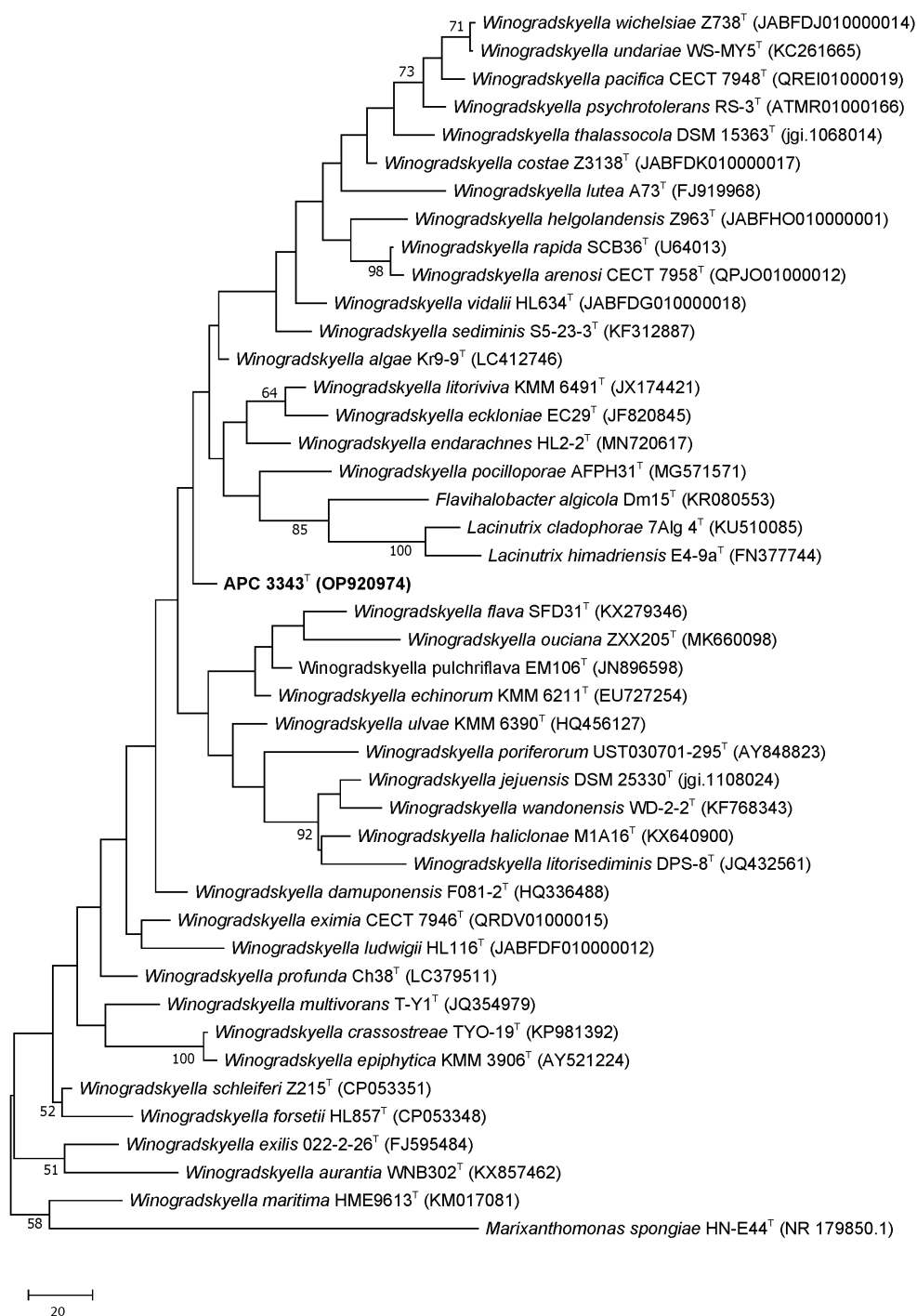


Fig. S2. Maximum-parsimony phylogenetic tree based on the 16S rRNA gene sequences of APC 3343^T and closely related hits identified through EzBioCloud. Distances were calculated based on the Kimura-2 parameter model and bootstrap analysis of 1000 replicates, values >50% are given at the branching points.

Table S2: Pairwise dDDH (d4) values and %G+C content differences calculated between APC 3343^T and closely related database hits from the Type Strain Genome Server (TYGS) database. Genome size (bp) of the hit strain is also given.

Query strain	Subject strain	Genome size (bp)	dDDH (d4, in %)	G+C content difference (in %)
APC 3343	<i>Winogradskyella forsetii</i> HL857	4512651	24.1	1.29
	<i>Winogradskyella schleiferi</i> Z215	4560973	23.8	1.2
	<i>Winogradskyella marina</i> F6397	3974081	23.5	0.16
	<i>Winogradskyella eximia</i> CECT 7946	4235755	23.4	1.09
	<i>Winogradskyella psychrotolerans</i> RS-3	3593001	22.6	0.05
	<i>Winogradskyella litoriviva</i> KMM 6491	3875632	22.6	1.65
	<i>Winogradskyella thalassocola</i> DSM 15363	4571886	22.5	0.15
	<i>Winogradskyella pacifica</i> CECT 7948	4275916	22.3	0.13
	<i>Winogradskyella echinorum</i> KCTC22026(T)	3784537	22.2	1.13
	<i>Winogradskyella echinorum</i> KCTC 22026	3778959	22.2	1.15
	<i>Winogradskyella wichelsiae</i> Z738	3893712	21.5	1.16
	<i>Winogradskyella endarachnes</i> HL2-2	3593001	21.5	1.57
	<i>Winogradskyella eckloniae</i> EC29	3687966	21.3	0.51
	<i>Winogradskyella epiphytica</i> KCTC 12220	3364767	21	0.74
	<i>Winogradskyella arenosi</i> CECT 7958	3666443	20.9	2.24

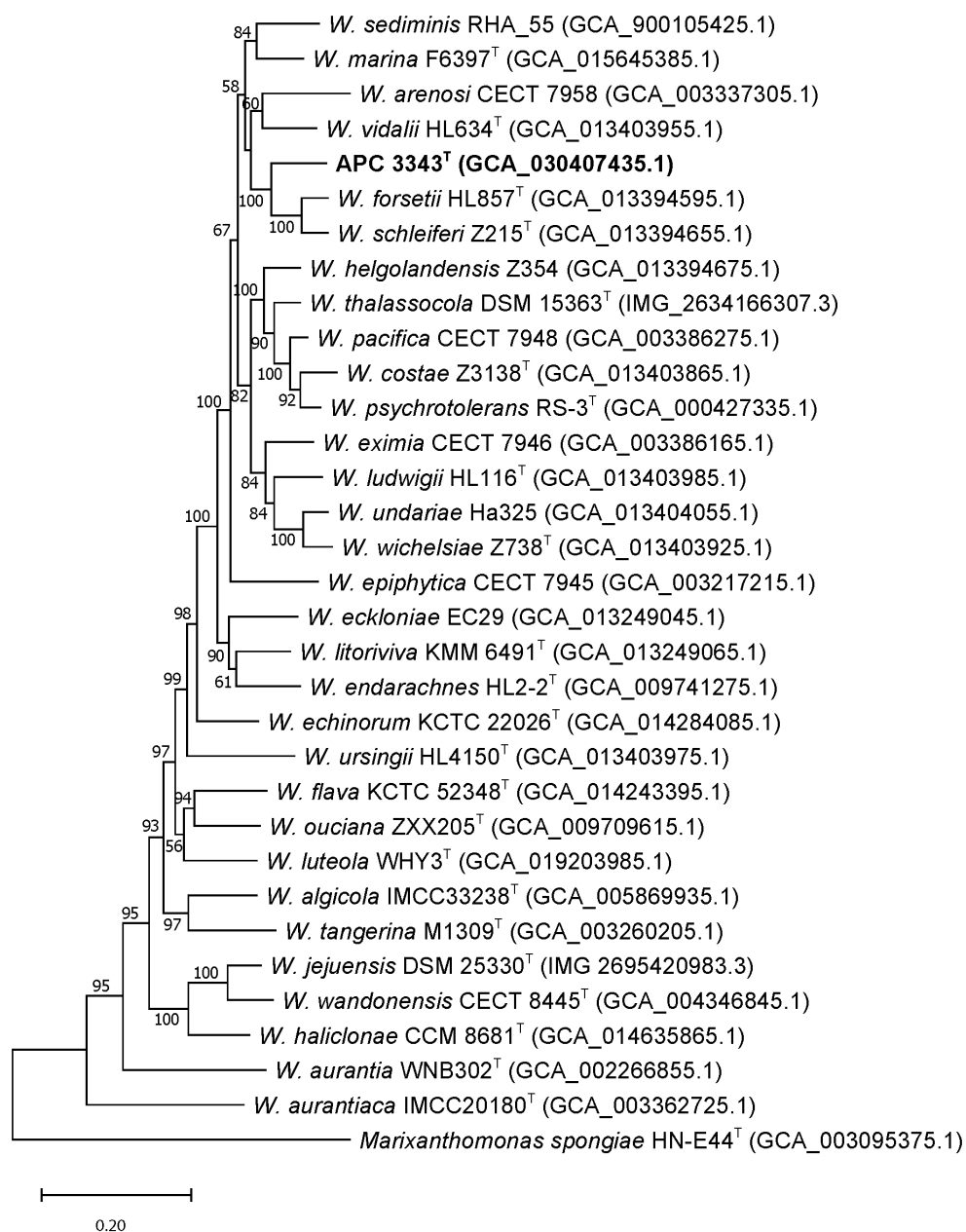


Fig. S3. Resulting maximum-likelihood phylogenetic tree based on the alignment of 33 core genes generated by ROARY from strain APC 3343^T and all available *Winogradskyella* spp. reference genomes. *Marixanthomonas spongiae* HN-E44 was included as an outlier.

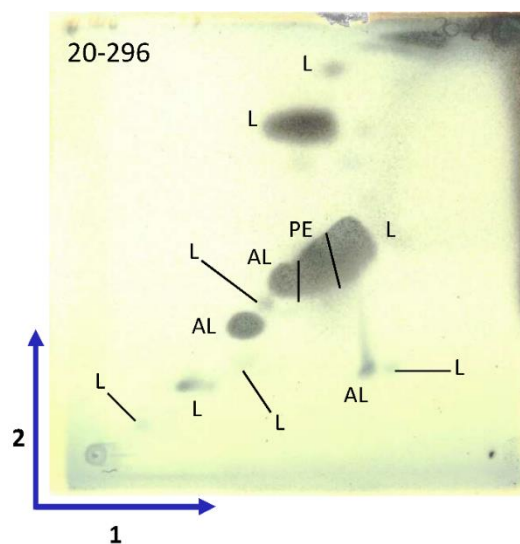


Fig. S4. Polar lipid profile of strain APC 3343^T after separation by two-dimensional thin layer chromatography.

PE = Phosphatidylethanolamine; AL= Aminolipid;
L = Lipid.

Table S3: Cellular polar lipid profiles of strain APC 3343^T and closely related *Winogradskyella* spp.: 1, APC 3343^T (data from this study); 2, *W. algae* Kr9-9^T (data from Kurilenko et al., [2019]); 3, *W. damuponensis* F081-2^T (Lee et al., 2013); 4, *W. eximia* KMM 3944^T (Nedashkovskaya et al., 2012); 5, *W. litoriviva* KMM 6491^T (Nedashkovskaya et al., 2015); 6, *W. endarachnes* HL2-2^T (Xu et al., 2020).

strain	polar lipid profile
1	phosphatidylethanolamine, 3x unknown aminolipids, 8x unknown lipids
2	phosphatidylethanolamine, 3x unidentified aminolipids, 1x unidentified lipid
3	phosphatidylethanolamine, 1x unidentified aminolipid, 1x unidentified phospholipid, 1x unidentified aminophospholipid, 6x unidentified lipids
4	phosphatidylethanolamine, 1x unknown aminolipid
5	phosphatidylethanolamine, 2x unknown aminolipids, 2x unknown lipids
6	phosphatidylethanolamine, 1x unidentified phospholipid, 2x unidentified aminolipids, 5x unknown polar lipids

Table S4: Antibiotic susceptibility profile of strain 3343^T with corresponding zone diameters (mm). Susceptibility and resistance were determined by no growth and growth, respectively.

Antibiotic	Disc content	Zone diameter (mm)	Result
ampicillin	10 µg	14 ±0.2	S
chloramphenicol	30 µg	18 ±1.6	S
erythromycin	15 µg	28 ±2.4	S
gentamicin	10 µg	0 ±0	R
kanamycin	30 µg	0 ±0	R
lincomycin	15 µg	20 ±3.4	S
neomycin	30 µg	0 ±0	R
novobiocin	5 µg	20 ±3.3	S
oleandomycin	15 µg	28 ±2.6	S
penicillin G	10 U	11 ±0.9	S
polymyxin B	300 U	0 ±0	R
rifampicin	30 µg	37 ±3.4	S
streptomycin	10 µg	0 ±0	R
tetracycline	30 µg	14 ±0.5	S

R = resistant; S = susceptible

Chapter 4

***Planococcus notacanthi* sp. nov.,**

Isolated from the Skin of a Deep-sea

Snub-nosed Spiny Eel

Abstract

A novel bacterial strain, APC 4016^T, was previously isolated from the skin of a snub-nosed spiny eel, *Notacanthus chemnitzii*, from a depth of 1000m in the northern Atlantic Ocean. Cells were aerobic, cocci, motile, Gram-positive to Gram-variable staining, and gave rise to orange-pigmented colonies. Growth occurred at 4–40 °C (optimum, 25–28 °C); pH 5.5–12 (optimum, pH 7–7.5), and 0–12 % (w/v) NaCl (optimum, 1 %). 16S rRNA gene phylogenetic analysis confirmed that strain APC 4016^T belonged to the genus *Planococcus* and was most closely related to *Planococcus okeanokoites* IFO 12536^T (98.98 % 16S similarity). However, digital dDDH and average nucleotide identity values between these two strains were low, at 20.1 % and 83.8 %, respectively. Major (>10 %) cellular fatty acids of strain APC 4016^T were iso-C_{14:0}, anteiso-C_{15:0} and C_{16:1}-ω-Alc. The predominant respiratory quinones were menaquinones 5, 6, 7 and 8. The major cellular polar lipids were phosphatidylglycerol, diphosphatidylglycerol and phosphatidylethanolamine, and three unknown lipids were also present. The draft genome sequence is 3.6 Mb with a G+C content of 45.25 mol%. This strain was previously shown to have antimicrobial activity and to encode bacteriocin and secondary metabolite biosynthetic gene clusters.

Based on the phylogenetic analysis and its distinct phenotypic characteristics, strain APC 4016^T is deemed to represent a novel species of the genus *Planococcus*, and for which the name *Planococcus notacanthi* sp. nov. is proposed. The type strain of this species is APC 4016^T (=DSM 115753^T = NCIMB 15463^T).

Author Notes

The GenBank accession numbers for the 16S rRNA gene sequence and the draft genome sequence of strain 4016^T are OQ456996 and JASDCQ000000000, respectively.

Abbreviations

ANI, average nucleotide identity; BGC, biosynthetic gene cluster; dDDH, digital DNA-DNA hybridisation; MA, marine agar; MK, menaquinone

4.1. Introduction

The genus *Planococcus* was first proposed by Migula (1894) in order to distinguish motile marine cocci from the genus *Micrococcus*. It has proven to be a difficult taxonomic group to describe; the genus has undergone several emendations (Nakagawa et al., 1996, Yoon et al., 2010) and members have been frequently reclassified, mainly between *Planococcus* and *Planomicrobium* (Yoon et al., 2001, Dai et al., 2005, Jung et al., 2009). Previously, members of the genera *Planococcus* and *Planomicrobium* were primarily distinguished based on signature nucleotides of the 16S rRNA gene sequence, particularly those at positions 183 and 190 (Rosenberg et al., 2014). However, an extensive demarcation of *Planococcaceae* and related families was recently proposed by Gupta and Patel (2020). They carried out extensive phylogenetic and comparative analysis of core protein sequences from 124 complete genomes from *Caryophanaceae*, *Planococcaceae* and selected *Bacillaceae* species. From their analysis, they proposed the consolidation of the families *Planococcaceae* and *Caryophanaceae* (under the name *Caryophanaceae*). They established that sequences from each species could be reliably grouped into distinctive clades, which were distinguished based on their phylogenetic grouping and on the presence of unique molecular markers ('conserved signature indels'). These conserved signature indels were specific to either *Caryophanaceae/Planococcaceae* members or to clades within this family. They also proposed the creation of a novel genus, *Metaplanococcus*, the reclassification of several species, and that the genera *Planococcus* and *Planomicrobium* be united under the name '*Planococcus*'. The emended genus *Planococcus* now includes most species from the genus *Planomicrobium*. To date, there are 26 valid type strains within the genus *Planococcus* (*Planococcus/Planomicrobium*). Species have been isolated from a wide variety of sources, primarily marine or halophilic environments, such as marine sediment (Yoon et al., 2010, Dai et al., 2005), algal mats (Reddy et al., 2002, Alam et al., 2003) and seafood (Yoon et al., 2001, KOMAGATA, 1985). Others have been found in extreme conditions including sub-zero temperatures (Mykytczuk et al., 2012, Zhang et al., 2009, Mayilraj et al., 2005), deep-sea sediment (Choi et al., 2007) and in polluted soils (Wang et al., 2017, Yang et al., 2020). Members of the genus

Planococcus are of great biotechnological interest because of their robustness and diverse biological activities, such as production of cold-adapted and halotolerant enzymes (Zhang et al., 2016, Huang et al., 2015, Hu et al., 2007), radiation tolerance (Rao et al., 2016), plant growth promotion (Rajput et al., 2013) and antioxidant and cytotoxic activities (Zhang et al., 2020, Shindo et al., 2008). They have also been extensively studied for their potential applications in bioremediation and polysaccharide degradation (Al-Awadhi et al., 2012, Vishnupriya et al., 2023, Gaur et al., 2022, Li et al., 2006).

Strain APC 4016^T was previously isolated during a study of antimicrobial-producing isolates from the microbiome of deep-sea fish, and subsequently shown to encode secondary metabolites and bacteriocins (Uniacke-Lowe et al., 2023). In this study, we report the taxonomic characterization of a *Planococcus* bacterial strain, APC 4016^T, isolated from the skin of a deep-sea snub-nosed spiny eel (*Notacanthus chemnitzii*).

4.2. Materials and Methods

4.2.1. Isolate information

Strain APC 4016^T was previously isolated from the skin of a snub-nosed spiny eel (*Notacanthus chemnitzii*; Uniacke-Lowe et al., 2023). In brief, the fish was collected by research vessels from a depth of approximately 1000m in international waters near the Grand Banks of Newfoundland in the north-western Atlantic Ocean (43.282 N 49.121 W; Collins et al., 2021). Swabs were taken of the skin and plated onto BD Difco marine agar 2216 (MA) and incubated aerobically for 3 weeks at 4 °C. An orange-coloured colony was selected and purified through sub-culturing on MA. The strain, designated APC 4016, was deposited into the APC Culture Collection, cells of which were preserved in 35 % (v/v) glycerol suspensions at –80 °C (Uniacke-Lowe et al., 2023). This strain has also been deposited into the DSMZ-German Collection of Microorganisms and Cell Cultures GmbH (= DSM 115753^T) and the National Collection of Industrial, Food and Marine Bacteria (= NCIMB 15463^T).

4.2.2. Genome sequencing, assembly and annotation

Strain APC 4016^T was cultured in BD Difco marine broth 2216 (MB) for 72hrs at 20 °C. Bacterial genomic DNA (gDNA) was extracted and sequenced, as previously described (Uniacke-Lowe et al., 2023). Briefly, gDNA was extracted using the GeneJET Genomic DNA Purification Kit (Thermo Scientific) and sequenced by MicrobesNG (<https://microbesng.com/> ; University of Birmingham, UK). The draft genome was assembled using SPAdes (*de novo* assembly method), and assembly quality was assessed using QUAST. The assembled contigs were submitted to GenBank and annotated upon submission using NCBI's Prokaryotic Annotation Pipeline. For this study, the draft assembly was annotated using PROKKA (Seemann, 2014). Additionally, functional subsystems were predicted using RASTtk (version 2.0) and SEED (Overbeek et al., 2014, Aziz et al., 2008). AntiSMASH and BAGEL4 were used to screen for the presence of secondary metabolite and bacteriocin biosynthetic gene clusters (BGCs), respectively, as previously reported (Uniacke-Lowe et al., 2023).

4.2.3. 16S rRNA gene phylogeny

The complete 16S rRNA gene sequence (1548 bp) was extracted from the gDNA sequence data using Barrnap (version 0.9; <https://github.com/tseemann/barrnap>). The 16S sequence was then queried using the 16S-based ID application on the EzBioCloud server (Yoon et al., 2017). Sequence files (.fasta) of the top hits were downloaded and imported into MEGA X (Kumar et al., 2018). Multiple alignments with the query sequence were created using ClustalW (Edgar, 2004) within the MEGA X program. Phylogenetic trees (neighbour-joining (Saitou and Nei, 1987), maximum-likelihood (Felsenstein, 1981) and maximum parsimony (Nei and Kumar, 2000)) were inferred using the Kimura-2-parameter method (Kimura, 1980) and bootstrap analyses was based on 1000 replicates (Felsenstein, 1985). The tree was rooted by including sequences from the family *Caryophanaceae* (*Caryophanaceae/Planococcaceae*) as outliers; namely, the 16S rRNA gene sequences from *Sporosarcina saromensis* HG645 (AB243859) and *Psychrobacillus lasiicapitis* NEAU-3TGS17 (KP219721).

4.2.4. Genomic phylogeny

For the following analyses, all available *Planococcus/Planomicrobium* valid type strain reference genomes (n=22) were downloaded from the GenBank (<https://www.ncbi.nlm.nih.gov/data-hub/genome/>) and JGI (<https://img.jgi.doe.gov/>) databases. Assemblies were annotated using PROKKA to obtain the GFF3 files for subsequent analysis. Pairwise average nucleotide identity (ANI) values were calculated between APC 4016^T and the *Planococcus/Planomicrobium* type strain reference assemblies with Pyani (version 0.2.12; Pritchard et al., 2016) using the ANI MUMer (ANIm) method (Richter and Rosselló-Móra, 2009). ROARY (Page et al., 2015) was used to generate multiple sequence alignments of the core genes of APC 4016^T and the type strain reference genomes. The annotated genome of *Sporosarcina aquimarina* S/N-308-OC-B4 was included as an outlier (ROARY parameters: 95 % 'identity for BLAST', 95 % 'percentage of isolates a gene must be in to be core', 100000 cluster limit). RaxML (Stamatakis, 2014) was used to generate a maximum-likelihood phylogenetic tree from the core gene alignment, which was then visualised in MEGA X. Additionally, the DSMZ Type

(Strain) Genome Server (TYGS) annotation platform (Meier-Kolthoff and Göker, 2019) was used to calculate digital DNA–DNA hybridization (dDDH) values between APC 4016^T and related type strains from the database.

4.2.5. Biochemical and phenotypic characterisation

Colony morphology was assessed by incubating aerobically at 25 °C on MA and on tryptic soy agar (TSA; Oxoid/ThermoFisher Scientific) and observed after 3 and 5 days. Growth at different temperatures was tested at 4, 9, 21, 25, 30, 37, 40 and 44 °C on MA and in MB. For the latter, growth was assessed by measuring optical density at OD_{600nm}. The following tests were performed on strain APC 4016^T and, unless stated otherwise, were carried out aerobically at 30 °C. Anaerobic growth was assessed on MA incubated in an AnaeroGen anaerobic system (Oxoid) for up to 14 days. The NaCl range for growth was tested in MB with the NaCl concentration adjusted to within the range 0–22 % w/v (in increments of 1 % from 0–10 % NaCl, and in increments of 2 % from 12–22 % NaCl). The pH range for growth was tested in MB from pH 4 to 12 (in increments of 0.5 from pH 4 to 9 and increments of 1 from pH 9 to 11). The pH was adjusted by the addition of 1M HCl and/or 1M NaOH. Growth was assessed by measuring optical density at OD_{600nm} daily for 4 days. Cells were examined using phase contrast microscopy and scanning electron microscopy (SEM). SEM imaging was carried out by UCD Conway Imaging Core (UCD Conway Institute of Biomolecular and Biomedical Research, University College Dublin, Ireland). Gram staining was carried out according to standard methods. Production of oxidase was tested using oxidase strips (Millipore) according to manufacturer's instructions. Production of catalase was tested by transferring a mass of colonies onto a glass slide and exposing them to 1–2 drops of 3 % hydrogen peroxide and observing for effervescence. Observation for motility was carried out using the hanging drop method according to the method outlined by Bernardet et al. (2002) except that a cavity microscope slide was used. The presence of flexirubin-type pigments was tested using the KOH method as described by Bernardet et al. (2002). Production of hydrogen sulphide was determined using Watman lead acetate strips (Sigma-Aldrich) suspended above an inoculum of the strain in MB in a test tube and was monitored for up to 7 days.

Artificial seawater (ASW) salts were prepared for the following assays according to Kurilenko et al. (2019) as follows: 30 g l⁻¹ NaCl, 5.94 g l⁻¹ MgSO₄·7H₂O, 4.53 g l⁻¹ MgCl₂·6H₂O, 0.64 g l⁻¹ KCl and 1.3 g l⁻¹ CaCl₂·2H₂O. Assimilation of Tween 80 and starch was tested by plating on agar-medium (15 g l⁻¹ agar) containing ASW salts, 5 g l⁻¹ peptone, 1 g l⁻¹ yeast extract, 0.1 g l⁻¹ K₂HPO₄, and 1 % v/v Tween 80 or 0.2 % w/v starch. Hydrolysis of DNA was determined using DNase Test Agar (Thermo Scientific) supplemented with ASW salts. Hydrolysis of casein was tested on agar medium containing ASW salts supplemented with 10 % w/v skimmed milk powder. Hydrolysis of cellulose was determined by culturing APC 4016^T on agar medium containing ASW salts, 0.2 % w/v cellulose and 0.5 % w/v peptone; following incubation the plate was then stained with 1 % Congo Red, washed with 1 M NaCl, and observed for zones of clearance around colonies.

Enzyme activity was assessed using the API ZYM (bioMérieux) test kit which was incubated at 30 °C for 18h. Other biochemical tests were carried out using the API 20E, API 20NE kits (bioMérieux) and the GENIII MicroPlate (Biolog). Cells were prepared in 0.85 % NaCl for the API kits. For the GENIII MicroPlate, protocol B was used with inoculating fluid B supplemented with 2 % NaCl. The results were recorded daily for up to 4 days. Antibiotic susceptibility was determined using the disc-diffusion method (Hudzicki, 2009). Muller-Hinton agar (Oxoid) was used with commercial antibiotic discs (Oxoid) with the following antibiotics: ampicillin (10 µg), chloramphenicol (30 µg) erythromycin (15 µg), gentamicin (10 µg), kanamycin (30 µg), lincomycin (15 µg), neomycin (30 µg), novobiocin (5 µg), oleandomycin (15 µg), penicillin G (10 U), polymyxin B (300 U), rifampicin (30 µg), streptomycin (10 µg) and tetracycline (30 µg). Zones of inhibition were measured after 36hrs at 30 °C. However, due to the lack of CLSI breakpoint data for interpretation criteria for the genus *Planococcus*, antibiotic susceptibility and resistance were interpreted as no growth or growth, respectively.

Analyses of cellular fatty acids, polar lipids and respiratory quinones was carried out by DSMZ (Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures GmbH, Braunschweig, Germany). For fatty acid and polar lipid analyses, strain APC 4016^T was prepared by culturing on MA at 25 °C. Cellular fatty acids were

determined using the Sherlock MIS (MIDI) system (version 6.1; TSBA40 method, TSBA6 calculation).

4.3. Results

4.3.1. Genome features and characteristics

The draft genome assembly of strain APC 4016^T was 3640984 bp in length with a G+C content of 45.25 mol%. An overview of the general characteristics of the draft genome is given in Table 1. Using RAST, 1603 subsystem features were identified across 24 subsystem categories. An overview of the composition of RAST subsystem features is shown in Fig. 1. The largest subsystem categories were dedicated to ‘amino acids and derivatives’, at 297 features/genes (18.5 % of the total), followed by ‘carbohydrates’, at 254 features (15.8 %) and ‘protein metabolism’ at 187 features (11.7 %). Three secondary metabolite BGCs were identified using antiSMASH, encoding two terpenes, one of which shared 28 % similarity to the BGC of carotenoid, and a type-3 polyketide synthase (T3PKS) which shared 8 % similarity to beta-lactam. A BGC encoding two lanthipeptides (class II/LanM-associated) was detected using BAGEL4, the biosynthetic genes of which shared 50 and 55.6 % match with cerecidin (*Bacillus cereus*), respectively (data from Uniacke-Lowe et al. [2023]). An overview of the antiSMASH and BAGEL screening results is given in Table S1 .

4.3.2. 16S rRNA gene and genome phylogenetic analysis

The 16S rRNA gene sequence of strain APC 4016^T is 1548 bp in length. Analysis of the 16S rRNA gene sequence using the EzBioCloud search tool indicated that APC 4016^T was most closely related to members of the genus *Planococcus* (*Planococcus/Planomicrobium*) within the family *Caryophanaceae*. The 16S rRNA gene sequence of APC 4016^T shared the highest similarity with *Planomicrobium okeanokoites* IFO 12536^T (98.98 % similarity) followed by *Planococcus plakortidis* DSM 23997^T (98.64 %), *Planococcus salinarum* DSM 23820^T (= ISL-16^T, 98.64 %), *Planococcus donghaensis* DSM 22276^T (98.57 %) and *Planococcus halotolerans* SCU63^T (98.52 %). Completeness values were all 100 %. These strains were chosen as reference strains for the comparative phenotypic and chemotaxonomic analyses. The maximum-likelihood phylogenetic tree created from the 16S rRNA gene sequences (Fig. 2) shows the evolutionary relatedness between APC 4016^T and the EzBioCloud hits. There was slight variability in topology between the maximum-likelihood, neighbour-joining (Fig. S1) and maximum parsimony (Fig. S2) trees, likely due to low

bootstrap confidence. However, in all three trees, APC 4016^T either formed its own clade or clustered with *P. okeanokoites* IFO 12536^T.

A summary of the overall genome relatedness between APC 4016^T and the reference strains is given in Table 2. The ANI values between APC 4016^T and the reference *Planococcus/Planomicrobium* genomes ranged from 82.95 to 83.76 %. The ANI coverage ranged from 6.2 to 52.7 %. The corresponding graphical output of ANI from PYANI is shown in Fig. S3. The dDDH (d4) values ranged from 19.5 to 22.2 %. These ANI and dDDH values are well below the general thresholds for taxonomic species of 95-96 % and 70 %, respectively (Chun et al., 2018). These values are especially important to consider when the 16S rRNA gene sequence % identity is ≥ 98.7 % (Chun et al., 2018). The maximum-likelihood tree generated from the ROARY core gene (n=67) alignments is shown in Fig. S4. This tree shows strain APC 4016^T clusters closely to a clade formed by *Planomicrobium stackebrandtii* K22-03^T, *Planococcus antarcticus* DSM 14505^T, *Planococcus faecalis* AJ003^T, *Planococcus kocurii* ATCC 43850^T, *Planococcus versutus* L10 15^T, *P. donghaensis* DSM 22276^T and *Planococcus halocryophilus* DSM 24743^T. The position of APC 4016^T within this tree indicates that it is a separate species amongst these strains. These results show that although APC 4016^T shared a high 16S rRNA gene sequence % identity to the reference strains, their genomes were not highly similar.

4.3.3. Cell morphology, biochemical and phenotypic characterisation

Cells of strain APC 4016^T are aerobic, motile cocci, 0.6-0.9 μm in diameter, occur in clusters of two to eight (Fig. 3) and are Gram-stain positive to variable (Fig. S5). Motility was observed as tumbling and spinning motions. After culturing on MA for 3 days, colonies of strain APC 4016^T appeared smooth circular and light orange with darker coloured edges, and 1–1.2 mm in diameter. Colonies cultured on TSA were paler in colour and slightly crateriform (Fig. 4). Flagella were not observed in SEM imaging (Fig. 3). The colony morphology after 5 days of incubation is shown in Fig. 4. From the API ZYM assay, strain APC 4016^T was positive for alkaline phosphatase, esterase, esterase lipase, leucine arylamidase, α -chymotrypsin, naphthol-AS-BI-phosphohydrolase and α -glucosidase activity; weakly positive for valine arylamidase, cystine arylamidase, acid phosphatase and β -galactosidase, and negative for lipase,

trypsin, α -galactosidase, β -glucuronidase, β -glucosidase, mannosidase and fucosidase activity. Strain APC 4016^T was susceptible to all of the tested antibiotics: ampicillin (10 μ g), chloramphenicol (30 μ g), erythromycin (15 μ g), gentamicin (10 μ g), kanamycin (30 μ g), lincomycin (15 μ g), neomycin (30 μ g), novobiocin (5 μ g), oleandomycin (15 μ g), penicillin G (10 U), polymyxin B (300 U), rifampicin (30 μ g), streptomycin (10 μ g) and tetracycline (30 μ g). The antibiotic susceptibility profile of strain 4016^T with corresponding zone of inhibition measurements is given in Table S2. Due to the lack of CLSI guidelines and breakpoint data for the genus *Planococcus*, susceptibility and resistance were interpreted as no growth or growth. We also note the lack of standardised antibiotic susceptibility testing methods for members of this genus in the literature, such as the use of alternative media instead of Mueller-Hinton, and varying growth temperatures. This represents a deficiency in the standardisation within this group, however, we have reported our results in line with existing literature.

Differential characteristics between APC 4016^T and the reference *Planococcus/Planomicrobium* strains are provided in Table 3. All the data for strain APC 4016^T is from this study, the data for the reference strains was acquired from the respective reference papers referred to in the Table 3 description. In the phenotypic assays, APC 4016^T shared a number of characteristics that are in accordance with members of the genus *Planococcus*, such as positive for catalase, negative for urease and hydrolysis of starch and a predominance of menaquinones 7 and 8 as the cellular respiratory quinones (Nakagawa et al., 1996). However, strain APC 4016^T differed in that it was positive for nitrate reduction and hydrolysis of Tween 80, weakly positive for hydrolysis of aesculin and contained menaquinones MK5 and MK6, as well as MK7 and MK8. Strain APC 4016^T demonstrated a relatively wide temperature range for growth of 4–40 °C, similar to the reference strains (Table 3). Strain APC 4016^T also demonstrated relatively wide %NaCl and pH ranges for growth of 0–12 % (optimum, 1 %) and 5.5–12 (inclusive; optimum, pH 7–7.5), respectively. Growth above pH 12 could not be determined due to instability of the growth media at higher pH. In contrast to the majority of the closely related reference strains, APC 4016^T tested negative for oxidase and hydrolysis of casein but positive for β -galactosidase.

The major (>10 % of total) cellular fatty acids of strain APC 4016^T were iso-C_{14:0} (12.5 %), anteiso-C_{15:0} (38.9 %), and C_{16:1} ω-Alc. (19.2 %). The overall cellular fatty acid composition profiles of strain APC 4016^T and the reference *Planococcus/Planomicrobium* strains were similar (Table 4). Strain APC 4016^T differed slightly in that it had very few straight-chain fatty acids and did not contain iso-C_{16:0} as a major component (>10 %) but contained a significant proportion of a unique summed feature, iso-C_{17:1} I/anteiso B (which could not be distinguished by the MIDI system). The cellular polar lipids of strain APC 4016^T were phosphatidylglycerol, diphosphatidylglycerol, phosphatidyl-ethanolamine and three unknown lipids (Fig. S6).

Based on the combination of phylogenetic distance, phenotypic and biochemical characteristics, strain APC 4016^T represents a novel species for which we propose the name *Planococcus notacanthi* sp. nov.

4.4. Description of *Planococcus notacanthi* sp. nov.

Planococcus notacanthi (not.a.can'thi. N.L. gen. n., *notacanthi*, pertaining to the fish it was isolated from, *Notacanthus chemnitzii*)

Cells are aerobic, motile cocci, 0.6 – 0.9 µm in diameter (arranged in clusters of 2 to 8), Gram-stain positive to variable. Colonies are light orange, circular, smooth on MA and slightly crateriform on TSA, raised, 1–1.2 mm in diameter. Catalase positive and oxidase negative. Flexirubin-type pigments are not detected. The temperature range for growth of is 4–40 °C (optimum, 25–28 °C). The NaCl range for growth is 0–12 % (optimum, 1 %). The pH range for growth is 5.5–12 (optimum, pH 7–7.5). Positive for hydrolysis of DNA and Tween 80. Negative for hydrolysis of agar, casein, cellulose and starch. Negative for production of H₂S. From API 20E: positive for β-galactosidase and weakly positive for tryptophan deaminase activity. Negative for arginine dihydrolase, lysine decarboxylase, ornithine decarboxylase, urease, citrate utilization, indole, acetoin, gelatinase and H₂S production. Negative for fermentation of/acid production from glucose, mannitol, inositol, sorbitol, rhamnose, sucrose, melibiose, amygdalin and arabinose. From API 20NE: Positive for nitrate reduction, production of β-galactosidase and weakly positive for hydrolysis of aesculin; negative for production of indole, urease, arginine dihydrolase, hydrolysis of gelatin and assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, N-acetyl-glucosamine, maltose, potassium gluconate, capric acid, adipic acid, malic acid, trisodium citrate and phenylacetic acid. From Biolog GENIII: Positive for oxidation of dextrin, L-fucose, D-glucose-6-PO₄, D-fructose-6-PO₄, D-glucuronic acid, glucuronamide, quinic acid, L-lactic acid, α-keto-glutaric acid, acetoacetic acid and acetic acid, and weakly positive for D-fucose, D-galacturonic acid and L-galacturonic acid. Major cellular fatty acids are iso-C_{14:0}, anteiso-C_{15:0}, and C_{16:1} ω-Alc. Respiratory quinones are MK5, MK6, MK7, and MK8. Cellular polar lipids are phosphatidylglycerol, diphosphatidylglycerol, phosphatidylethanolamine and three unknown lipids.

The type strain is APC 4016^T (= DSM 115753^T = NCIMB 15463^T), isolated from the skin of a deep-sea snub-nosed spiny eel, *Notacanthus chemnitzii*, from a depth of 1000 m in the Northwest Atlantic Ocean (43.282 N 49.121 W). The genome of the type strain is 3.6 Mb with a DNA G+C content of 45.25 mol%.

The GenBank accession numbers for the 16S rRNA gene sequence and the draft genome of strain APC 4016^T are OQ456996 and JASDCQ000000000, respectively.

4.5. References

- AL-AWADHI, H., DASHTI, N., KANSOUR, M., SORKHOH, N. & RADWAN, S. 2012. Hydrocarbon-utilizing bacteria associated with biofouling materials from offshore waters of the Arabian Gulf. *International Biodeterioration & Biodegradation*, 69, 10-16.
- ALAM, S. I., SINGH, L., DUBE, S., REDDY, G. S. & SHIVAJI, S. 2003. Psychrophilic *Planococcus maitriensis* sp. nov. from Antarctica. *Systematic and Applied Microbiology*, 26, 505-10.
- AZIZ, R. K., BARTELS, D., BEST, A. A., DEJONGH, M., DISZ, T., EDWARDS, R. A., FORMSMA, K., GERDES, S., GLASS, E. M., KUBAL, M., MEYER, F., OLSEN, G. J., OLSON, R., OSTERMAN, A. L., OVERBEEK, R. A., MCNEIL, L. K., PAARMANN, D., PACZIAN, T., PARRELLO, B., PUSCH, G. D., REICH, C., STEVENS, R., VASSIEVA, O., VONSTEIN, V., WILKE, A. & ZAGNITKO, O. 2008. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*, 9, 75.
- BERNARDET, J.-F., NAKAGAWA, Y., HOLMES, B. & PROKARYOTES, S. O. T. T. O. F. A. C.-L. B. O. T. I. C. O. S. O. 2002. Proposed minimal standards for describing new taxa of the family *Flavobacteriaceae* and emended description of the family. *International Journal of Systematic and Evolutionary Microbiology*, 52, 1049-1070.
- CHOI, J.-H., IM, W.-T., LIU, Q.-M., YOO, J.-S., SHIN, J.-H., RHEE, S.-K. & ROH, D.-H. 2007. *Planococcus donghaensis* sp. nov., a starch-degrading bacterium isolated from the East Sea, South Korea. *International Journal of Systematic and Evolutionary Microbiology*, 57, 2645-2650.
- CHUN, J., OREN, A., VENTOSA, A., CHRISTENSEN, H., ARAHAL, D. R., DA COSTA, M. S., ROONEY, A. P., YI, H., XU, X.-W., DE MEYER, S. & TRUJILLO, M. E. 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*, 68, 461-466.
- COLLINS, F. W. J., WALSH, C. J., GOMEZ-SALA, B., GUIJARRO-GARCÍA, E., STOKES, D., JAKOBSDÓTTIR, K. B., KRISTJÁNSSON, K., BURNS, F., COTTER, P. D., REA, M. C., HILL, C. & ROSS, R. P. 2021. The microbiome of deep-sea fish reveals new microbial species and a sparsity of antibiotic resistance genes. *Gut Microbes*, 13, 1-13.
- DAI, X., WANG, Y.-N., WANG, B.-J., LIU, S.-J. & ZHOU, Y.-G. 2005. *Planomicrobium chinense* sp. nov., isolated from coastal sediment, and transfer of *Planococcus psychrophilus* and *Planococcus alkanoclasticus* to *Planomicrobium* as *Planomicrobium psychrophilum* comb. nov. and *Planomicrobium alkanoclasticum* comb. nov. *International journal of systematic and evolutionary microbiology*, 55, 699-702.

- EDGAR, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32, 1792-1797.
- FELSENSTEIN, J. 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. *Journal of Molecular Evolution*, 17, 368-376.
- FELSENSTEIN, J. 1985. Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, 39, 783-791.
- GAN, L., ZHANG, Y., ZHANG, L., LI, X., WANG, Z., HE, L., LI, Z. & TIAN, Y. 2018. *Planococcus halotolerans* sp. nov., isolated from a saline soil sample in China. *International Journal of Systematic and Evolutionary Microbiology*, 68, 3500-3505.
- GAUR, V. K., GUPTA, P., TRIPATHI, V., THAKUR, R. S., REGAR, R. K., PATEL, D. K. & MANICKAM, N. 2022. Valorization of agro-industrial waste for rhamnolipid production, its role in crude oil solubilization and resensitizing bacterial pathogens. *Environmental Technology & Innovation*, 25, 102108.
- GUPTA, R. S. & PATEL, S. 2020. Robust demarcation of the family *Caryophanaceae* (*Planococcaceae*) and its different genera including three novel genera based on phylogenomics and highly specific molecular signatures. *Frontiers in microbiology*, 10, 2821.
- HU, J. M., LI, H., CAO, L. X., WU, P. C., ZHANG, C. T., SANG, S. L., ZHANG, X. Y., CHEN, M. J., LU, J. Q. & LIU, Y. H. 2007. Molecular cloning and characterization of the gene encoding cold-active β -galactosidase from a psychrotrophic and halotolerant *Planococcus* sp. L4. *Journal of Agricultural and Food Chemistry*, 55, 2217-2224.
- HUANG, X., LIN, J., YE, X. & WANG, G. 2015. Molecular characterization of a thermophilic and salt-and alkaline-tolerant xylanase from *Planococcus* sp. SL4, a strain isolated from the sediment of a soda lake. *Journal of Microbiology and Biotechnology*, 25, 662-671.
- HUDZICKI, J. 2009. Kirby-Bauer disk diffusion susceptibility test protocol. *American society for microbiology*, 15, 55-63.
- JUNG, Y.-T., KANG, S.-J., OH, T.-K., YOON, J.-H. & KIM, B.-H. 2009. *Planomicrobium flavidum* sp. nov., isolated from a marine solar saltern, and transfer of *Planococcus stackebrandtii* Mayilraj et al. 2005 to the genus *Planomicrobium* as *Planomicrobium stackebrandtii* comb. nov. *International journal of systematic and evolutionary microbiology*, 59, 2929-2933.
- KAUR, I., DAS, A. P., ACHARYA, M., KLENK, H.-P., SREE, A. & MAYILRAJ, S. 2012. *Planococcus plakortidis* sp. nov., isolated from the marine sponge *Plakortis simplex* (Schulze). *International Journal of Systematic and Evolutionary Microbiology*, 62, 883-889.
- KIMURA, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16, 111-120.

- KOMAGATA, M. V. H. A. K. 1985. A new species of *Planococcus*, *P. Kocurii* isolated from fish, frozen foods, and fish curing brine. *The Journal of General and Applied Microbiology*, 31, 441-455.
- KUMAR, S., STECHER, G., LI, M., KNYAZ, C. & TAMURA, K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, 35, 1547-1549.
- KURILENKO, V. V., ROMANENKO, L. A., ISAEVA, M. P., SVETASHEV, V. I. & MIKHAILOV, V. V. 2019. *Winogradskyella algae* sp. nov., a marine bacterium isolated from the brown alga. *Antonie Van Leeuwenhoek*, 112, 731-739.
- LI, H., LIU, Y. H., LUO, N., ZHANG, X. Y., LUAN, T. G., HU, J. M., WANG, Z. Y., WU, P. C., CHEN, M. J. & LU, J. Q. 2006. Biodegradation of benzene and its derivatives by a psychrotolerant and moderately haloalkaliphilic *Planococcus* sp. strain ZD22. *Research in Microbiology*, 157, 629-636.
- LUO, X., ZHANG, J., LI, D., XIN, Y., XIN, D. & FAN, L. 2014. *Planomicrobium soli* sp. nov., isolated from soil. *International Journal of Systematic and Evolutionary Microbiology*, 64, 2700-2705.
- MAYILRAJ, S., PRASAD, G. S., SURESH, K., SAINI, H. S., SHIVAJI, S. & CHAKRABARTI, T. 2005. *Planococcus stackebrandtii* sp. nov., isolated from a cold desert of the Himalayas, India. *International Journal of Systematic and Evolutionary Microbiology*, 55, 91-94.
- MEIER-KOLTHOFF, J. P. & GÖKER, M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nature Communications*, 10, 2182.
- MIGULA 1894. Ueber ein neues System der Bakterien. Arb. bakt. Inst. Karlsruhe, 1, 235-238.
- MYKYTCZUK, N. C. S., WILHELM, R. C. & WHYTE, L. G. 2012. *Planococcus halocryophilus* sp. nov., an extreme sub-zero species from high Arctic permafrost. *International Journal of Systematic and Evolutionary Microbiology*, 62, 1937-1944.
- NAKAGAWA, Y., SAKANE, T. & YOKOTA, A. 1996. Emendation of the genus *Planococcus* and transfer of *Flavobacterium okeanokoites* Zobell and Upham 1944 to the genus *Planococcus* as *Planococcus okeanokoites* comb. nov. *International Journal of Systematic and Evolutionary Microbiology*, 46, 866-870.
- NEI, M. & KUMAR, S. 2000. *Molecular evolution and phylogenetics*, Oxford university press.
- OVERBEEK, R., OLSON, R., PUSCH, G. D., OLSEN, G. J., DAVIS, J. J., DISZ, T., EDWARDS, R. A., GERDES, S., PARRELLO, B., SHUKLA, M., VONSTEIN, V., WATTAM, A. R., XIA, F. & STEVENS, R. 2014. The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). *Nucleic Acids Research*, 42, D206-14.

- PAGE, A. J., CUMMINS, C. A., HUNT, M., WONG, V. K., REUTER, S., HOLDEN, M. T. G., FOOKES, M., FALUSH, D., KEANE, J. A. & PARKHILL, J. 2015. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics*, 31, 3691-3693.
- PRITCHARD, L., GLOVER, R. H., HUMPHRIS, S., ELPHINSTONE, J. G. & TOTH, I. K. 2016. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. *Analytical Methods*, 8, 12-24.
- RAJPUT, L., IMRAN, A., MUBEEN, F. & HAFEEZ, F. Y. 2013. Salt-tolerant PGPR strain *Planococcus rifietoensis* promotes the growth and yield of wheat (*Triticum aestivum* L.) cultivated in saline soil. *Pakistan Journal of Botany*, 45, 1955-1962.
- RAO, S., CHAN, O. W., LACAP-BUGLER, D. C. & POINTING, S. B. 2016. Radiation-tolerant bacteria isolated from high altitude soil in Tibet. *Indian Journal of Microbiology*, 56, 508-512.
- REDDY, G. S., PRAKASH, J. S., VAIRAMANI, M., PRABHAKAR, S., MATSUMOTO, G. I. & SHIVAJI, S. 2002. *Planococcus antarcticus* and *Planococcus psychrophilus* spp. nov. isolated from cyanobacterial mat samples collected from ponds in Antarctica. *Extremophiles*, 6, 253-61.
- RICHTER, M. & ROSSELLÓ-MÓRA, R. 2009. Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences*, 106, 19126-19131.
- ROSENBERG, E., DELONG, E. F., LORY, S., STACKEBRANDT, E. & THOMPSON, F. 2014. The prokaryotes: Firmicutes and tenericutes. *Prokaryotes*, 1, 3.
- VISHNUPRIYA, S., JABIR, T., ADARSH, B. M., KATTATHEYIL, H., SHAHANA KABEER, S., KRISHNAN, K. P., RADHAKRISHNAN, C. K. & MOHAMED HATHA, A. A. 2023. Diversity of complex polysaccharide degrading bacteria from the sediments of interlinked high Arctic fjords, Svalbard. *Regional Studies in Marine Science*, 102989.
- SAITOU, N. & NEI, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4, 406-425.
- SEE-TOO, W.-S., EE, R., MADHAIYAN, M., KWON, S.-W., TAN, J. Y., LIM, Y. L., CONVEY, P., PEARCE, D. A., YIN, W. F. & CHAN, K.-G. 2017. *Planococcus versutus* sp. nov., isolated from soil. *International Journal of Systematic and Evolutionary Microbiology*, 67, 944-950.
- SEEMANN, T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 30, 2068-9.
- SHINDO, K., ENDO, M., MIYAKE, Y., WAKASUGI, K., MORRITT, D., BRAMLEY, P. M., FRASER, P. D., KASAI, H. & MISAWA, N. 2008. Methyl Glucosyl-3,4-dehydro-apo-8'-lycopenoate, a Novel Antioxidative Glyco-C30-carotenoic Acid Produced by a Marine Bacterium *Planococcus maritimus*. *The Journal of Antibiotics*, 61, 729-735.

- STAMATAKIS, A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30, 1312-1313.
- UNIACKE-LOWE, S., COLLINS, F. W. J., HILL, C. & ROSS, R. P. 2023. Bioactivity Screening and Genomic Analysis Reveals Deep-Sea Fish Microbiome Isolates as Sources of Novel Antimicrobials. *Marine Drugs*, 21, 444.
- WANG, X., WANG, Z., ZHAO, X., HUANG, X., ZHOU, Y. & LI, W.-J. 2017. *Planococcus ruber* sp. nov., isolated from a polluted farmland soil sample. *International Journal of Systematic and Evolutionary Microbiology*, 67, 2549-2554.
- YANG, R., ZHANG, B., WANG, J., TAI, X., SUN, H., ZHANG, G., ZHANG, W., CHEN, T. & LIU, G. 2020. *Planococcus lenghuensis* sp. nov., an oil-degrading bacterium isolated from petroleum-contaminated soil. *Antonie van Leeuwenhoek*, 113, 839-850.
- YOON, J.-H., KANG, S.-J., LEE, S.-Y., OH, K.-H. & OH, T.-K. 2010. *Planococcus salinarum* sp. nov., isolated from a marine solar saltern, and emended description of the genus *Planococcus*. *International journal of systematic and evolutionary microbiology*, 60, 754-758.
- YOON, J. H., KANG, S. S., LEE, K. C., LEE, E. S., KHO, Y. H., KANG, K. H. & PARK, Y. H. 2001. *Planomicrobium koreense* gen. nov., sp. nov., a bacterium isolated from the Korean traditional fermented seafood jeotgal, and transfer of *Planococcus okeanoikoites* (Nakagawa et al. 1996) and *Planococcus mcmeekinii* (Junge et al. 1998) to the genus *Planomicrobium*. *International Journal of Systematic and Evolutionary Microbiology*, 51, 1511-1520.
- YOON, S.-H., HA, S.-M., KWON, S., LIM, J., KIM, Y., SEO, H. & CHUN, J. 2017. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *International Journal of Systematic and Evolutionary Microbiology*, 67, 1613-1617.
- ZHANG, B., YANG, R., ZHANG, G., LIU, Y., ZHANG, D., ZHANG, W., CHEN, T. & LIU, G. 2020. Characteristics of *Planococcus antioxidans* sp. nov., an antioxidant-producing strain isolated from the desert soil in the Qinghai-Tibetan Plateau. *MicrobiologyOpen*, 9, 1183-1196.
- ZHANG, D. C., LIU, H. C., XIN, Y. H., YU, Y., ZHOU, P. J. & ZHOU, Y. G. 2009. *Planomicrobium glaciei* sp. nov., a psychrotolerant bacterium isolated from a glacier. *International Journal of Systematic and Evolutionary Microbiology*, 59, 1387-90.
- ZHANG, H., MU, H., MO, Q., SUN, T., LIU, Y., XU, M., WANG, H., DAI, Y. & LU, F. 2016. Gene cloning, expression and characterization of a novel cold-adapted protease from *Planococcus* sp. *Journal of Molecular Catalysis B: Enzymatic*, 130, 1-8.

4.6. Chapter 4 Figures and Tables

Table 1. General characteristics of the draft genome of strain APC 4016^T.

Genomic feature	
Draft genome size (bp)	3640984
G+C content (mol%)	45.25
N50 (kb)	727.4
L50 (kb)	2
Number of contigs	21
Protein-coding genes (CDS)	3614
rRNA genes	5
tRNA genes	40
tmRNA genes	1

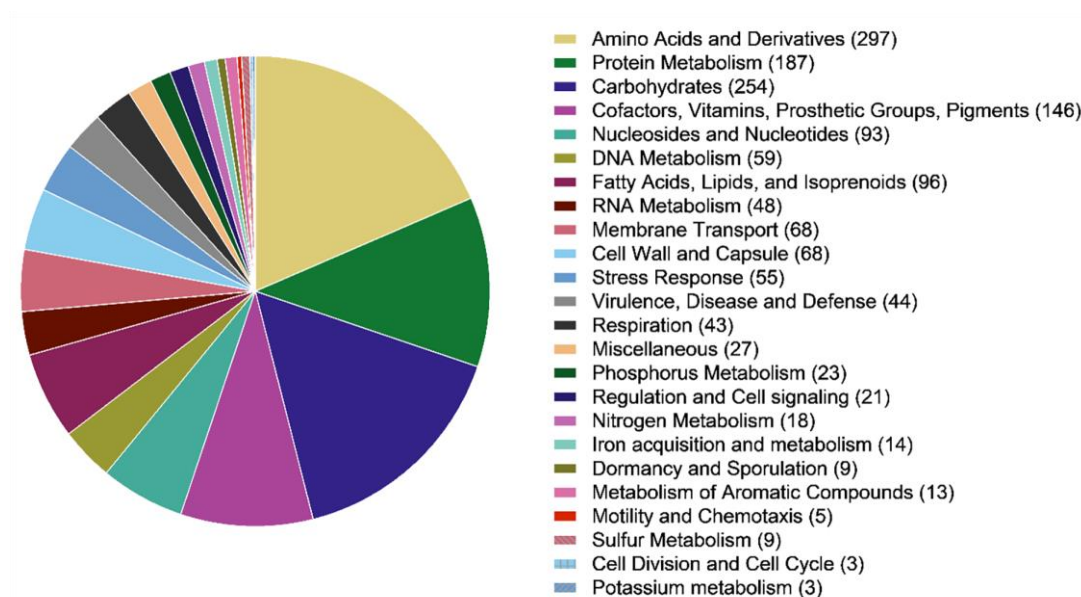


Fig. 1. Overview of the subsystem features of the draft genome of strain APC 4016^T based on the RAST annotation and SEED database. The counts of features per subsystem category are given. The total number of subsystem features was 1603.

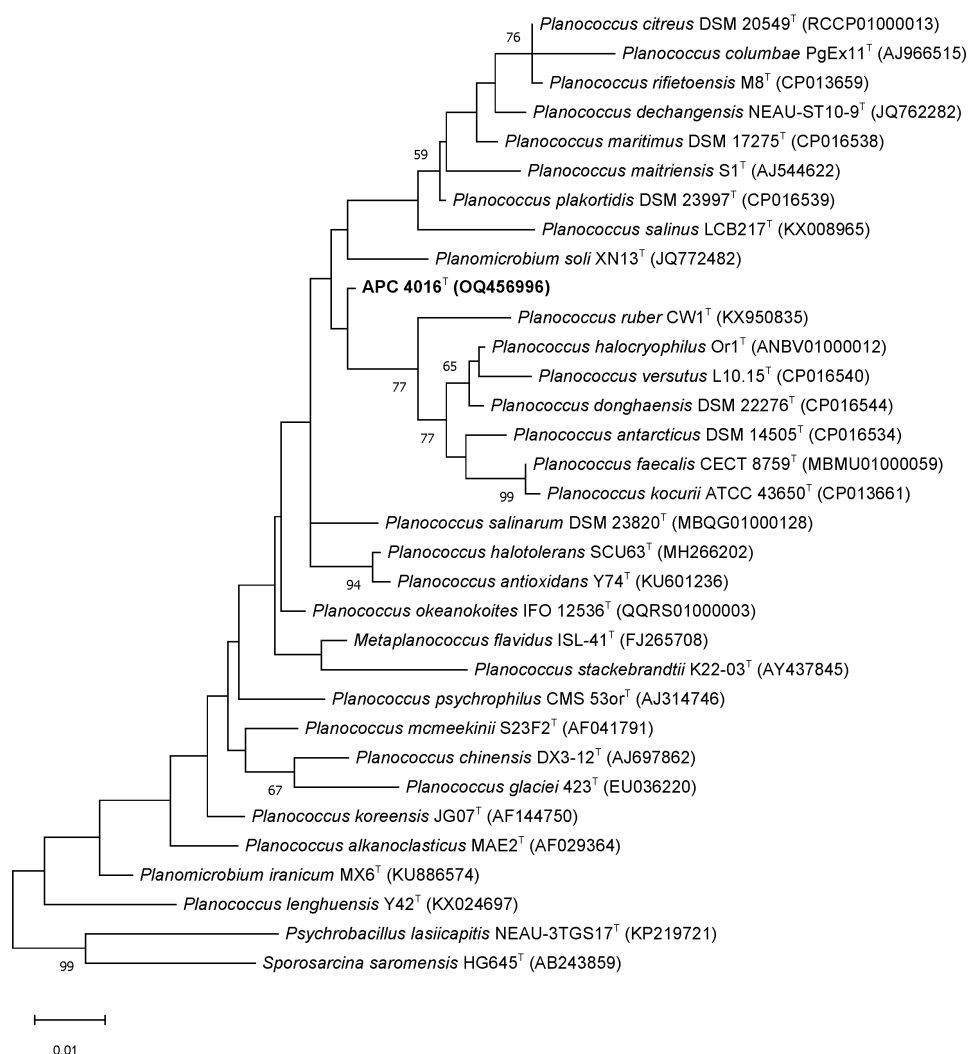


Fig. 2: Maximum-likelihood evolutionary tree of distances based on 16S rRNA gene sequences, showing the relationship of strain APC 4016^T and related species. Bootstrap values based on 1000 replicates (>50 %) are shown. The tree was rooted by 16S rRNA gene sequences from *Psychrobacillus lasiacapitis* NEAU-3TGS17 and *Sporosarcina saromensis* HG645 of the family *Caryophanaceae*.

Table 2: Summary of genomic comparisons between APC 4016^T and closely related type strains. dDDH = digital DNA-DNA hybridisation; ANI = average nucleotide identity; coverage = ANI alignment coverage.

Query strain	Subject strain	16S ID (%)	dDDH (d ₄)	ΔG+C (mol%)	ANI (%)	Coverage (%)
APC 4016 ^T	<i>Planococcus okeanokoites</i> IFO 12536 ^T	98.98	20.1	5.9	83.8	8.6
	<i>Planococcus plakortidis</i> DSM 23997 ^T	98.64	19.5	4.7	82.9	6.9
	<i>Planococcus donghaensis</i> DSM 22276 ^T	98.64	22.2	5.2	83.2	27.7
	<i>Planococcus salinarum</i> DSM 23820 ^T	98.64	21.9	2.7	83.8	7.7
	<i>Planococcus halotolerans</i> SCU63 ^T	98.57	19.9	0.7	83.7	8.8

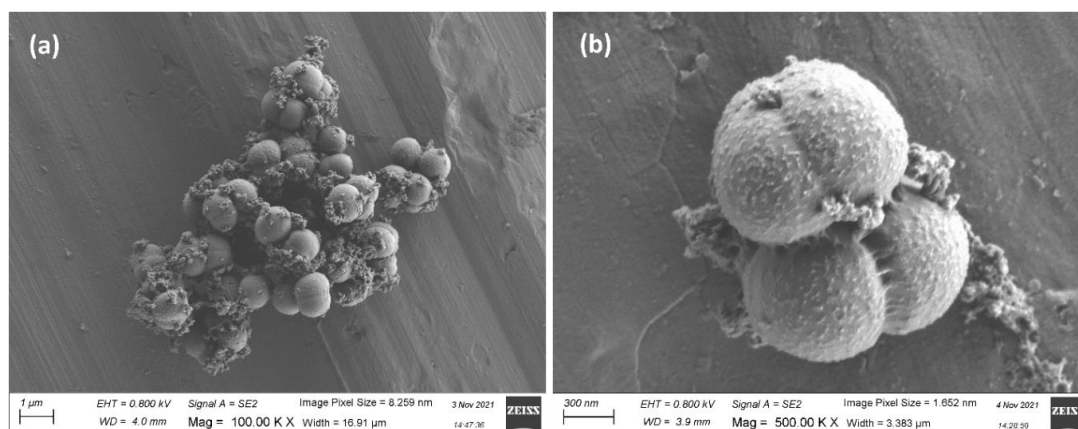


Fig. 3. Scanning electron microscopy image of cells of APC 4016^T at x100K (a) and x500K (b).

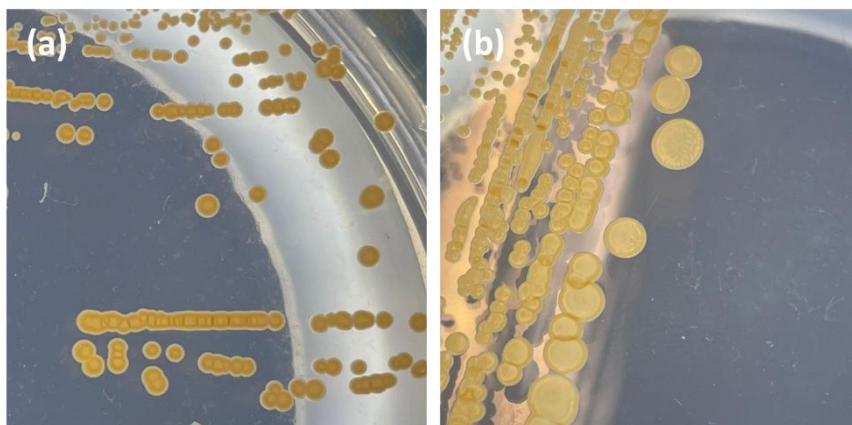


Fig. 4. Colonies of APC 4016^T after culturing at 25 °C for 5 days on marine agar (a) and tryptic soy agar (b).

Table 3: Differential characteristics of APC 4016^T and the reference *Planococcus/Planomicrobium* strains

Strains: 1, APC 4016^T (data from this study); 2, *P. okeanokoites* IFO 12536^T (data from Nakagawa et al., [1996], Yoon et al., [2001], Zhang et al., [2020]) ; 3, *P. plakortidis* DSM 2399^T (data from Kaur et al., [2012]); 4, *P. salinarum* DSM 23820^T (data from See-Too et al., [2017], Yoon et al., [2010]); 5, *P. donghaensis* DSM 22276^T (data from Choi et al., [2007]); 6, *P. halotolerans* SCU63^T (data from Gan et al., [2018]). +, Positive; –, negative; W, weakly positive; ND, not determined; V, variable; *, only minimum value for growth reported; †, only optimum range reported.

Characteristic	1	2	3	4	5	6
Cell shape	coccoid	rods	coccoid	coccoid, short rods or rods	coccoid	coccoid, short rod
Gram stain	+ to V	+ to V	+	+	+	+
Colony colour	light orange	bright yellow/orange	dull yellow/orange	pale yellow	orange	moderate orange
Motility	+	+	+	–	+	+
Range for growth						
Temperature (°C)	4–40	20–37 †	15–37	4–38	4–37	0–40
NaCl (% w/v)	0–12	0–7	2–7	0–13	0–12	15
pH	5.5–12	ND	6–10	>6 *	>7 *	6.5–9.0
Oxidase	–	W	+	+	+	+
Nitrate reduction	+	–	–	–	ND	–
Hydrolysis of:						
Tween 80	+	–	ND	+	–	–
Casein	–	+	–	+	+	+
Gelatin	–	+	–	–	–	–
Starch	–	–	–	–	+	–
Aesculin	W	–	–	–	+	ND
Utilisation of:						
D-glucose	–	–	–	–	+	+
D-maltose	–	–	–	–	ND	+
D-mannitol	–	–	–	–	+	+
D-mannose	–	–	–	–	–	+
D-raffinose	–	–	+	–	ND	ND
Sucrose	–	–	–	–	+	+
Acid production from:						
Glucose	–	–	+	–	–	ND
Enzyme Activity:						
β-galactosidase	+	ND	–	–	+	–
Predominant menaquinones	MK-5, MK-6, MK-7, MK-8	MK-7, MK-8	MK-6, MK-7, MK-8	MK-7, MK-8	MK-7, MK-8	MK-7, MK-8

Table 4: Cellular fatty acid profiles of APC 4016^T and the reference *Planococcus/Planomicrobium* type strains: 1, APC 4016^T (data from this study); 2, *P. okeanokoites* IFO 12536^T (data from Luo et al., [2014], Yoon et al., [2001]) ; 3, *P. plakortidis* DSM 2399^T (data from Kaur et al., [2012]); 4, *P. salinarum* DSM 23820^T (data from Kaur et al., [2012]); 5, *P. donghaensis* DSM 22276^T (data from Choi et al., [2007]); 6, *P. halotolerans* SCU63^T (data from Gan et al., [2018]). –, Not detected; TR, trace amount (<1 %).

Fatty Acid	1*	2†	3*	4*	5*	6‡
Straight-chain:						
C _{12:0}	–	–	–	9.6	–	–
C _{14:0}	–	–	1.5	5.5	–	–
C _{15:0}	–	–	3.3	–	1.7	–
C _{16:0}	TR	1.1-2.2	1.9	4.0	6.3	6.3
C _{17:0}	1.1	–	4.0	–	1.3	3.7
C _{18:0}	–	–	TR	7.4	–	1.8
Branched:						
iso-C _{14:0}	12.5	39.1-40.4	9.6	5.0	3.4	10.0
iso-C _{15:0}	2.9	0-1.1	13.2	–	4.7	5.4
anteiso-C _{15:0}	38.9	3.2-4.0	34.2	35.1	43.8	39.7
iso-C _{16:0}	6.3	24.5-29.0	11.7	11.8	9.4	10.6
iso-C _{17:0}	1.4	–	2.9	–	6.4	–
anteiso-C _{17:0}	2.3	–	5.5	7.7	15.5	5.6
iso-C _{17:1} -ω10c	2.7	–	–	–	TR	–
iso-C _{18:0}	1.1	TR	1.9	–	TR	TR
Unsaturated:						
C _{16:1} -ω7c alcohol	19.2	17.3-28.1	7.0	10.8	1.6	12.9
C _{16:1} ω11c	2.4	2.1-5.4	TR	–	1.3	–
C _{18:1} ω9c	–	0-1.0	–	–	–	2.7
Summed feature 4:						
C _{17:1} I/anteiso-B	8.7	–	–	–	–	–
Summed feature 5:						
iso-C _{17:1} I and/or anteiso-C _{17:1} B	–	–	1.2	–	1.9	–

*Analysis carried out at 25 °C.

†Analysis carried out at 28 °C.

‡Analysis carried out at 29 °C.

4.7. Chapter 4 Supplementary Material

Table S1. Overview of the secondary metabolite BGCs and bacteriocin gene clusters encoded by strain APC 4016^T identified using antiSMASH and BAGEL, respectively.

Sample	hit type	node	most similar known hit / database hit	similarity / match (%)	accession / motif	database
APC 4016	lanthipeptide class II	2	cerecidin	70	BGC0000502	antiSMASH
	terpene	2	carotenoid	100	BGC0000645	antiSMASH
	terpene	3	none	-	-	antiSMASH
	lanthipeptide class II x2	1	cerecidin	50,55.6	PF16934	BAGEL4

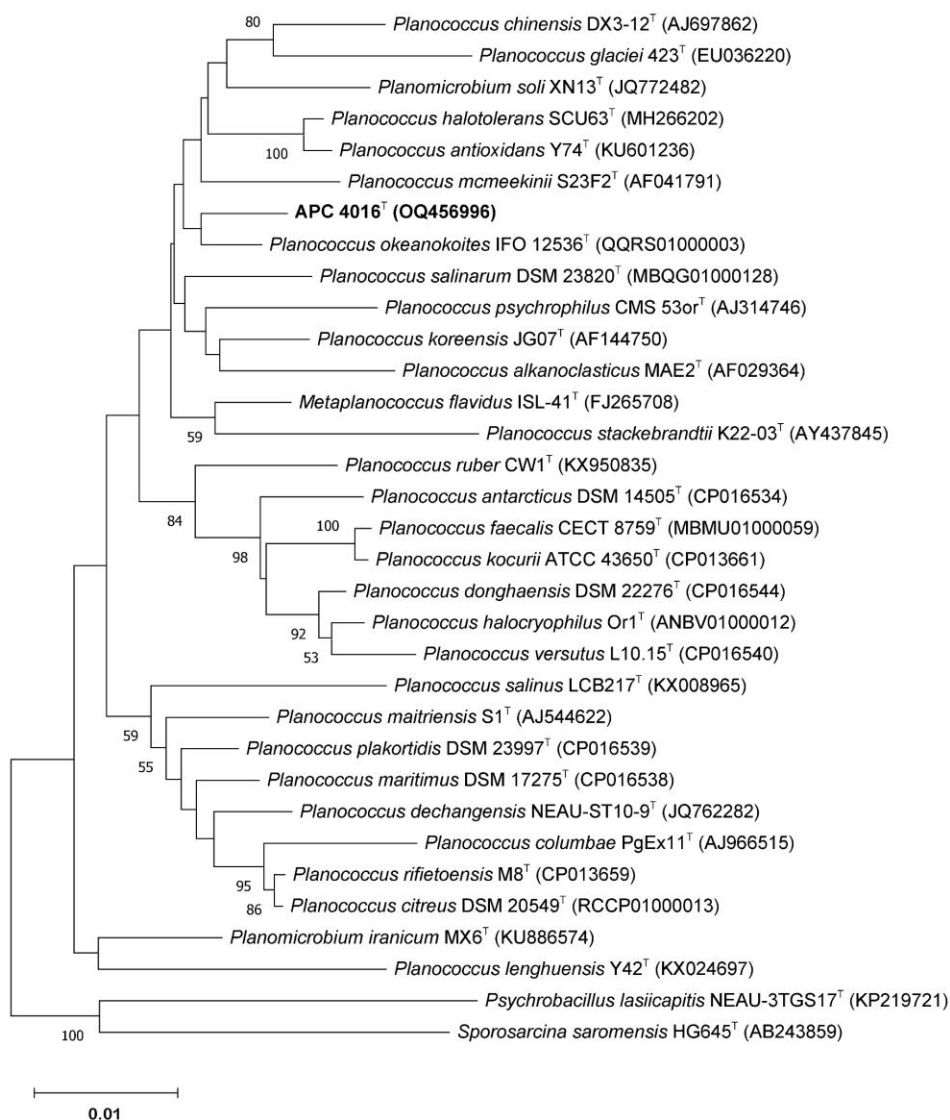


Fig. S1. Neighbor-joining tree of evolutionary tree of distances based on 16S rRNA gene sequences of APC 4016^T and related type strains. Bootstrap values based on 1000 replicates (>50 %) are shown. The tree was rooted by 16S rRNA gene sequences from *Psychrobacillus lasiicapitis* NEAU-3TGS17 and *Sporosarcina saromensis* HG645 of the family *Caryophanaceae*.

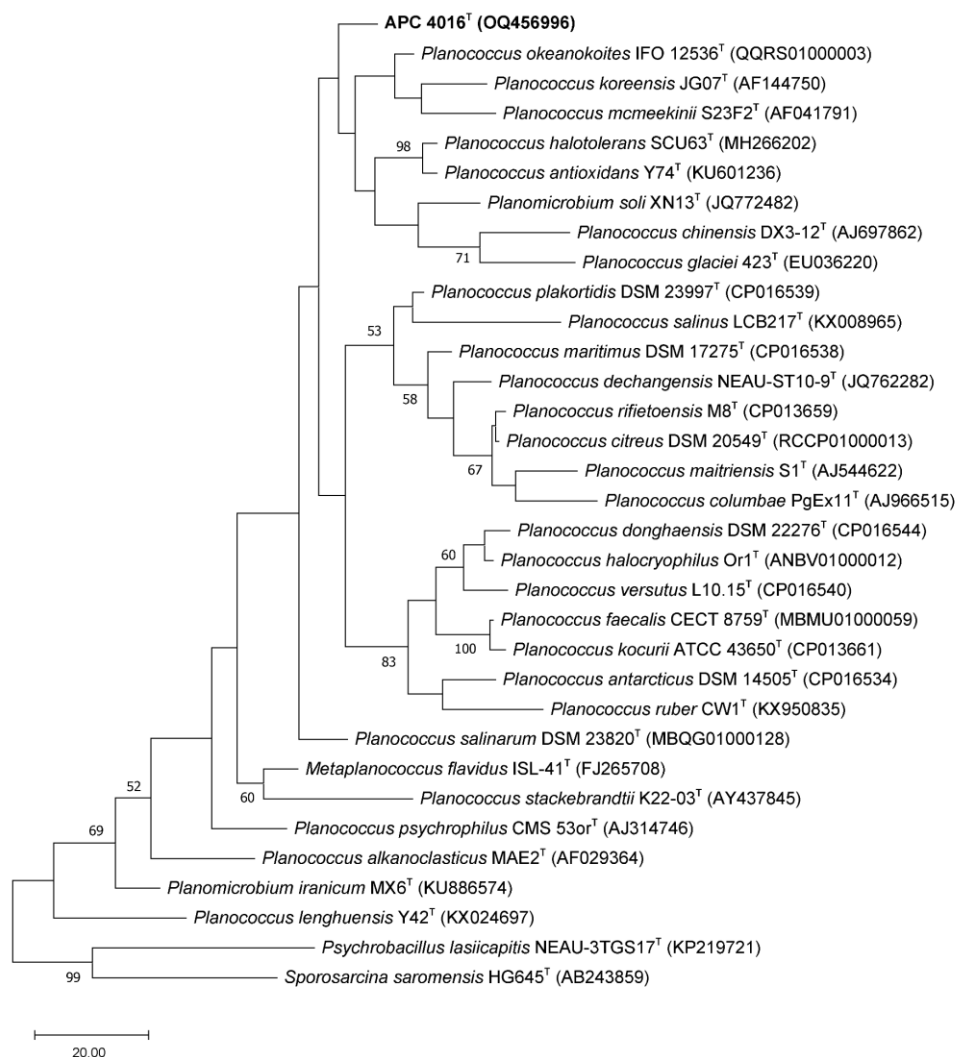


Fig. S2. Max-parsimony evolutionary tree of distances based on 16S rRNA gene sequences from APC 4016^T and related type strains. Bootstrap values based on 1000 replicates (>50 %) are shown. The tree was rooted by 16S rRNA gene sequences from *Psychrobacillus lasiicapitis* NEAU-3TGS17^T and *Sporosarcina saromensis* HG645^T of the family *Caryophanaceae*.

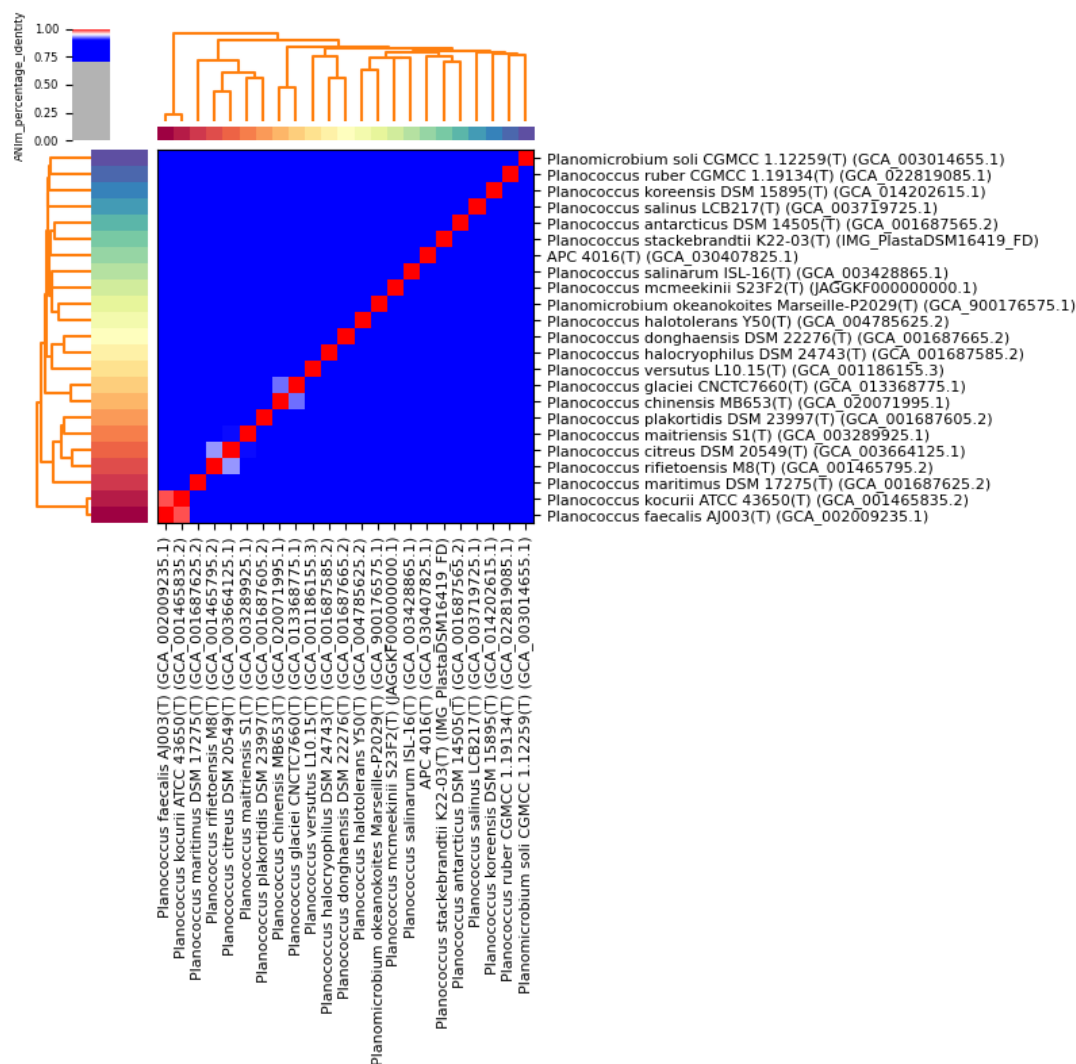


Fig. S3. Pairwise average nucleotide identity (ANI) values between APC 4016^T and all available *Planococcus*/*Planomicrobium* valid type strain genomes from the GenBank and JGI databases. ANI was calculated using PYANI using the MUMer method (ANIm). Blue cells correspond to an ANI value of < 95 %, indicating separate species. Red cells indicate correspond to an ANI value of >95 %, indicating the same species. The colour intensities lighten as the ANI value approaches 95 %. The dendrograms are created by hierarchal clustering of the pairwise ANI values.

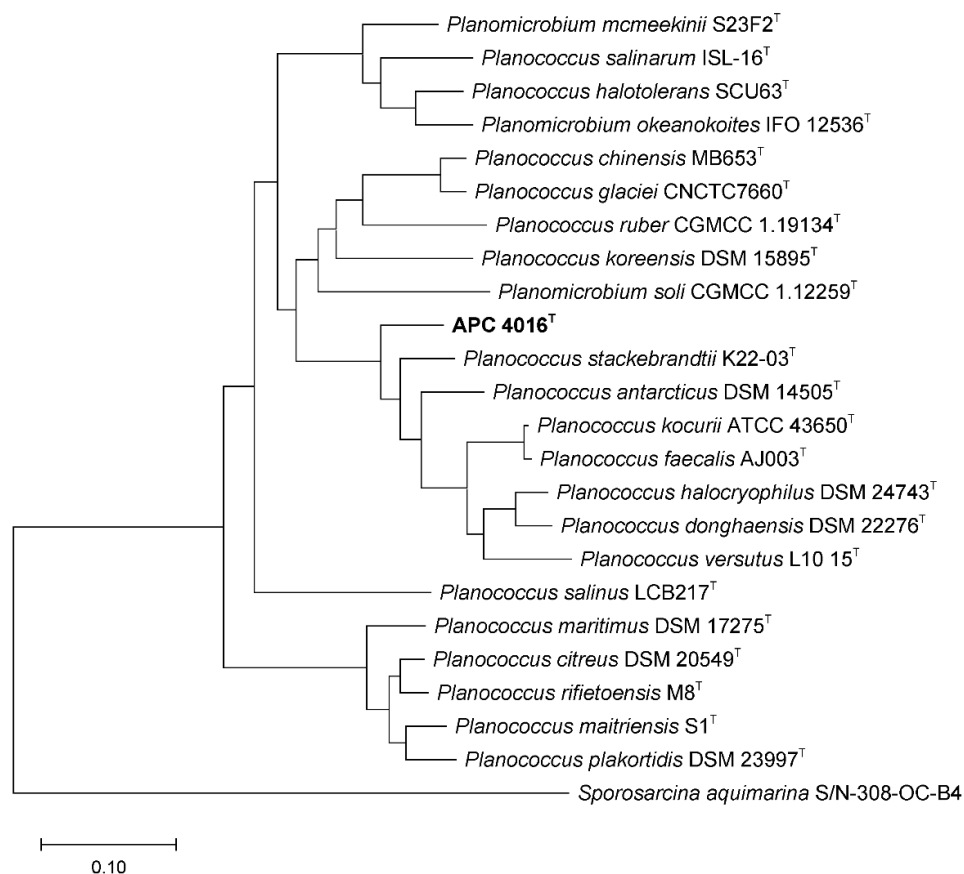


Fig. S4. Resulting maximum-likelihood phylogenetic tree created from the ROARY alignment of 67 core genes from APC 4016^T and *Planococcus/Planomicrobium* type strain reference genomes. *Sporosarcina aquimarina* S/N-308-OC-B4 was included as an outlier group.

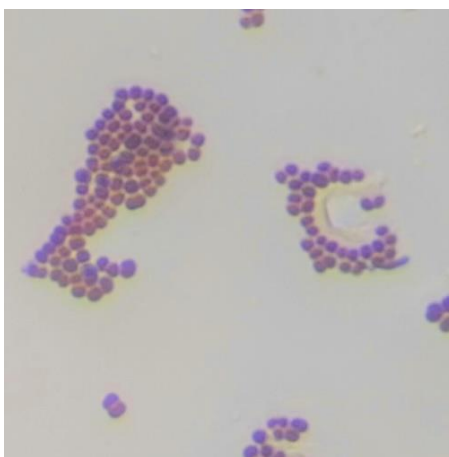


Fig. S5. Cells of strain APC 4016^T after Gram staining. Cells are Gram-positive to Gram-variable.

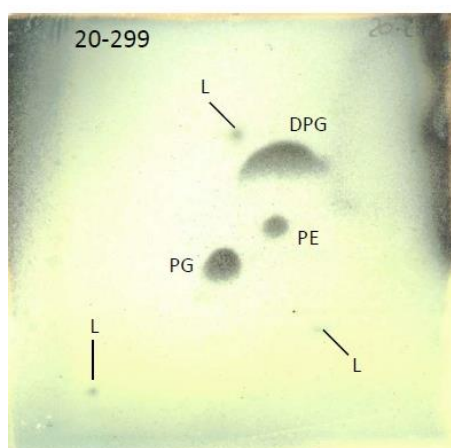


Fig. S6. Cellular polar lipid profile of strain APC 4016^T after separation by two-dimensional thin layer chromatograph. L = lipid; PE = phosphatidylethanolamine; PG = phosphatidylglycerol; DPG = diphosphatidylglycerol.

Table S2. Antibiotic susceptibility profile of strain 4016^T with corresponding zone of inhibition diameters (mm). Susceptibility and resistance were determined by no growth and growth, respectively.

Antibiotic	Disc content	Zone of inhibition (mm)	Result
ampicillin	10 µg	43 ±1.9	S
chloramphenicol	30 µg	33 ±2.5	S
erythromycin	15 µg	43 ±1.0	S
gentamicin	10 µg	20 ±1.6	S
kanamycin	30 µg	18 ±3.3	S
lincomycin	15 µg	10 ±4.0	S
neomycin	30 µg	17 ±5.4	S
novobiocin	5 µg	22 ±1.0	S
oleandomycin	15 µg	24 ±4.9	S
penicillin G	10 U	41 ±1.4	S
polymyxin B	300 U	20 ±0	S
rifampicin	30 µg	41 ±1.0	S
streptomycin	10 µg	16 ±2.6	S
tetracycline	30 µg	37 ±2.5	S

S = susceptible

Appendix to Chapter 4

In the following Appendix SMU-L performed the culture optimization and provided culture preparations for spiking and DNA isolation for the spiking experiments.

Application of *Planococcus* sp. APC 3900 and *Pseudoalteromonas* sp. APC 3896 for quantification of bacterial taxa during human microbiome sequencing

The development of quantified DNA from deep-sea strains, APC 3896 and APC 3900, for use as internal DNA controls in determining absolute abundance of bacteria in human faecal metagenomic sequencing experiments. APC 3896 and APC 3900 were selected on the basis of one being Gram-positive and the other Gram-negative, respectively, and also on their absence in human metagenomic databases. APC 3893 shared 100 % 16S rRNA gene sequence identity to *Pseudoalteromonas issachenkonii* strain KMM 3549 (CP011030.1). APC 3900 shared 100 % identity to *Planococcus notacanthi* APC 4016 (from this study). These strains were not detected in the sequencing data from the non-spiked faecal samples used in this study (Figure 1).

Growth curves to determine optimum growth of the strains at 30 °C were generated (Figures 2 and 3). APC 3896 was estimated to contain 8 16S copy numbers, while APC 3900 was estimated to contain 9. Copy numbers per 1 ng of extracted DNA were calculated. The two marine bacteria in different 16S rDNA copy number concentrations (ranging from 10^4 to 10^8) were added to each sample in a series of experiments in order to identify which concentration would represent 1–10 % of the total bacterial concentration, as measured by 16S rRNA amplicon sequencing (Illumina). Initially these experiments evaluated the utilisation of spike controls in the form of either cells from the two strains or pure genomic DNA (gDNA) into DNA isolated from the mother/infant pair faecal samples.

Initial results using the cell spike-in method show that the output relative abundance of *Pseudoalteromonas* sp. APC 3893 and *Planococcus* sp. APC 3900 were in the range 0.45–2.77 % and 3.17–9.50 %, respectively (Figures 4 and 5). The results outputs from using the cell spike-ins were relatively inconsistent, presumably due to the relative

ease of lysing Gram-negative over Gram-positive bacteria or an error in estimating 16S copy and cell numbers. Therefore, the experimental approach outlined in figure 6 using DNA spike-ins was adopted.

Based on the results from these experiments, we determined that by spiking in 2.5 μL of gDNA from each of the two marine bacteria (in a concentration of 10^7 copy numbers) into 20 μL of extracted sample DNA, the relative abundance of the 16S sequencing amplicons would range between 1 % to 10 % of the total. This was examined first using the DNA spike-in from *Pseudoalteromonas* sp. APC 3896. The relative abundance output of APC 3896 in the spiked samples was consistently detected in the desired range (Figure 7). Optimisation of the *Planococcus* sp. APC 3900 DNA spike and evaluation of these results using qPCR is to be carried out.

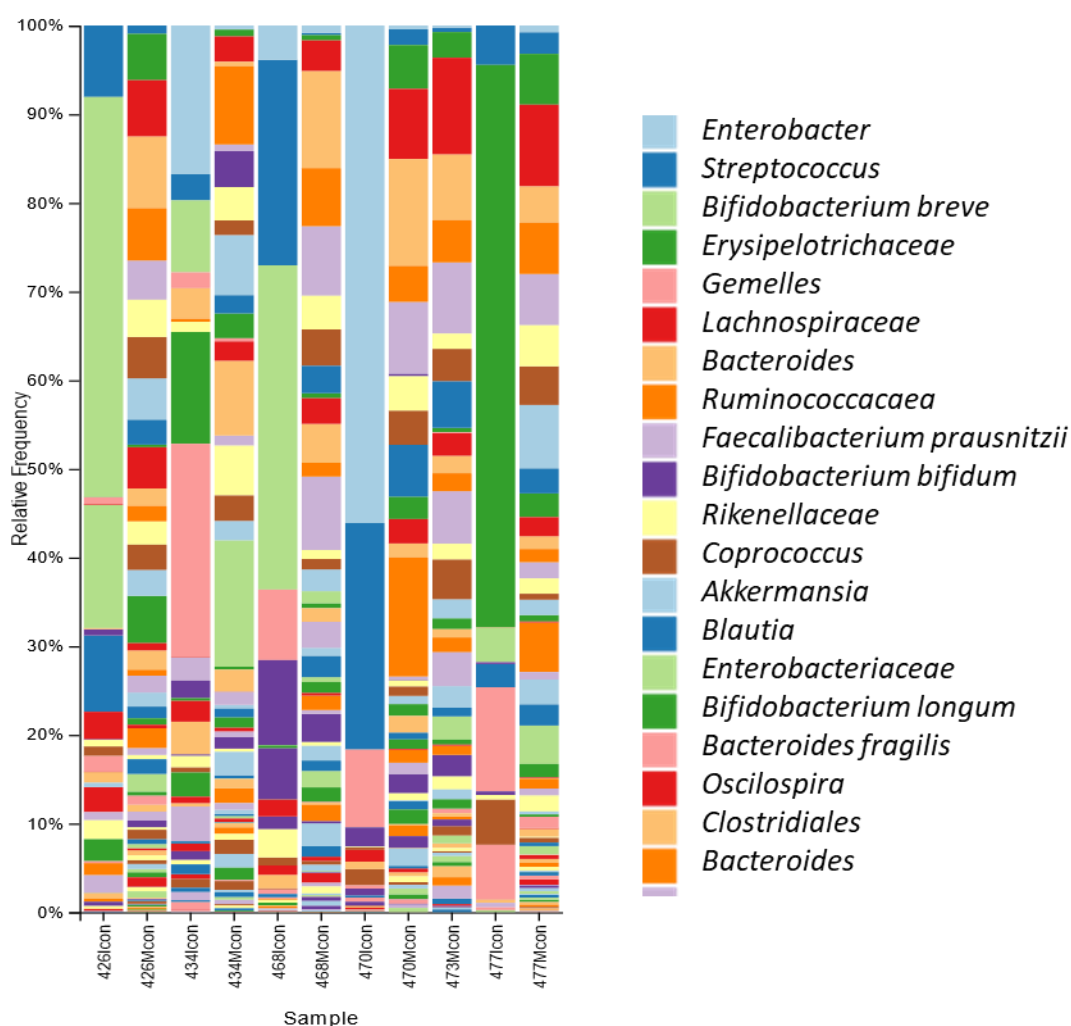


Figure 1: Relative frequency of taxa determined using 16S rRNA amplicon sequencing (Illumina) in control (no spike-in) mother/infant faecal samples. *Planococcus* and *Pseudoalteromonas* sequences were not detected.

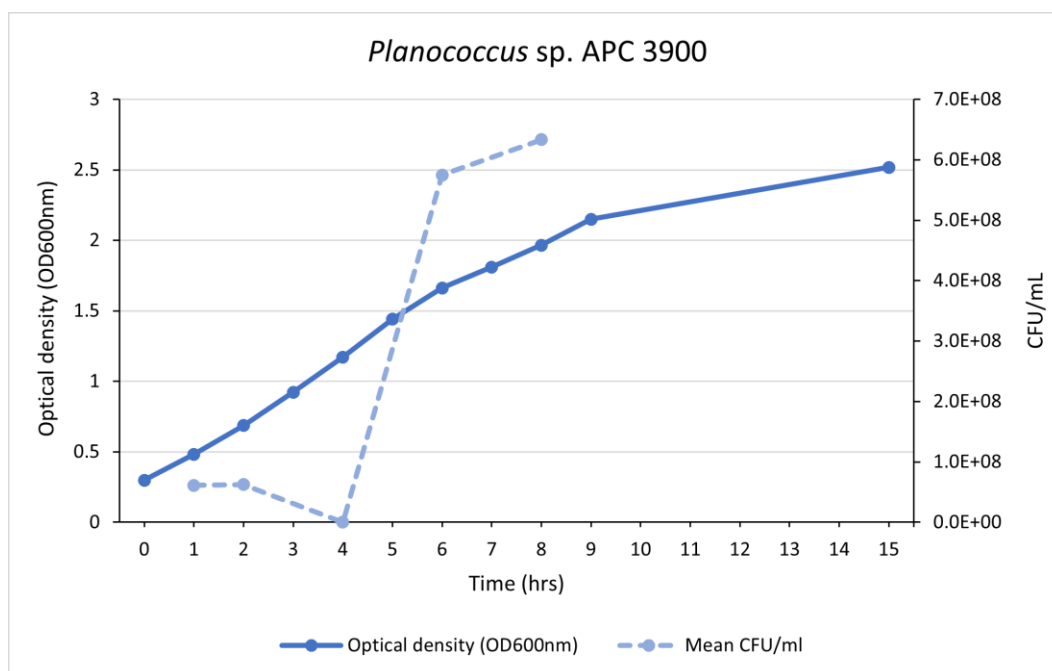


Figure 2: Growth curve of strain *Planococcus* sp. APC 3900 (lab ID: 26E) was used to determine the optimal growth of the strain. Data was generated from CFU/mL counts and measuring optical density at OD_{600nm} over time.

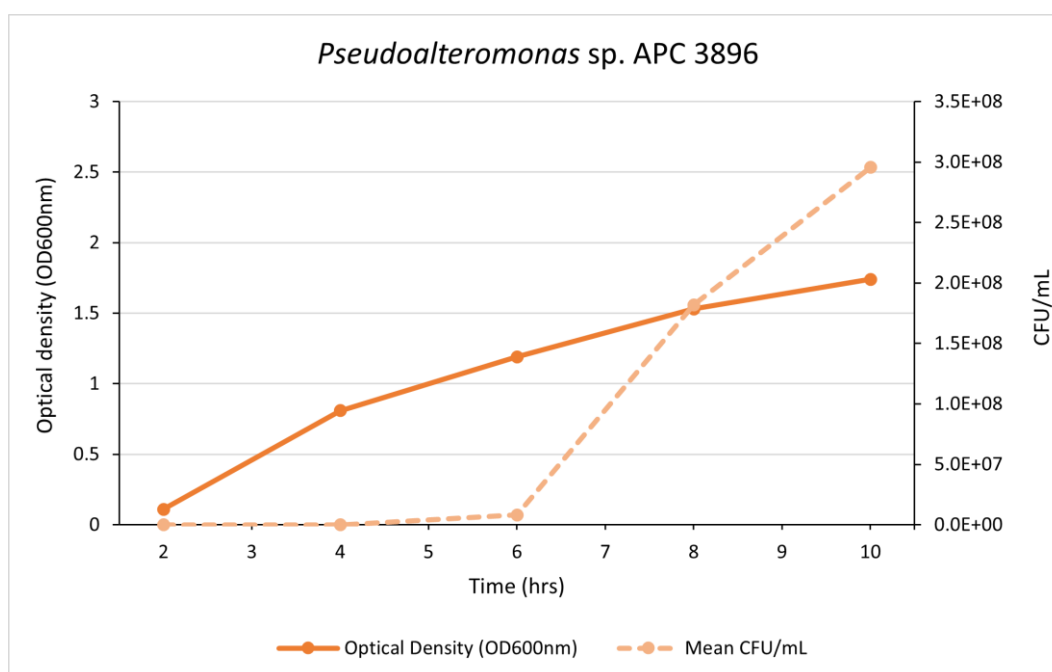


Figure 3: Growth curve of strain *Pseudoalteromonas* sp. APC 3896 (lab ID: 23E) was used to determine the optimal growth of the strain. Data was generated from CFU/mL counts and measuring optical density at OD_{600nm} over time.

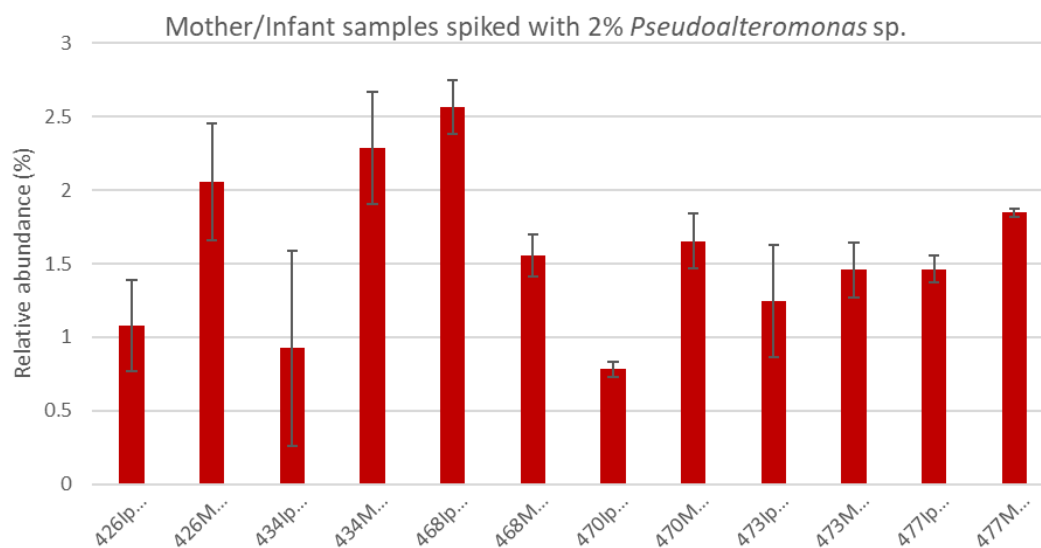


Figure 4: Relative abundance of *Pseudoalteromonas* APC 3896 after spiking cells into mother/infant faecal samples at 2 % estimated output. The actual output mean relative abundance was in the range of 0.45–2.77 %.

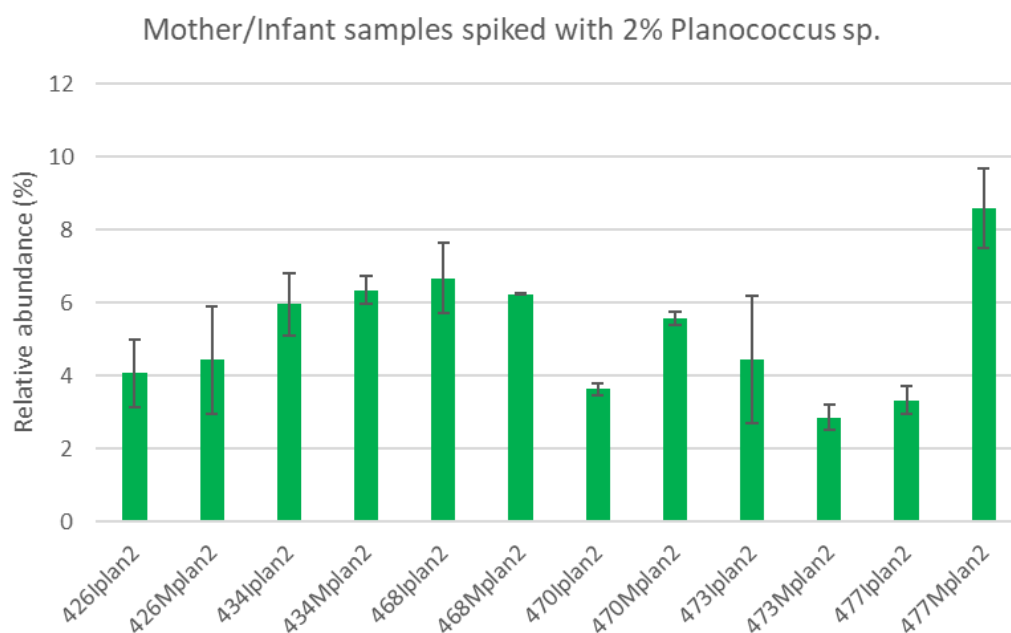


Figure 5: Relative abundance of *Planococcus* sp. APC 3900 after spiking cells into mother/infant faecal samples at 2 % estimated output. The actual output mean relative abundance was detected in the range of 3.17–9.50 %.

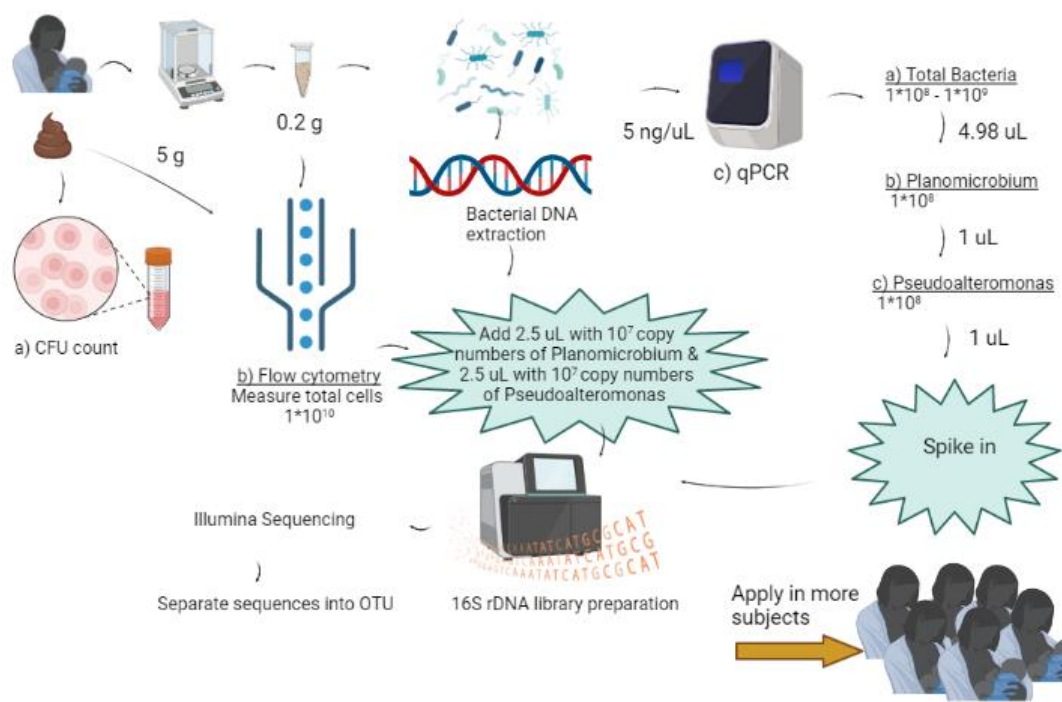


Figure 6: Flowchart showing general overview of the experimental process. Cell counts of deep-sea spike-in DNA were calculated based on copy number per in a known quantity of DNA CFU counts of bacteria from infant/mother faecal samples (a). Flow-cytometry cell counts of known quantity of sample (b). Bacterial genomic DNA is extracted, normalised, and subjected to 16S rDNA HTS, with and without spike-in DNA, which could be evaluated with qPCR (c).

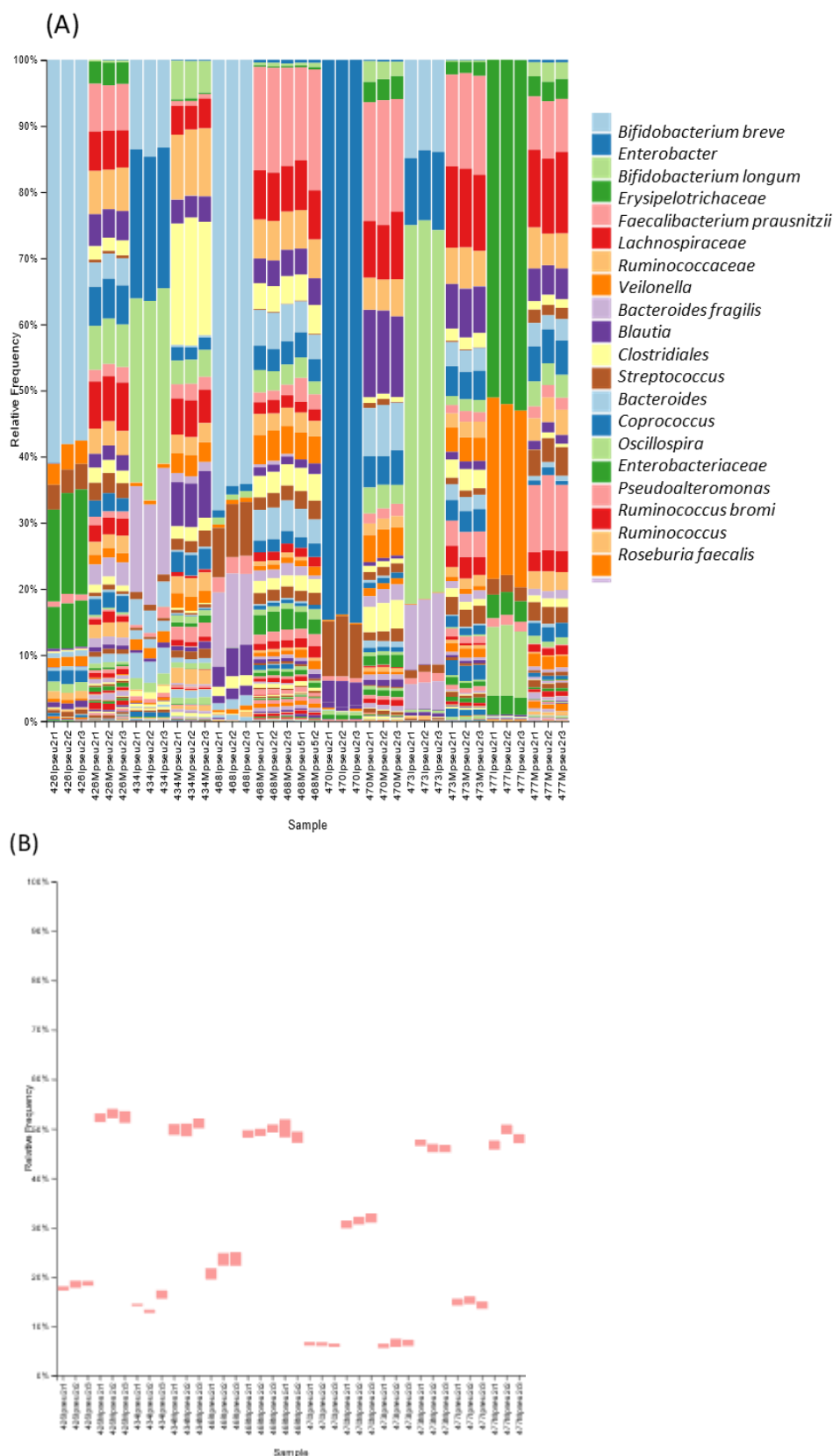


Figure 7: Relative frequency of taxa in mother/infant faecal samples after spiking with 2 % gDNA from *Pseudoalteromonas* sp. APC 3896. The spike-in was detected in each sample in the range of approximately 1–10 %. (a) Full results output, (b) *Pseudoalteromonas* results only.

Chapter 5

***In silico* Characterisation and
Molecular Cloning of Cold-Active β -
agarases from Deep-sea
Pseudoalteromonas sp. APC 4017**

Abstract

In this study we identified an agar-degrading *Pseudoalteromonas* strain, APC 4017, which had previously been isolated from the skin of a deep-sea fish. Crude extract from a culture of this strain retained agar-degrading activity at 5 °C. We identified two putative β -agarase genes, *aga4017A* and *aga4017B*, in the genome of strain APC 4017. The first encodes a protein, Aga4017A, of 445 amino acids in length, which shared high sequence homology to characterised GH16 β -agarases and was predicted to contain conserved features and domains characteristic of GH16 β -agarases. Aga4017B is 773 amino acids in length and was predicted to contain an agarase carbohydrate-binding domain module, characteristic of GH50 β -agarases, yet showed low sequence homology to characterised GH50 β -agarase. A recombinant plasmid pCR2.1-TOPO_aga4017A was created which contained *aga4017A* and was expressed in *E. coli* TOP10. Transformed cells were positive for agar-degradation. The 3D structural protein model of Aga4017A was predicted using Alphafold2 which demonstrated that the structure was organised into three distinct modules corresponding to the predicted signal peptide sequence, conserved GH16 β -agarase catalytic domain and the carbohydrate-binding domain. This study represents one of the first to characterise a β -agarase protein sequence found in a deep-sea fish microbiome isolate.

5.1. Introduction

Agar is a sulphated galactan composed mainly of two polysaccharides, agarose and agarpectin, and constitutes part of the cell wall of red algae (Rhodophyta). Agarose itself is a linear molecule consisting of alternating repeats of 3,6-anhydro- α -L-galactose and β -D-galactose residues linked by α -1,3 and β -1,4 bonds (Chi et al., 2012). Agar-degrading enzymes (agarases) are members of the glycoside hydrolase (GH) enzyme class, and act on agarose. They are classified into two main groups: α -agarases (EC 3.2.1.158) that cleave the α -1,3 linkages of agarose to yield agar-oligosaccharides (Flament et al., 2007), and β -agarases (EC 3.2.1.81) that cleave the β -1,4 linkages of agarose to yield neoagar-oligosaccharides (NAOs) (Ohta et al., 2004). According to the database of carbohydrate-active enzymes (CAZy; [<http://www.cazy.org>]) agarases can be further classified into 5 specific glycoside hydrolase families (GH96 for α -agarase, GH16, GH50, GH86, and GH118 for β -agarase) based on their amino acid sequence, and which can generate end products with varying degrees of polymerisation (Li et al., 2018).

Agarases can be an important biological tool in research, for example, in the recovery of oligonucleotides and nucleic acid fragments from agarose gels, and to investigate the structure of galactans and seaweed cell walls (Yaphe, 1957, Fleurence et al., 1995). Furthermore, the oligosaccharide products derived from agar degradation, particularly NAOs, can have a wide variety of health-benefiting biological activities such as anti-inflammatory (Wang et al., 2017), anti-oxidant (Jonnadula et al., 2020), cell whitening, moisturising (Yun et al., 2013), prebiotic (Hu et al., 2006), anti-obesity and anti-diabetic (Hong et al., 2017). The enzymatic degradation of agar is often preferred to chemical degradation due to the reaction being more specific, environmentally friendly, and possibly more effective at producing more functional NAOs (Xu et al., 2018, Chen et al., 2021). Much of current research is investigating the expression of agarase genes in heterologous hosts, such as *Escherichia coli* and *Bacillus subtilis*, to potentially produce commercial food-grade products (Cui et al., 2018). Current commercially-available β -agarases include native and recombinant products from *Pseudoalteromonas* and *Zobellia* species.

Agar-degrading bacteria have been isolated from a wide variety of marine habitats, such as seawater, sediment, soil, and algae (Fu and Kim, 2010). Deep-sea habitats are characterised by extreme environmental conditions such as low temperatures, no sunlight and high hydrostatic pressure and therefore are particularly interesting as sources of novel enzymes with unique properties. Novel β -agarases have been isolated from deep-sea sediment (Chen et al., 2016) and deep-sea waters (Ohta et al., 2004). However, to date, no studies have reported the characterisation or isolation of agarases from deep-sea fish microbiota.

In this study, we describe the identification and *in silico* characterisation of two putative β -agarases encoded by *Pseudoalteromonas* sp. APC 4017, an isolate from the skin of a deep-sea fish. We also predict the 3D protein model structure of β -agarase 4017A using *in silico* methods and describe its subsequent cloning and expression in the heterologous host *E. coli*.

5.2. Materials and Methods

5.2.1. Bacterial isolates and identification of agarolytic strains

Bacterial strains were originally isolated as part of a screening study for antimicrobial-producing isolates from the skin and intestines of deep-sea fish (Uniacke-Lowe et al., 2023). It was noted that a number of the isolates demonstrated agar “pitting” when cultured on BD Difco Marine Agar 2216 (MA). One of the agarolytic isolates, strain APC 4017, was selected for further analysis for this study.

5.2.2. Taxonomic analysis:

Strain APC 4017 was subjected to genomic DNA (gDNA) extraction, sequencing and assembly, as previously described (Uniacke-Lowe et al., 2023). The genome assembly sequence is available from NCBI under the accession JASDCC000000000.1. The 16S rRNA gene sequence was extracted from the whole genome sequence data using barrnap (version 0.9, <https://github.com/tseemann/barrnap>). 16S rRNA gene -based taxonomy was inferred using the 16S-based ID application on the EzBioCloud server (Yoon et al., 2017). Multiple sequence alignments with the query and hit sequences were created using Muscle (Edgar, 2004) within MEGA X (Kumar et al., 2018). A neighbour-joining tree was constructed and inferred using the Kimura-2-parameter method (Kimura, 1980). Bootstrapping was based on 500 replicates. The tree was rooted by including the 16S rRNA gene sequence from *Algicola bacteriolytica* E8R (accession NR_036838.1) as an outlier.

5.2.3. Preparation of crude enzyme agar extract

Crude supernatant containing the agarase enzymes was prepared using culture freeze-thaw extraction as follows: An overnight culture of APC 4017 was prepared in BD Difco Marine Broth 2216 (MB). A dense lawn of culture was then spread onto MA (1.5 % agar) and cultured at 25 °C for 4 days. Following incubation, the plate was frozen at –20 °C overnight and allowed to thaw. Liquid from the thawed plate was collected and centrifuged at 4500 x g for 20 min at 5 °C. This freeze-thaw process was repeated several times until no more liquid could be extracted. Cells were removed by filtration using a 0.45µm filter (Filtropur S PES, Sarstedt). The cell-free extract was concentrated by ultra filtration using a molecular weight cut-off (MWCO) of 10K

(Pierce Protein Concentrator PES, 10K MWCO, Thermo Fisher Scientific) and the retentate was resuspended in half the volume of phosphate buffer saline solution (PBS) to achieve a 2X concentration. Aliquots of the retentate in PBS and filtrate were stored at 4 °C (short term) and at –20 °C (long term).

To confirm the presence of agarase activity, the retentate and filtrate were spotted (15µL) onto 2 % agar supplemented with 2 % w/v NaCl and incubated at 5 °C and 25 °C overnight. The plate was stained by flooding with Lugol's iodine and observed for zones of clearance (occurring within 2-3 minutes).

5.2.4. *In silico* agarase gene identification and characterisation

Bacterial genome assembly was annotated using PROKKA (version 1.14.6; Seemann, 2014). Agarase protein sequences predicted from the annotation were queried using NCBI's BLASTp tool against the non-redundant (nr) and Protein DataBank (PDB) databases. Multiple sequence alignments (MSAs) were made in Jalview (version 2.11.2.7; Waterhouse et al., 2009) using Muscle (Edgar, 2004). Pairwise amino acid percentage identities were calculated within Jalview. Signal sequences of the putative agarase proteins was predicted using SignalP (version 6.0; Teufel et al., 2022). Molecular mass and theoretical isoelectric point (pI) were predicted using the ExPASy-ProtParam tool (<https://web.expasy.org/protparam/>, accessed on 11/09/2023). Conserved domains and features were queried using NCBI's web-based conserved domain search tool against the Conserved Domain Database (CDD, version 3.2; Yang et al., 2020).

5.2.5. Cloning

5.2.5.1. Agarase gene PCR amplification

Genomic DNA (gDNA) from strain APC 4017 was used as the template for agarase gene amplification following extraction using the GeneJET Genomic DNA Purification Kit (Thermo Fisher Scientific) and Gram-negative protocol. The primers used for amplification of the β -agarase genes are shown in Table 1. Amplification of the agarase genes was carried out using Biomix Red DNA polymerase (Meridian bioscience) and PCR with the following parameters: initial denaturation, 95 °C x 60s; 32 PCR cycles (95 °C for 10s, anneal* at 53 °C for 20s, 72 °C for 70s); final extension, 72 °C for 10 min. PCR products were analysed by running on a 1 % agarose gel, stained

with MIDORI Green and subject to gel electrophoresis. PCR products were purified using the GeneJET PCR purification kit (Thermo Fisher Scientific).

*A gradient PCR using annealing temperatures of 55 ± 5 °C was initially carried out to optimise the annealing temperature.

5.2.5.2. Competent cells and cloning vector

One Shot™ TOP10 chemically competent *E. coli* cells, pCR2.1-TOPO TA cloning vector and TOPO-TA cloning kit were acquired from Invitrogen (cat# K4500J10, Thermo Fisher Scientific).

5.2.5.3. Plasmid construction and transformation

Production of PCR products for the cloning reaction was carried out according to the TOPO TA Cloning kit manual using gDNA from APC 4017 as the template, the agarase primers (Table 1) and PCR cycle parameters as above. The cloning reaction was carried out using the TOPO TA blunt-end cloning kit and method for TOP10 chemically competent cells (reaction time: 15 min), according to manufacturer's instructions. Transformation was carried out using the regular transformation protocol for One Shot chemically competent *E. coli*. Resulting cells from the transformation reaction were plated on LB containing 50 µg/mL ampicillin and X-Gal (40 µL of 40mg/mL spread onto plate) and incubated overnight at 37 °C. Following incubation, several positive clones (white colonies) were picked and subcultured on LB medium containing 50 µg/mL ampicillin.

5.2.5.4. Transformant analysis

Colony PCR on selected clones was used to confirm the presence of the agarase gene using the relevant insert primers (Table 1).

5.2.5.5. Plasmid extraction, sequencing, and assembly

A single colony from a transformant positive for agar-degradation (and insertion of fragment) was selected for plasmid extraction using the PureLink Quick Plasmid Miniprep Kit (Invitrogen), as per manufacturer's guidelines. Plasmid DNA was sequenced, assembled, and annotated by PlasmidSaurus (<https://www.plasmidsaurus.com/>). The annotated plasmid assembly was visualised in SnapGene. The 4017A agarase gene was mapped onto the resulting plasmid assembly following nucleotide alignment in SnapGene.

5.2.6. Predictive 3D modelling

The predicted 3D protein model structures were generated using AlphaFold2 (Jumper et al., 2021) on our laboratory's local server and visualised in Mol* Viewer (Sehna et al., 2021). 3D model confidence was determined using AlphaFold's per-residue confidence metric: predicted local distance difference test (pLDDT). Conserved features were mapped onto the 3D structure sequence in Mol* Viewer. The predicted 3D structure model of Aga4017A was superimposed with the crystal structure model of the catalytic β -agarase domain from *Microbulbifer thermotolerans* JAMB-A94 (PDB 3WZ1) using PyMOL (The PyMOL Molecular Graphics System, version 2.5.5, Schrödinger, LLC).

5.3. Results

5.3.1. Identification and selection of agarolytic strains

Colonies of marine fish bacterial isolates displayed agar pitting after culturing on MA, which when stained with Lugol's iodine, resulted in clear zones in the areas where the agar had been degraded (Figure 1a). APC 4017, originally isolated from the skin of *Bathysaurus ferox*, exhibited suspected agarase activity after culturing on MA at 21 ± 2 °C for up to 10 days and at 5 ± 1 °C for up to 3 weeks (Figure 1b). Based on the 16S rRNA gene analysis, strain APC 4017 was most closely related to *Pseudoalteromonas carrageenovora* IAM 12662^T (accession no. X82136) at 99.12 % sequence similarity (97.7 % completeness) and *Pseudoalteromonas atlantica* NBRC 103033^T (BJUT01000111) at 99.0 % similarity (100 % completeness). APC 4017 formed its own clade in a rooted neighbour-joining tree of 16S rRNA gene sequences, indicating that it may be a novel species (Figure S1).

5.3.2. Crude agarase agar extract

Crude agarase prepared from the extract of strain APC 4017 cultured on solid media was spotted onto 2 % agar and incubated overnight. After staining with Lugol's iodine clear/lighter coloured zones were formed where the supernatant had been spotted, demonstrating agar degradation. This activity was observed when the spotted supernatant was incubated at 5 °C and 25 °C (Figure 2).

5.3.3. Molecular characterisation of Aga4017A

From the PROKKA genome annotation, a putative β -agarase-encoding gene of 1338 bp, denoted *aga4017A*, was detected in the genome of the strain APC 4017. This gene encoded a 445 amino acid product. The molecular mass and pI were predicted to be 50.21 kDa and 4.75, respectively. A Sec-secretion signal peptide cleaved by signal peptidase I (Sec/SPI) and a cut site were predicted from residues 1 to 19 and between residues 19 and 20, respectively (Figure S2). Specific conserved domains were predicted in the Aga4017A sequence including a GH16 family β -agarase domain (GH16_beta_agarase [cd02178]) and a carbohydrate-binding ricin-type beta-trefoil lectin domain (Ricin B lectin [Pfam PF00652]) from residue 18 to 294 and from 308 to 441, respectively (Table 2). The full list of CDD hits identified is given in Table S1. The

start position of the GH16_beta_agarase domain (residue 18) is consistent with the cut site proceeding the predicted signal peptide position (residue 19).

Five conserved features were predicted in the Aga4017A protein sequence, including active sites, catalytic residues, and sugar-binding sites (Table 3). Regarding the predicted Q-X-W motifs in the Aga4017A sequence, tryptophan (W) has been replaced by phenylalanine (F), both of which are hydrophobic amino acids.

BLASTp analysis against the nr database revealed that Aga4017A has the highest % identity to several uncharacterised ricin domain-containing proteins from *Pseudoalteromonas* spp. and to Aga862, a characterised GH16 family β -agarase from *Pseudoalteromonas* sp. Q30F (AXK59832.1) at 93.48 %. Other characterised hits included β -agarases from *Pseudoalteromonas* sp. H9 (AIW52355.1), *Janthinobacterium* sp. SY12 (ABO93616.2) and *Pseudoalteromonas* sp. JT-6 (ACF40223.1). All sequences were 445 amino acids in length and the % identities ranged from 86.97 to 96.63 % (Table 4).

The multiple sequence alignment of Aga4017A and the characterised hit protein sequences showed that the region of highest similarity among the sequences occurs between residues 33 to 300 (approximately), which corresponds to the catalytic agarase domain according to the CDD analysis (Figure 3). The least conserved residues were found in the predicted carbohydrate-binding (C terminal) domain. The MSA also shows the presence of a highly conserved sequence pattern (EXDXXE), which has previously been reported to contain key catalytic residues that are characteristic of GH16 family β -agarases (Michel et al., 2006). The predicted catalytic residues of Aga4017A (E143, D145, E148 [Table 3]) most likely correspond to this sequence pattern.

When queried against the PDB database, Aga4017A shared the highest % identity to the catalytic domain of a GH16 β -agarase from *Microbulbifer thermotolerans* (3WZ1), a deep-sea sediment isolate (Takagi et al., 2015, Ohta et al., 2004). The sequences shared 42.07 % identity at 62 % coverage, according to the BLASTp results.

A putative protein coding sequence (CDS) of 2322 bp encoding a product of 773 amino acids was identified immediately downstream of the *aga4017A*. Both genes

appeared to be located within a large gene cluster encoding enzymes involved in hydrolysis of agarose and agarose-derivatives, such as galactose (Figure 4).

5.3.4. Molecular characterisation of Aga4017B

BLASTp analysis of this second gene, denoted *aga4017B*, found that it was most similar to putative uncharacterised agarase_CBM(carbohydrate-binding module)-like domain-containing proteins from various *Pseudoalteromonas* species. Identities ranged from 83.69-98.58 % (at 100 % coverage). The agarase_CBM domain (Pfam PF17992) forms the active site of exo β -agarases (Pluvinau et al., 2013). When queried against the PDB database, Aga4017B was most closely related to a GH50 family β -agarase from *Pseudoalteromonas fuliginea* (PDB 6XJ9) at 51.81 % identity and 91 % coverage, followed by β -agarases from *Agarivorans gilvus* (PDB 5Z6P) and *Saccharophagus degradans* 2-40 (PDB 4BQ2 and PDB 4BQ4) and at 50.71 %, 49.87 % and 49.73 % identity, respectively (Figure 5).

Aga4017B had an estimated molecular mass and pI of 88.02 kDa and 6.22, respectively. A Sec lipoprotein signal peptide cleaved by signal peptidase II (Sec/SPII) was predicted from residues 1 to 22 with cut site occurring between residues 22 and 23 (Figure S3). A conserved agarase_CBM domain (PF17992, superfamily cl40253) was predicted from residues 46 to 223. No conserved features or functional sites were predicted in Aga4017B.

Pairwise sequence analysis of Aga4017A and Aga4017B found that they shared only 20 % identity suggesting that they represent very different GH families. An overview of the predicted molecular characteristics of Aga4017A and Aga4017B, including molecular mass (kDa), pI and signal peptide prediction, is given in Table 6.

5.3.5. Heterologous expression of Aga4017A

The agarase PCR products were each inserted into a pCR2.1-TOPO vector and subsequently used to transform the *E. coli* TOP10 cells. Transformation to generate *E. coli* TOP10 pCR2.1-TOPO_aga4017A resulted in colonies that exhibited agar pitting (Figure 6), confirming that Aga4017A is an agar-hydrolytic enzyme. The agar-pitting, and hence enzyme expression, was observed without induction. Colony PCR on randomly selected transformants exhibiting agar-pitting confirmed the presence of

the *aga4017A* gene (Figure 7). Plasmid sequencing and assembly produced a plasmid 5324 bp in size and contained the insert of interest, *aga4017A*, on the reverse strand (Figure 8).

5.3.6. Predicted 3D structure of GH16 agarase Aga4017A

The Aga4017A protein was selected for further analysis. The predictive 3D protein modelling of Aga4017A using AlphaFold2 produced a globular protein with distinctive modules, which correspond to the conserved domains predicted from the NCBI CDD search (Figure 9). The secondary structure of the predicted model largely consisted of β -strands (Figure 10).

The pLDDT scoring per residues plotted within AlphaFold is illustrated in Figure S4. A pLDDT score of >90 is considered the cut-off for high prediction accuracy, whilst a cut off of >70 corresponds to a “generally correct” prediction (Tunyasuvunakool et al., 2021). Overall, 81 % of the residues in the Aga4017A-predicted model yielded a pLDDT score of >90, whilst 94 % had a score of >70, indicating a strong model confidence. The lowest scoring region was seen within the first 18 residues (mean score of 56), which corresponds to the predicted signal sequence of 19 residues. A second low scoring region occurred between residues 296-306 (average score of 59), which likely corresponds to the linker region between the agarase and ricin B lectin - like carbohydrate binding domain of the Aga4017A protein.

Coordinates of the predicted functional sites were mapped onto the predicted 3D model (Figure 11). The active sites and catalytic residues were localised within the GH16 β -agarase domain. More specifically, the active sites were located within a concave channel/groove formed by β -strands in this protein domain.

The observed 3D structure and arrangement of the β -strands of the predicted carbohydrate-binding domain of Aga4017A is consistent with it being a β -trefoil fold (Pfam PFCL0066). This type of fold, which largely consists of β -strands forming six distinctive two-stranded hairpins, is found in the carbohydrate-binding module (CBM) family 13. The CBM13 family belongs to the same Pfam family as ricin B lectin (Fujimoto, 2013). This structural model result is in agreement with the predicted conserved domain results derived from the protein sequence analysis (Table 2). The localisation of the predicted sugar-binding sites to within the ricin B lectin domain of

4017A is also consistent with it being a carbohydrate-binding domain. The calcium-binding sites were also concentrated within the ricin B lectin -like domain (Figure 11). When superposed with the crystal structure of the GH16 domain of β -agarase MtAgaA (PDB 3WZ1), from *M. thermotolerans* JAMB-A94, a clear overlapping of the structures could be seen, demonstrating their high structural similarity (Figure 12). Both structures contained the distinctive concave region containing the active sites. Crucially, the highly conserved EXDXXE catalytic residues within this concave region aligned completely. However, when aligned with Muscle, the amino acid sequences only shared 41.08 % identity over 297 amino acids (Figure S5).

5.4. Discussion

The aim of this study was to characterise putative agar-degrading enzymes identified in bacteria isolated from deep-sea fish. Strain APC 4017 was deemed to be a good candidate for further analysis. Extreme environments are of great bioprospecting interest in terms of finding novel enzymes that are active, if not at least stable, under a wide range of conditions such as extreme of temperatures, pH, hydrostatic pressure, and variable salinity (Littlechild, 2015, Mesbah, 2022). *Pseudoalteromonas* species, in particular, are of great biotechnological interest due to their tolerance of extreme environmental conditions and ability to produce a wide range of bioactive compounds (Parrilli et al., 2021). 16S rRNA gene sequence analysis showed that strain APC 4017 is a member of the genus *Pseudoalteromonas*, and most closely related to *P. carrageenovora* IAM 12662^T and *P. atlantica* NBRC 103033^T, two species known for their ability to degrade various polysaccharides, including agar (MORRICE et al., 1983) carrageenan (Gobet et al., 2018). The topology of the phylogenetic tree generated from alignment of the 16S rRNA gene sequences suggests that APC 4017 is a novel *Pseudoalteromonas* species, as it does not cluster closely to the other sequences, however further analysis is needed to confirm this.

Crude supernatant from APC 4017 exhibited agar-degrading activity at low temperatures of 5 °C and 25 °C. This may indicate an adaptation to the low temperatures experienced in the deep sea. Cold-active enzymes may be particularly desirable for their efficiency at lower temperatures and, therefore, can be used in more cost effective and environmentally friendly reactions (Barroca et al., 2017). Cold-active enzymes can also be used in the preparation of thermolabile food substances, such as the production of lactose-free milk using β -galactosidase (Mangiagalli and Lotti, 2021). Cold-active β -agarases, specifically, have applications for carrying out hydrolysis of agar/seaweed compounds and production of NAOs under low-temperature conditions. Current research is also looking into the application of agarases in the development of seaweed-derived greener renewables, such as biodegradable plastics (Farghali et al., 2023) and biofuels (Park et al., 2020). Preliminary *in silico* analysis of the genome of strain APC 4017 found two agarase genes, denoted *aga4017A* and *aga4017B*, encoded in tandem. The protein sequence

of Aga4017A was 445 amino acids in length, with an estimated molecular mass of 50.21 kDa. The most closely related characterised agarases to Aga4017A from the NCBI nr and PDB databases were Aga862 from *Pseudoalteromonas* sp. Q30F (AXK59832.1), at 93.48 % identity, and MtAgaA from *Microbulbifer thermotolerans* JAMB-A94 (3WZ1) at 42.07 % identity, respectively, both of which are GH16 family β -agarases (Cui et al., 2018, Ohta et al., 2004). Notably, *M. thermotolerans* JAMB-A94 is also a deep-sea isolate, having been originally isolated from deep-sea sediment. Conserved domains and features predicted in Aga4017A included an N-terminal secretory signal peptide (Sec/SPI) sequence, an GH16 β -agarase domain, a C-terminal ricin-type beta-trefoil lectin (carbohydrate-binding) domain, and a highly conserved catalytic sequence pattern (EXDXXE) containing glutamic acid and aspartic acid residues, which is characteristic of GH16 family β -agarases (Michel et al., 2006). Mutations to these residues in β -agarase Aga0917 resulted in improper binding of the substrate to the catalytic sites and loss of hydrolytic activity (Li et al., 2022). An alignment generated from the sequences of Aga4017A and closely related characterised β -agarases from the NCBI nr database showed a high level of sequence similarity, particularly between residues 33 to 300, the region corresponding to the agarase domain. The least conserved residues were found in the N-terminal signal sequence and C-terminal carbohydrate-binding domain, possibly indicating differences in substrate binding affinity.

The second agarase of APC 4017, Aga4017B, was much larger, at 773 amino acids in length and with an estimated molecular mass of 88.02 kDa. The sequence was most closely related to PfGH50B, a GH50 β -agarase from *Pseudoalteromonas fuliginea* (PDB 6XJ9), at only 51.91 % identity. The structure of PfGH50B consists of an N-terminal carbohydrate binding module, a linker region and a larger C-terminal catalytic domain and was noted as being unique amongst the GH50 β -agarases (Pluvinaige et al., 2020). Aga4017B was predicted to contain a lipoprotein (Sec/SPII) signal sequence and a conserved agarase_CBM domain from residues 46 to 223, consistent with corresponding to the N-terminal CBM domain of PfGH50B from *P. fuliginea*. Interestingly, no other conserved features or functional sites were predicted in the Aga4017B sequence. In the BLASTp sequence similarity search

results all identified hits were hypothetical/uncharacterised agarase_CBM domain-containing proteins. This suggests that Aga4017B is a novel GH50 β -agarase.

Agar-degrading systems involving multiple β -agarases has been reported in many bacterial species and is understood to enable progressive degradation of agarose to small oligosaccharides (Ekborg et al., 2006). Several more recent genomics-based studies of agarolytic strains reported large (>100 kbp) 'polysaccharide utilisation loci' (PUL) that encoded multiple agarases representing different GH families, and other hydrolytic enzymes (Schultz-Johansen et al., 2018). Kang et al. (2023), for example, showed that the PUL of *Aquimarina* sp. ERC-38 included six β -agarases (three GH86 and three GH16 family β -agarases). Thin layer chromatography analysis showed that the mechanism of action of these agarases was different, with some being more effective at hydrolysing NAOs of different lengths, demonstrating substrate specificity. It could also be seen that several of these agarase were located within very close proximity of each other on the PUL. It is therefore likely that Aga4017A and Aga4017B are involved in a complex agarolytic system involving multiple enzymes.

The *aga4017A* and *aga4017B* genes were cloned into *E. coli* TOP10 using the pCR2.1-TOPO cloning vector. Significantly, pCR2.1-TOPO_aga4017A transformants were positive for agarolytic activity despite lacking dedicated promoter for expression of the insert. Isolation and subsequent sequencing of pCR2.1-TOPO_aga4017A confirmed insertion of the complete *aga4017A* gene on the reverse strand. Aga4017A was selected for further characterisation due to its effective agar-degrading activity and apparent non-toxic nature to the *E. coli* host.

The predicted 3D model of Aga4017A had a globular structure but with distinctive modules in the N- and C-terminal domains formed by the tertiary structure. Mapping of the predicted conserved domains from the NCBI CDD search onto the predicted 3D model confirmed these modular regions as the N-terminal GH16 catalytic and ricin B lectin (carbohydrate-binding) domains. The topology of the GH16 catalytic domain formed a concave groove and contained the predicted active sites and catalytic sequences, which is consistent with previous reports on GH16-family enzyme structures (Michel et al., 2006). It has been proposed that the active site of glycoside hydrolases can be one of three general topologies: pocket, tunnel and cleft/groove, with the latter being most commonly found on endo-acting enzymes (Davies and

Henrissat, 1995), suggesting that Aga4017A is an endo-type β -agarase. This was further supported by our superposition of Aga4017A with the crystal structure of the catalytic domain of the deep-sea agarase MtAgaA (*M. thermotolerans* JAMB-A94), in which a clear overlapping of the secondary structures, concave regions and key catalytic residues was observed. MtAgaA is a highly thermostable endo-type GH16 β -agarase that produces neoagarotetraose as its end product (Ohta et al., 2004). The topology of the C-terminal module of the Aga4017A 3D model was consistent with being a ricin B lectin trefoil fold (Pfam pf00652), which agrees with the NCBI CDD search. An interesting aspect of the APC 4017 agarases to explore, which we have not done in this study, would be their full temperature range of activity, thermostability, and hydrostatic pressure tolerance. The results are consistent and strongly suggest that Aga4017A is an endo-type GH16 β -agarase.

5.5. Conclusions

In this study we identified a potentially novel agar-degrading strain, *Pseudoalteromonas* sp. APC 4017, originally isolated from the gut of a deep-sea fish. Supernatant from strain APC 4017 demonstrated agar hydrolysis at low temperatures which suggests that it may have applications for industrial processing of agar and agar-containing biological substrates. Two putative β -agarases, Aga4017A and Aga4017B, were identified in the genome of APC 4017, which were found to be encoded, in tandem, on a large agarolytic gene cluster. Sequence analysis showed that Aga4017A was most likely a GH16 family β -agarase and contained conserved characteristic GH16 domains and catalytic residues. *E. coli* TOP10 transformants expressing the Aga4017A were capable of agar degradation as evidenced by liquid pools around the colonies. Sequence analysis and 3D modelling showed Aga4017A to be a modular enzyme, with distinctive domains involved in signalling, biocatalysis and carbohydrate-binding. The second putative agarase, Aga4017B, shared some sequence homology with GH50 β -agarases and was predicted to contain a C-terminal agarase_CBM domain, a characteristic of GH50 β -agarases. Yet, Aga4017B only shared 51.81 % identity to its nearest characterised relative, a GH50 β -agarase from *P. fuliginea*. It is, therefore, likely that either Aga4017B is a GH50 β -agarase with an unusual structure or a related protein that binds to the same substrate as GH50 β -agarase. Purification and further analysis of these β -agarases, and their end products, is necessary to fully characterise them and determine their potential downstream applications. Cold-active agarases are of great biotechnological interest for their application in various industrial processes and production of bioactive oligosaccharides from agar. This study presents the first to report on the preliminary characterisation and cloning of an agarase derived from a deep-sea fish microbiome isolate.

5.6. References

- BARROCA, M., SANTOS, G., GERDAY, C. & COLLINS, T. 2017. Biotechnological Aspects of Cold-Active Enzymes. In: MARGESIN, R. (ed.) *Psychrophiles: From Biodiversity to Biotechnology*. Cham: Springer International Publishing. 461–475.
- CHEN, X., FU, X., HUANG, L., XU, J. & GAO, X. 2021. Agar oligosaccharides: A review of preparation, structures, bioactivities and application. *Carbohydrate Polymers*, 265, 118076.
- CHEN, X.-L., HOU, Y.-P., JIN, M., ZENG, R.-Y. & LIN, H.-T. 2016. Expression and characterization of a novel thermostable and ph-stable β -agarase from deep-sea bacterium *Flammeovirga* Sp. OC4. *Journal of Agricultural and Food Chemistry*, 64, 7251-7258.
- CHI, W. J., CHANG, Y. K. & HONG, S. K. 2012. Agar degradation by microorganisms and agar-degrading enzymes. *Applied Microbiology and Biotechnology*, 94, 917-30.
- CUI, X., JIANG, Y., CHANG, L., MENG, L., YU, J., WANG, C. & JIANG, X. 2018. Heterologous expression of an agarase gene in *Bacillus subtilis*, and characterization of the agarase. *International Journal of Biological Macromolecules*, 120, 657-664.
- DAVIES, G. & HENRISSAT, B. 1995. Structures and mechanisms of glycosyl hydrolases. *Structure*, 3, 853-859.
- EDGAR, R. C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32, 1792-1797.
- EKBORG, N. A., TAYLOR, L. E., LONGMIRE, A. G., HENRISSAT, B., WEINER, R. M. & HUTCHESON, S. W. 2006. Genomic and proteomic analyses of the agarolytic system expressed by *Saccharophagus degradans* 2-40. *Applied and Environmental Microbiology*, 72, 3396-405.
- FARGHALI, M., MOHAMED, I. M. A., OSMAN, A. I. & ROONEY, D. W. 2023. Seaweed for climate mitigation, wastewater treatment, bioenergy, bioplastic, biochar, food, pharmaceuticals, and cosmetics: a review. *Environmental Chemistry Letters*, 21, 97-152.
- FLAMENT, D., BARBEYRON, T., JAM, M., POTIN, P., CZIZEK, M., KLOAREG, B. & MICHEL, G. 2007. Alpha-agarases define a new family of glycoside hydrolases, distinct from beta-agarase families. *Applied and Environmental Microbiology*, 73, 4691-4.
- FLEURENCE, J., MASSIANI, L., GUYADER, O. & MABEAU, S. 1995. Use of enzymatic cell wall degradation for improvement of protein extraction from *Chondrus crispus*, *Gracilaria verrucosa* and *Palmaria palmata*. *Journal of Applied Phycology*, 7, 393-397.

- FU, X. T. & KIM, S. M. 2010. Agarase: Review of major sources, categories, purification method, enzyme characteristics and applications. *Marine Drugs*, 8, 200-218.
- FUJIMOTO, Z. 2013. Structure and function of carbohydrate-binding module families 13 and 42 of glycoside hydrolases, comprising a β -trefoil fold. *Bioscience, Biotechnology, and Biochemistry*, 77, 1363-71.
- GOBET, A., BARBEYRON, T., MATARD-MANN, M., MAGDELENAT, G., VALLENET, D., DUCHAUD, E. & MICHEL, G. 2018. Evolutionary evidence of algal polysaccharide degradation acquisition by *Pseudoalteromonas carrageenovora* 9T to Adapt to Macroalgal Niches. *Frontiers in Microbiology*, 9.
- HONG, S. J., LEE, J.-H., KIM, E. J., YANG, H. J., PARK, J.-S. & HONG, S.-K. 2017. Anti-obesity and anti-diabetic effect of neoagarooligosaccharides on high-fat diet-induced obesity in mice. *Marine Drugs*, 15, 90.
- HU, B., GONG, Q., WANG, Y., MA, Y., LI, J. & YU, W. 2006. Prebiotic effects of neoagaro-oligosaccharides prepared by enzymatic hydrolysis of agarose. *Anaerobe*, 12, 260-266.
- JONNADULA, R., IMRAN, M., VASHIST, P. & GHADI, S. C. 2020. Production of agar-derived antioxidants and single cell detritus from *Gracilaria corticata* using agarase from *Microbulbifer* sp. CMC-5. *Proceedings of the National Academy of Sciences, India Section B: Biological Sciences*, 90, 73-78.
- JUMPER, J., EVANS, R., PRITZEL, A., GREEN, T., FIGURNOV, M., RONNEBERGER, O., TUNYASUVUNAKOOL, K., BATES, R., ŽÍDEK, A., POTAPENKO, A., BRIDGLAND, A., MEYER, C., KOHL, S. A. A., BALLARD, A. J., COWIE, A., ROMERA-PAREDES, B., NIKOLOV, S., JAIN, R., ADLER, J., BACK, T., PETERSEN, S., REIMAN, D., CLANCY, E., ZIELINSKI, M., STEINEGGER, M., PACHOLSKA, M., BERGHAMMER, T., BODENSTEIN, S., SILVER, D., VINYALS, O., SENIOR, A. W., KAVUKCUOGLU, K., KOHLI, P. & HASSABIS, D. 2021. Highly accurate protein structure prediction with AlphaFold. *Nature*, 596, 583-589.
- KANG, J. Y., SONG, H.-Y. & KIM, J.-M. 2023. Agarolytic pathway in the newly isolated *Aquimarina* sp. bacterial strain ERC-38 and characterization of a putative β -agarase. *Marine Biotechnology*, 25, 314-327.
- KIMURA, M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *Journal of Molecular Evolution*, 16, 111-120.
- KUMAR, S., STECHER, G., LI, M., KNYAZ, C. & TAMURA, K. 2018. MEGA X: Molecular evolutionary genetics analysis across computing platforms. *Molecular Biology and Evolution*, 35, 1547-1549.
- LI, R.-K., CHEN, Z., YING, X.-J., NG, T. B. & YE, X.-Y. 2018. A novel GH16 beta-agarase isolated from a marine bacterium, *Microbulbifer* sp. BN3 and its

- characterization and high-level expression in *Pichia pastoris*. *International Journal of Biological Macromolecules*, 119, 1164-1170.
- LI, T., CUI, N., FAN, J., LIU, X., GUAN, J., ZHANG, W., LI, Y., GAO, Q., LIU, Y., XU, Y., WANG, L., GUO, C. & WANG, Y. 2022. Engineering β -agarase Aga0917 from *Pseudoalteromonas fuliginea* YTW1-15-1 to improve the thermostability and enzyme activity. *ACS Sustainable Chemistry & Engineering*, 10, 15437-15449.
- LITTLECHILD, J. 2015. Enzymes from extreme environments and their industrial applications. *Frontiers in Bioengineering and Biotechnology*, 3.
- MANGIAGALLI, M. & LOTTI, M. 2021. Cold-active β -galactosidases: insight into cold adaptation mechanisms and biotechnological exploitation. *Marine Drugs*, 19, 43.
- MESBAH, N. M. 2022. Industrial biotechnology based on enzymes from extreme environments. *Front Bioeng Biotechnol*, 10, 870083.
- MICHEL, G., NYVAL-COLLEN, P., BARBEYRON, T., CZIZEK, M. & HELBERT, W. 2006. Bioconversion of red seaweed galactans: a focus on bacterial agarases and carrageenases. *Applied Microbiology and Biotechnology*, 71, 23-33.
- MORRICE, L. M., MCLEAN, M. W., WILLIAMSON, F. B. & LONG, W. F. 1983. β -Agarases I and II from *Pseudomonas atlantica*. Purifications and some properties. *European Journal of Biochemistry*, 135, 553-558.
- OHTA, Y., NOGI, Y., MIYAZAKI, M., LI, Z., HATADA, Y., ITO, S. & HORIKOSHI, K. 2004. Enzymatic properties and nucleotide and amino acid sequences of a thermostable beta-agarase from the novel marine isolate, JAMB-A94. *Bioscience, Biotechnology, and Biochemistry*, 68, 1073-81.
- PARK, S. H., LEE, C.-R. & HONG, S.-K. 2020. Implications of agar and agarase in industrial applications of sustainable marine biomass. *Applied Microbiology and Biotechnology*, 104, 2815-2832.
- PARRILLI, E., TEDESCO, P., FONDI, M., TUTINO, M. L., LO GIUDICE, A., DE PASCALE, D. & FANI, R. 2021. The art of adapting to extreme environments: The model system *Pseudoalteromonas*. *Physics of Life Reviews*, 36, 137-161.
- PLUVINAGE, B., HEHEMANN, J.-H. & BORASTON, A. B. 2013. Substrate recognition and hydrolysis by a family 50 exo- β -agarase, Aga50D, from the marine bacterium *Saccharophagus degradans*. *The Journal of biological chemistry*, 288, 28078-28088.
- PLUVINAGE, B., ROBB, C. S., JEFFRIES, R. & BORASTON, A. B. 2020. The structure of PfGH50B, an agarase from the marine bacterium *Pseudoalteromonas fuliginea* PS47. *Acta Crystallographica. Section F, Structural Biology Communications*, 76, 422-427.
- SCHULTZ-JOHANSEN, M., BECH, P. K., HENNESSY, R. C., GLARING, M. A., BARBEYRON, T., CZIZEK, M. & STOUGAARD, P. 2018. A novel enzyme portfolio for red algal polysaccharide degradation in the marine bacterium *Paraglaciecola hydrolytica* S66T encoded in a sizeable polysaccharide

- utilization locus. *Frontiers in microbiology*, 9, 839. doi: 10.3389/fmicb.2018.00839. PMID: 29774012; PMCID: PMC5943477.
- SEEMANN, T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*, 30, 2068-9.
- SEHNAL, D., BITTRICH, S., DESHPANDE, M., SVOBODOVÁ, R., BERKA, K., BAZGIER, V., VELANKAR, S., BURLEY, S. K., KOČA, J. & ROSE, A. S. 2021. Mol* Viewer: modern web app for 3D visualization and analysis of large biomolecular structures. *Nucleic Acids Research*, 49, W431-W437.
- TAKAGI, E., HATADA, Y., AKITA, M., OHTA, Y., YOKOI, G., MIYAZAKI, T., NISHIKAWA, A. & TONOZUKA, T. 2015. Crystal structure of the catalytic domain of a GH16 β -agarase from a deep-sea bacterium, *Microbulbifer thermotolerans* JAMB-A94. *Bioscience, Biotechnology, and Biochemistry*, 79, 625-632.
- TEUFEL, F., ALMAGRO ARMENTEROS, J. J., JOHANSEN, A. R., GÍSLASON, M. H., PIHL, S. I., TSIRIGOS, K. D., WINTHER, O., BRUNAK, S., VON HEIJNE, G. & NIELSEN, H. 2022. SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nature Biotechnology*, 40, 1023-1025.
- TUNYASUVUNAKOOL, K., ADLER, J., WU, Z., GREEN, T., ZIELINSKI, M., ŽÍDEK, A., BRIDGLAND, A., COWIE, A., MEYER, C., LAYDON, A., VELANKAR, S., KLEYWEGT, G. J., BATEMAN, A., EVANS, R., PRITZEL, A., FIGURNOV, M., RONNEBERGER, O., BATES, R., KOHL, S. A. A., POTAPENKO, A., BALLARD, A. J., ROMERA-PAREDES, B., NIKOLOV, S., JAIN, R., CLANCY, E., REIMAN, D., PETERSEN, S., SENIOR, A. W., KAVUKCUOGLU, K., BIRNEY, E., KOHLI, P., JUMPER, J. & HASSABIS, D. 2021. Highly accurate protein structure prediction for the human proteome. *Nature*, 596, 590-596.
- UNIACKE-LOWE, S., COLLINS, F. W. J., HILL, C. & ROSS, R. P. 2023. Bioactivity screening and genomic analysis reveals deep-sea fish microbiome isolates as sources of novel antimicrobials. *Marine Drugs*, 21, 444.
- WANG, W., LIU, P., HAO, C., WU, L., WAN, W. & MAO, X. 2017. Neoagaro-oligosaccharide monomers inhibit inflammation in LPS-stimulated macrophages through suppression of MAPK and NF- κ B pathways. *Scientific Reports*, 7, 44252.
- WATERHOUSE, A. M., PROCTER, J. B., MARTIN, D. M. A., CLAMP, M. & BARTON, G. J. 2009. Jalview Version 2—a multiple sequence alignment editor and analysis workbench. *Bioinformatics*, 25, 1189-1191.
- XU, X.-Q., SU, B.-M., XIE, J.-S., LI, R.-K., YANG, J., LIN, J. & YE, X.-Y. 2018. Preparation of bioactive neoagaroligosaccharides through hydrolysis of *Gracilaria lemaneiformis* agar: A comparative study. *Food Chemistry*, 240, 330-337.
- YANG, M., DERBYSHIRE, M. K., YAMASHITA, R. A. & MARCHLER-BAUER, A. 2020. NCBI's conserved domain database and tools for protein domain analysis. *Current Protocols in Bioinformatics*, 69, e90.

- YAPHE, W. 1957. The use of agarase from *Pseudomonas atlantica* in the identification of agar in marine algae (*Rhodophyceae*). *Canadian Journal of Microbiology*, 3, 987-993.
- YOON, S.-H., HA, S.-M., KWON, S., LIM, J., KIM, Y., SEO, H. & CHUN, J. 2017. Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *International Journal of Systematic and Evolutionary Microbiology*, 67, 1613-1617.
- YUN, E. J., LEE, S., KIM, J. H., KIM, B. B., KIM, H. T., LEE, S. H., PELTON, J. G., KANG, N. J., CHOI, I.-G. & KIM, K. H. 2013. Enzymatic production of 3,6-anhydro-l-galactose from agarose and its purification and in vitro skin whitening and anti-inflammatory activities. *Applied Microbiology and Biotechnology*, 97, 2961-2970.

5.7. Chapter 5 Figures and Tables

Table 4: Primers used in this study for amplification of agarase genes and corresponding gene targets. T_m = melting temperature.

Primer ID	Primer Sequence (5' – 3')	T _m (°C)	Gene target	Product length (bp)
23B_B_agaF	GACAACATCAATTTAAAGAGGC	59.2	<i>aga4017A</i> (forward strand)	1393
23B_B_agaR	CACTCAATGTAAAACTTTGC	56.5	<i>aga4017A</i> (reverse strand)	
23B_H_agaF	TAGGCTTGCTATGAAATACACA	58.9	<i>aga4017B</i> (forward strand)	2351
23B_H_agaR	TACCCAAGCCTGAATATACTAG	58.0	<i>aga4017B</i> (reverse strand)	

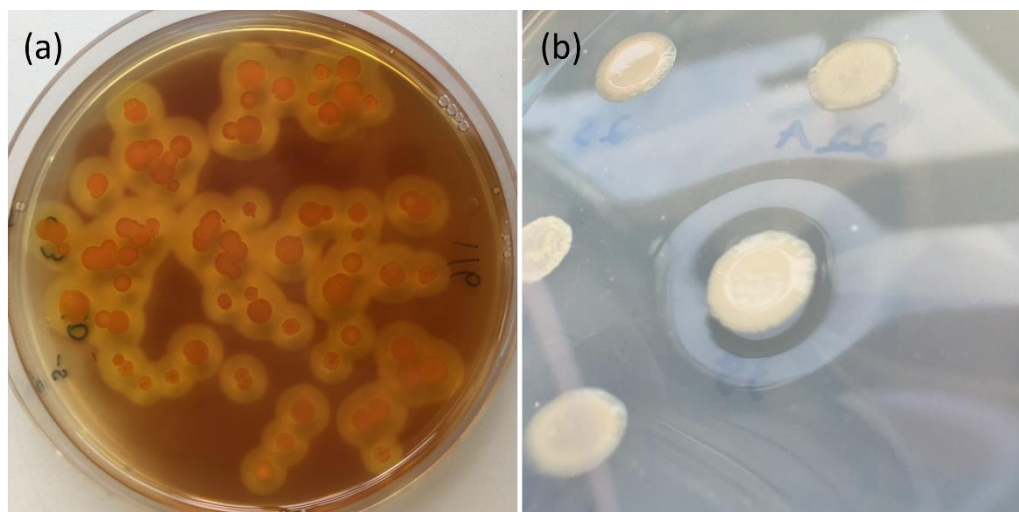


Figure 4: Colonies of the deep-sea fish isolates on marine agar after staining with Lugol's iodine. Clear zones around the colonies indicate where agar degradation has occurred (a). A single colony of *Pseudoalteromonas* sp. APC 4017 exhibiting agar "pitting" on marine agar (b).

Table 5: Conserved domains predicted in Aga4017A.

Query	From	To	E-Value	Bitscore	Accession	Short name
Aga4017A	18	294	6.21E-98	293.873	cd02178	GH16_beta_agarase
	308	441	4.06E-20	85.6611	PF00652	Ricin_B_lectin

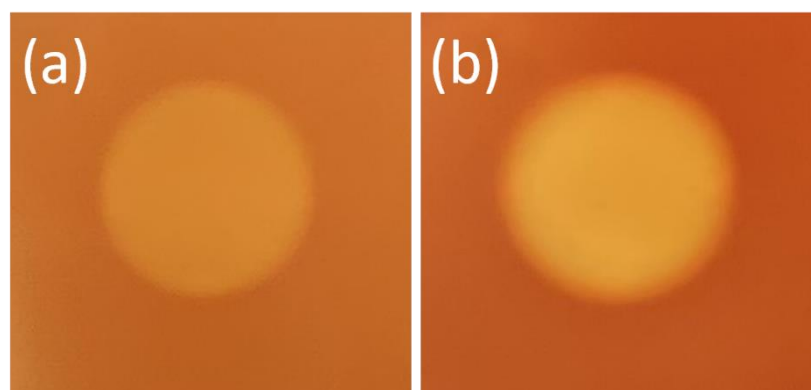


Figure 5: Crude agarase prepared from APC 4017 was spotted on 2 % agar medium and incubated at (a) 5 °C and (b) 25 °C. Staining with Lugol's iodine resulted in clear zones forming in the areas where the agar had been degraded.

Table 6: Conserved features and corresponding residue coordinates predicted in β -agarase Aga4017A.

Query	Type	Title	coordinates
Aga4017 A	specific	active site	N66, W68, W134, D140, E143, D145, E148, F169, R171, E260, H262
	specific	calcium binding site	D21, N48, D288
	specific	catalytic residues	E143, D145, E148
	specific	Q-X-W motif	Q344, Q345, F346; Q392, R393, F394; Q439, R440, F441
	specific	putative sugar binding sites	D320, Q333, W335, N343, Q344, Q368, Q381, W383, S391, Q392, D416, V429, W431, Q439

Table 7: Comparison of Aga4017A to most similar sequences and selected characterised hits from the NCBI nr database queried using BLASTp.

Query	Query len (AA)	Hit description [strain]	% Cov.	% Id	len (AA)	accession	status
Aga4017A	445	Ricin domain-containing protein [<i>Pseudoalteromonas</i> sp. A601]	100	96.63	445	WP_087533949.1	uncharacterised
		Ricin domain-containing protein [<i>Pseudoalteromonas telluritireducens</i>]	100	93.71	445	WP_064385033.1	uncharacterised
		Ricin domain-containing protein [<i>Pseudoalteromonas</i> sp. Xi13]	100	93.48	445	WP_126007335.1	uncharacterised
		beta-agarase [<i>Pseudoalteromonas</i> sp. Q30F]	100	93.48	445	AXK59832.1	characterised
		beta-agarase [<i>Pseudoalteromonas</i> sp. H9]	100	87.42	445	AIW52355.1	characterised
		extracellular agarase precursor [<i>Janthinobacterium</i> sp. SY12]	100	87.19	445	ABO93616.2	characterised
		beta-agarase [<i>Pseudoalteromonas</i> sp. JT-6]	100	86.97	445	ACF40223.1	characterised

len = length; cov = coverage; Id = identity

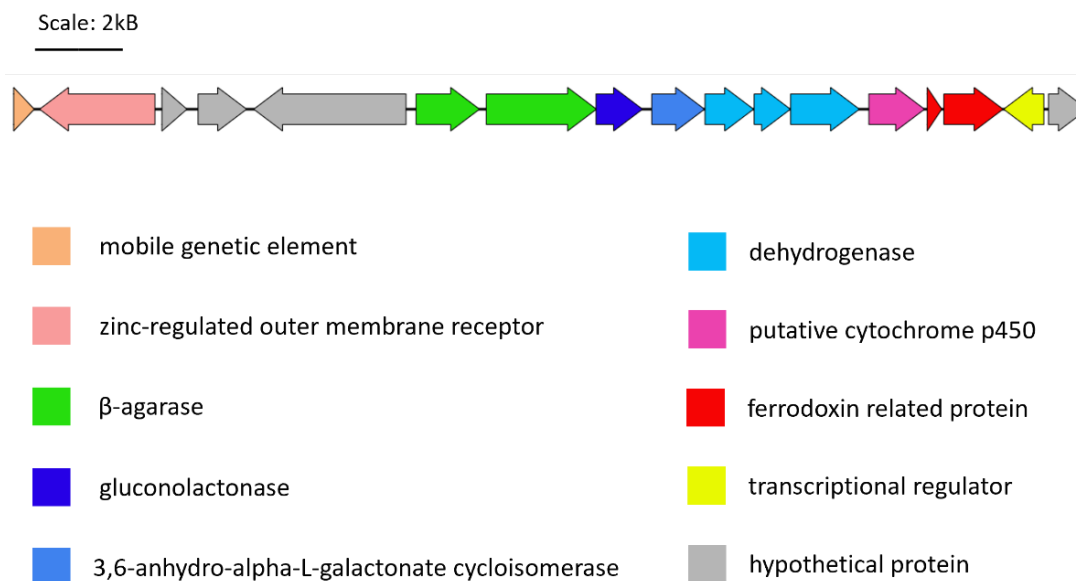


Figure 6: The two putative β-agarase genes, *aga4017A* and *aga4017B*, were located within a large gene cluster encoding enzymes likely involved metabolism of agarose and agarose-derivates. Figure was created using Gene Graphics (genegraphics.net).

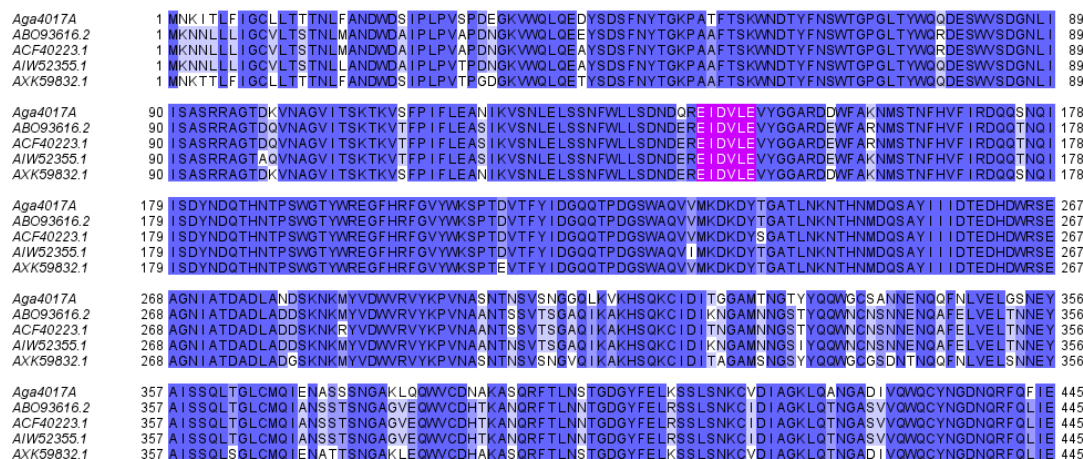


Figure 7: Multiple sequence alignment of Aga4017A and similar characterised agarases from the NCBI nr database: *Pseudoalteromonas* sp. Q30F, AXK59832.1; *Pseudoalteromonas* sp. H9, AIW52355.1; *Janthinobacterium* sp. SY12, ABO93616.2; *Pseudoalteromonas* sp. JT-6, ACF40223.1. Residues are coloured according to % identity (>50 %), with the colour intensity increasing as the value approaches 100 %. A highly conserved catalytic sequence pattern (EXDXXE), characteristic of GH16 family β -agarases, is highlighted in magenta.

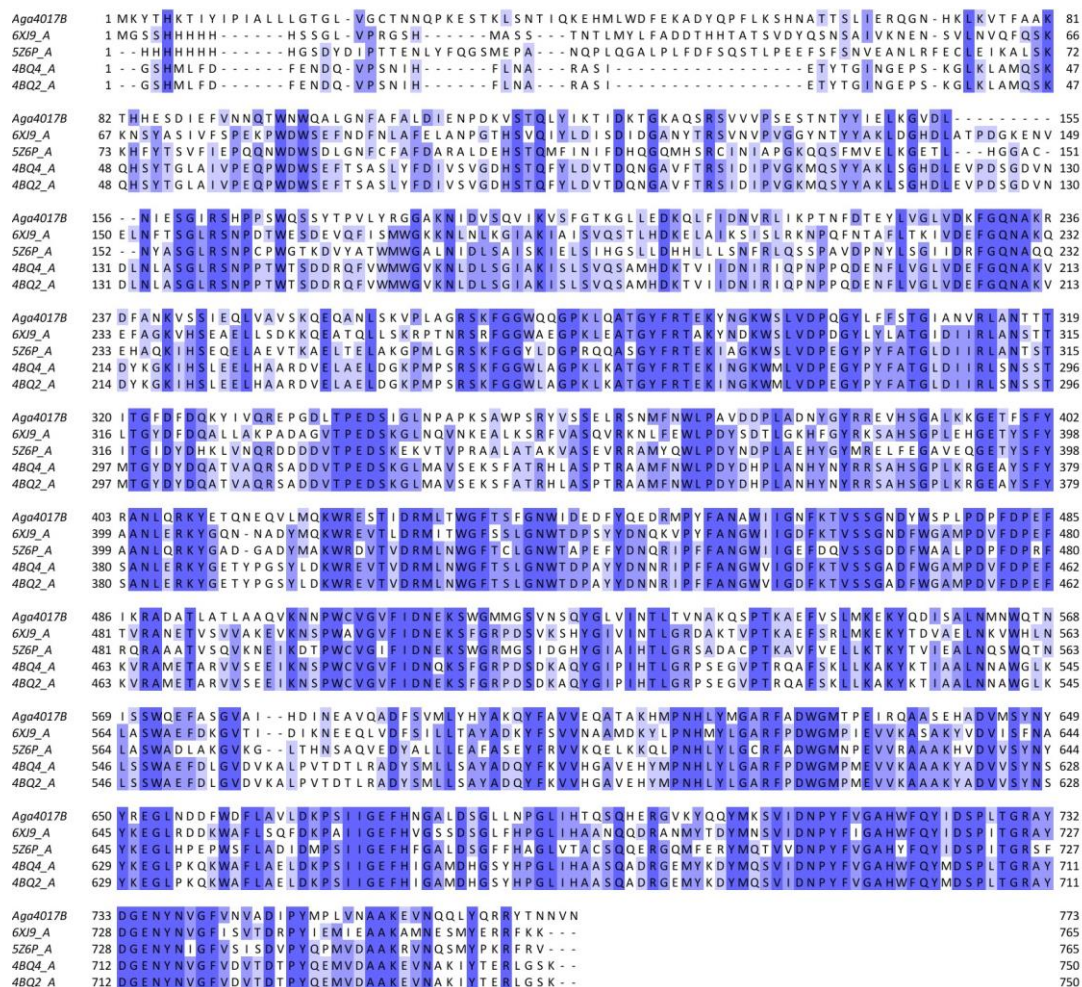


Figure 8: Multiple sequence alignment of Aga4017B and hit sequence from the NCBI PDB database: *Pseudalteromonas fuliginea* (6XJ9), *Agarivorans gilvus* (5Z6P) and *Saccharophagus degradans* 2-40 (4BQ2 and 4BQ4). Residues are coloured according to % identity (>50 %), with the colour intensity increasing as the value approaches 100 %.

Table 8: The sole predicted conserved domain in Aga4017B was Agarase_CBM (Pfam PF17992), a carbohydrate binding module -like domain frequently found in exo β -agarases.

Query	From	To	E-Value	Bitscore	Accession	Short name	Superfamily
Aga4017B	46	223	2.34E-64	212.224	PF17992	Agarase_CBM	cl40253

Table 9: Overview of molecular characteristics of putative β -agarases, Aga4017A and Aga4017B, from *Pseudoalteromonas* sp. APC 4017.

Name	Length (AA)	Mr (kDa)	pI	Signal peptide type	Signal peptide length (AA)	predicted GH family
Aga4017A	445	50.21	4.75	Sec/SPI	19	GH16
Aga4017B	773	88.02	6.22	Sec/SPII lipopeptide	22	GH50

Mr = molecular mass; pI = theoretical isoelectric point; GH = glycoside hydrolase

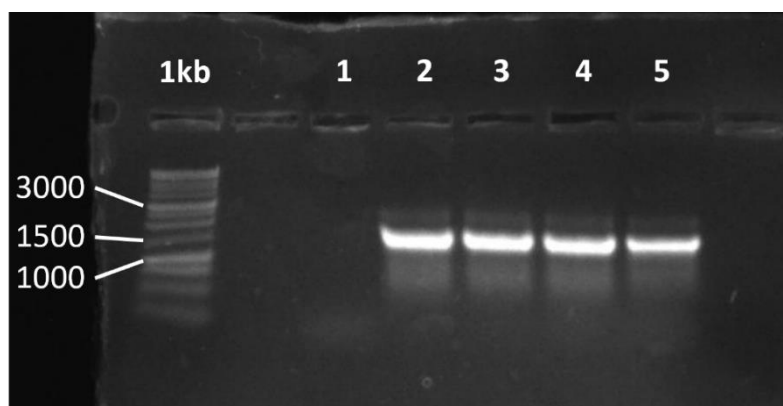


Figure 9: Gel electrophoresis following colony PCR on transformed *E. coli* TOP10 using *aga4017A* primers. Lane 1 = negative control (no DNA); 2-4 = randomly selected transformed *E. coli* TOP10 pCR2.1-TOPO_aga4017a colonies exhibiting agar-pitting; 5 = positive control (APC 4017 DNA as template). A 1kb molecular ladder was used. The expected PCR product size was 1393 kb.



Figure 10: Map of cloning vector pCR2.1-TOPO_aga4017A encoding β -agarase (*aga4017A*) from *Pseudoalteromonas* sp. APC 4017. Vector map was created in SnapGene.

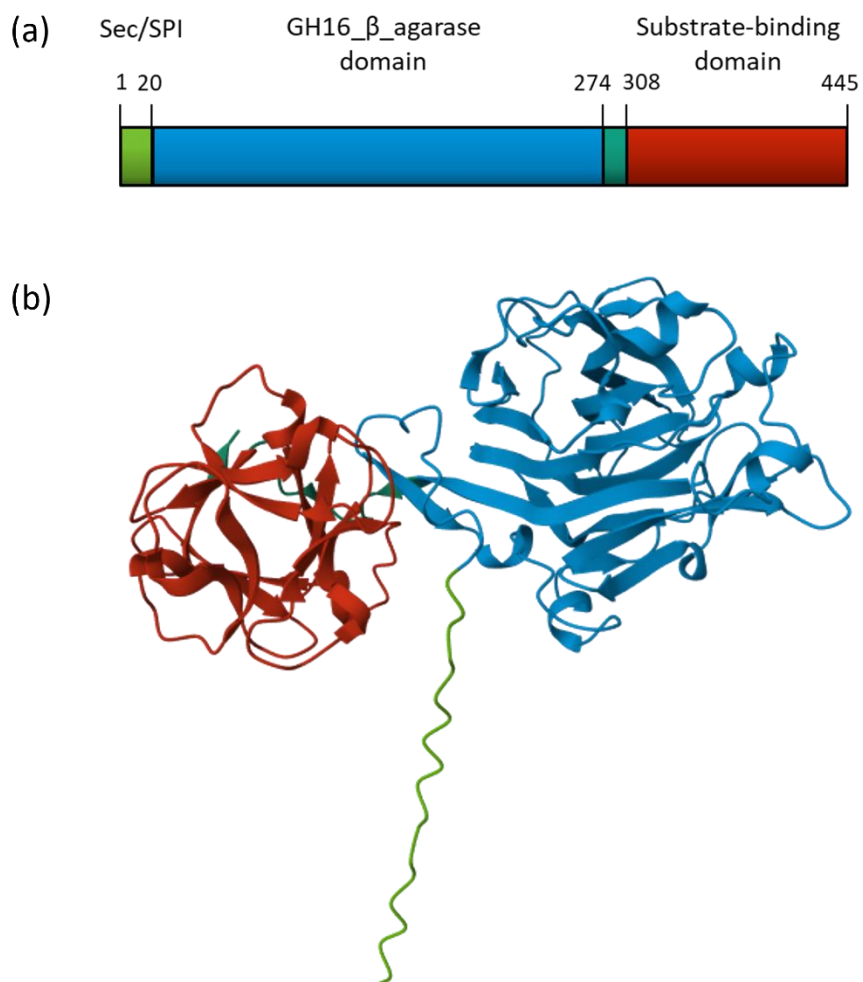


Figure 11: (a) Illustration of the predicted conserved domains and corresponding amino acid positions of the Aga4017A agarase sequence. Shown are the N terminal Sec/SPI signal peptide (light green), a GH16 β -agarase domain (blue), the C terminal Ricin B lectin like substrate-binding domain (Red) and a linker region (dark green). (b) Predicted 3D structure of Aga4017A, created using AlphaFold2 and visualised using Mol* 3D viewer (Protein Data Bank). The sequence is coloured according to predicted domain.



Figure 12: Predicted 3D structure of beta agarase Aga4017A, including the signal sequence, created using Alphafold2 and visualised using Mol* 3D viewer (Protein Data Bank). The sequence is coloured according to predicted secondary structure: Pink = α -helix; purple = 3_{10} -helix; yellow = β -strand; white = coil; green = turn/bend.

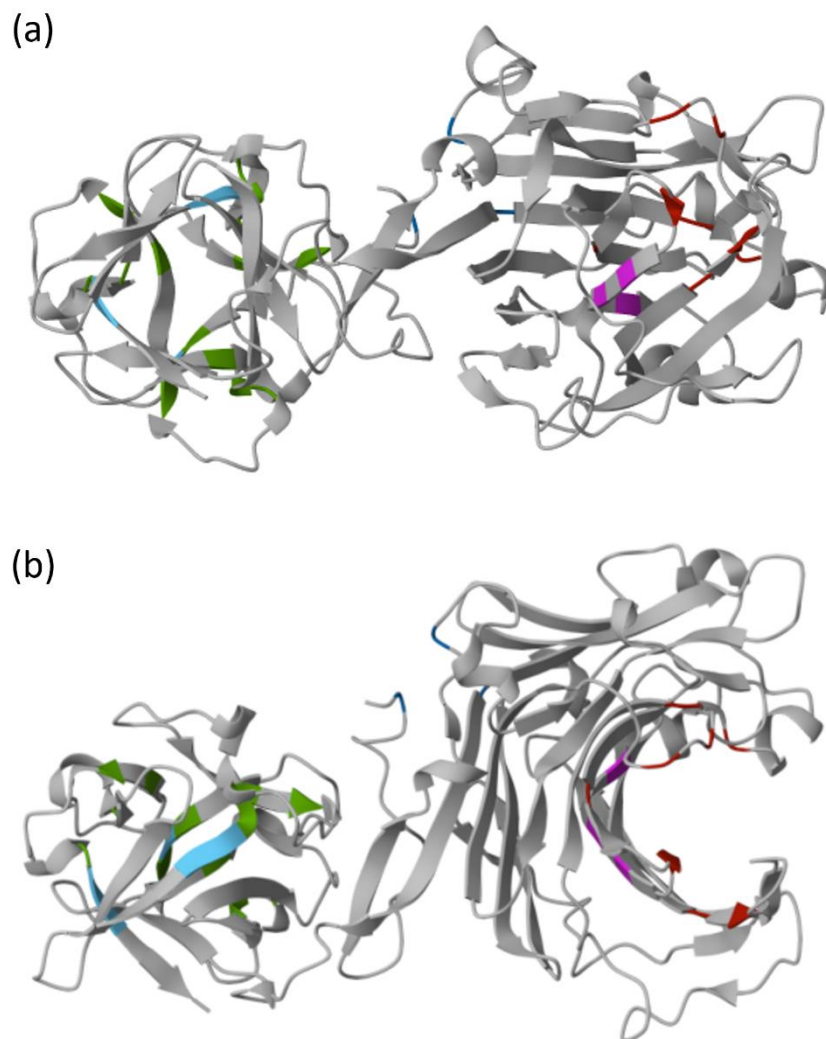


Figure 13: Predicted 3D structure of beta agarase Aga4017A (excluding signal sequence). Default view (a); top-down view (b). Predicted conserved residues are highlighted: Green = putative sugar-binding sites; light blue = Q-X-W- motif; red = active site; magenta = catalytic residues; marine blue = calcium-binding residues. Predicted models and conserved sites were visualised in Mol* 3D View (PDB).

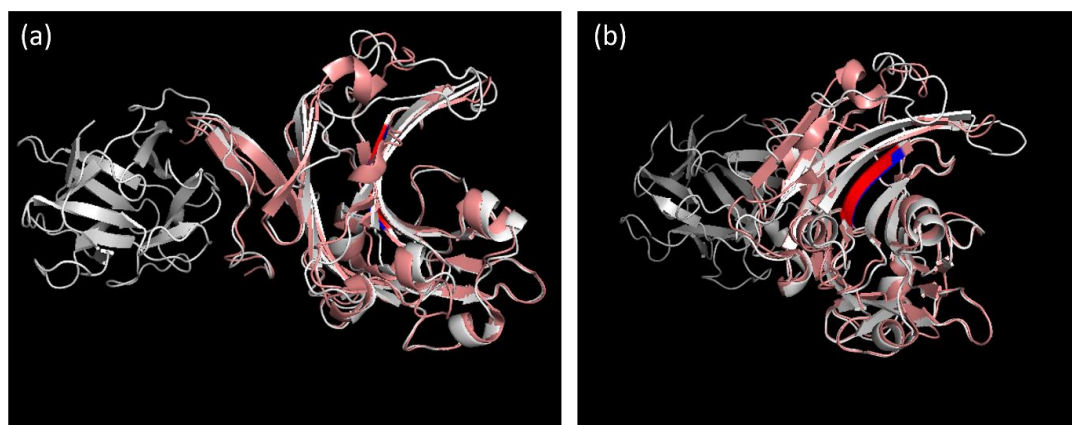


Figure 14: Superposition of the predicted 3D model of Aga4017A (white) with the crystal structure of the GH16 catalytic domain of β -agarase MtAga from *Microbulbifer thermotolerans* JAMB-A94 (light pink). The regions shared 42.07% sequence identity yet are structurally very similar and both contain the distinctive concave topology in this region (a). Significantly, the highly conserved EXDXXE catalytic residue sequence present in both proteins (red and blue, respectively) overlap perfectly within this concave region (b). Figure was created using PyMOL.

5.8. Chapter 5 Supplementary Material

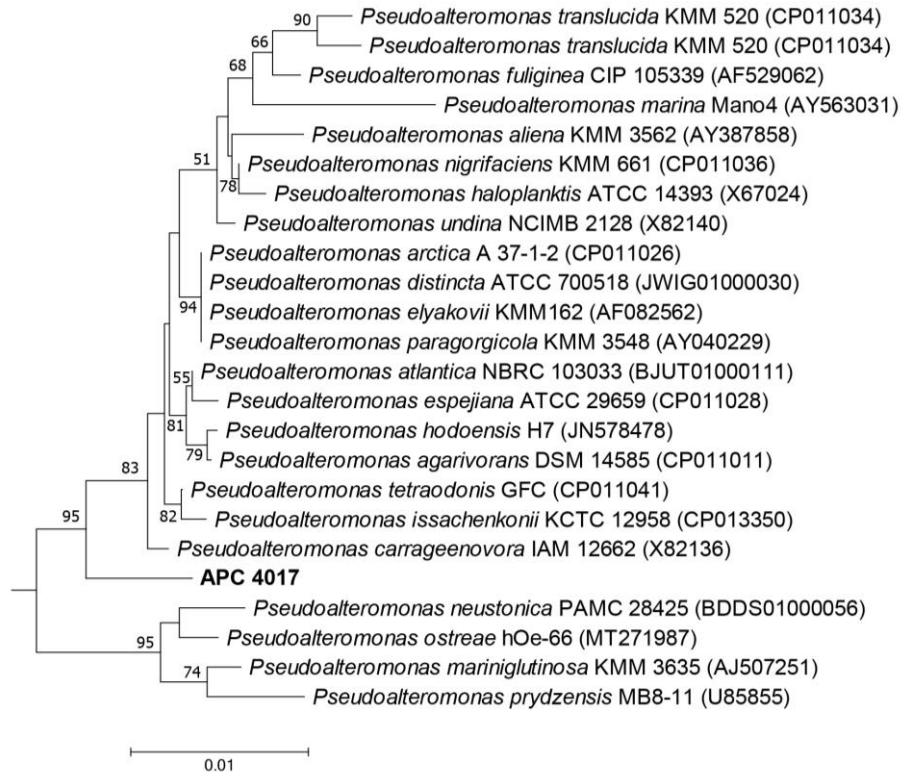


Figure S1: Subtree of rooted neighbour-joining phylogenetic tree based on the alignment of 16S rRNA gene sequences from strain APC 4017 and hits from the EzBioCloud taxon database. APC 4017 was most closely related to members of the genus *Pseudoalteromonas*.

Table S1: Full output of NCBI CDD predicted conserved domains in APC 4017 agarase sequences.

Query	Hit type	From	To	E-Value	Bitscore	Accession	Short name	Superfamily
Aga4017A	non-specific	44	293	2.35E-22	95.00	cd08023	GH16_laminarinase_like	cl22861
	non-specific	45	293	1.18E-18	84.02	cd00413	Glyco_hydrolase_16	cl22861
	non-specific	343	431	7.59E-18	78.19	pfam14200	RicinB_lectin_2	cl23784
	non-specific	301	383	5.77E-13	64.32	pfam14200	RicinB_lectin_2	cl23784
	non-specific	306	441	1.66E-11	64.03	NF035930	lectin_2	cl41551
	non-specific	356	445	2.37E-11	61.01	pfam00652	Ricin_B_lectin	cl40832
	non-specific	303	394	5.76E-09	54.07	pfam00652	Ricin_B_lectin	cl40832
	non-specific	390	445	1.52E-08	51.61	pfam14200	RicinB_lectin_2	cl23784
	non-specific	358	445	7.09E-08	50.97	cd00161	RICIN	cl23784
	non-specific	76	220	1.65E-06	47.97	pfam00722	Glyco_hydro_16	cl22861
	non-specific	182	294	0.000644	41.46	cd08024	GH16_CCF	cl22861
	non-specific	403	443	0.008404	36.36	pfam00652	Ricin_B_lectin	cl40832
	specific	18	294	6.21E-98	293.87	cd02178	GH16_beta_agarase	cl22861
	specific	308	441	4.06E-20	85.66	pfam00652	Ricin_B_lectin	cl40832
	specific	315	443	3.24E-13	65.99	cd00161	RICIN	cl23784
	specific	312	444	6.00E-13	65.23	smart00458	RICIN	cl40779
	superfamily	18	294	6.21E-98	293.87	cl22861	LamG superfamily	-
	superfamily	308	441	4.06E-20	85.66	cl40832	Ricin_B_lectin superfamily	-
	superfamily	343	431	7.59E-18	78.19	cl23784	RICIN superfamily	-
	superfamily	312	444	6.00E-13	65.23	cl40779	RICIN superfamily	-
	superfamily	306	441	1.66E-11	64.03	cl41551	lectin_2 superfamily	-
Aga4017B	non-specific	46	223	2.34E-64	212.22	pfam17992	Agarase_CBM	cl40253
	superfamily	46	223	2.34E-64	212.22	cl40253	Agarase_CBM superfamily	-

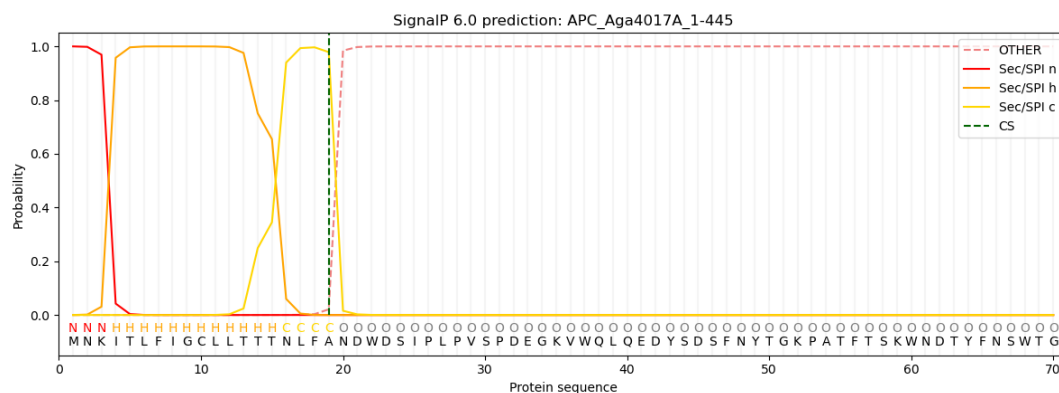


Figure S2: Output of SignalP analysis of GH16- family β -agarase Aga4017A. The sequence was predicted to contain a Sec signal peptidase II signal sequence (Sec/SPII) and a cut site (CS) between residues 19 and 20.

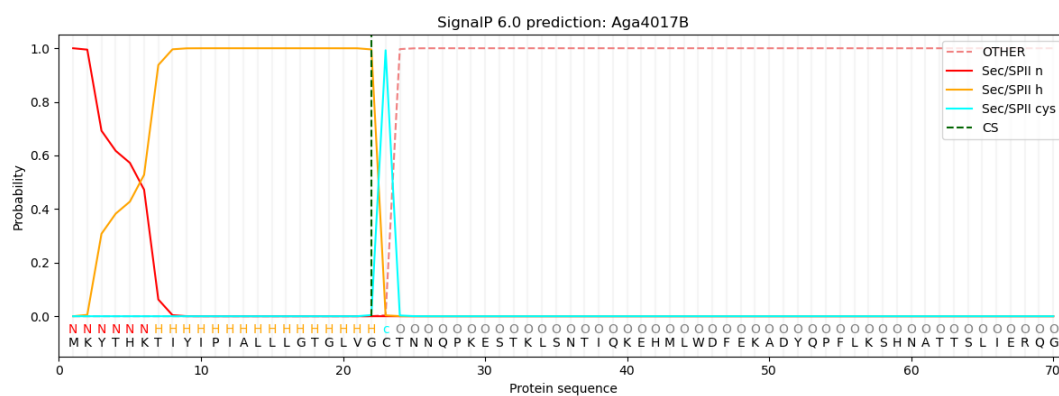


Figure S3: Output of SignalP analysis of GH50-family β -agarase Aga4017B. The sequence was predicted to contain a Sec signal peptidase II signal sequence (Sec/SPII) and a cut site (CS) between residues 18 and 19.

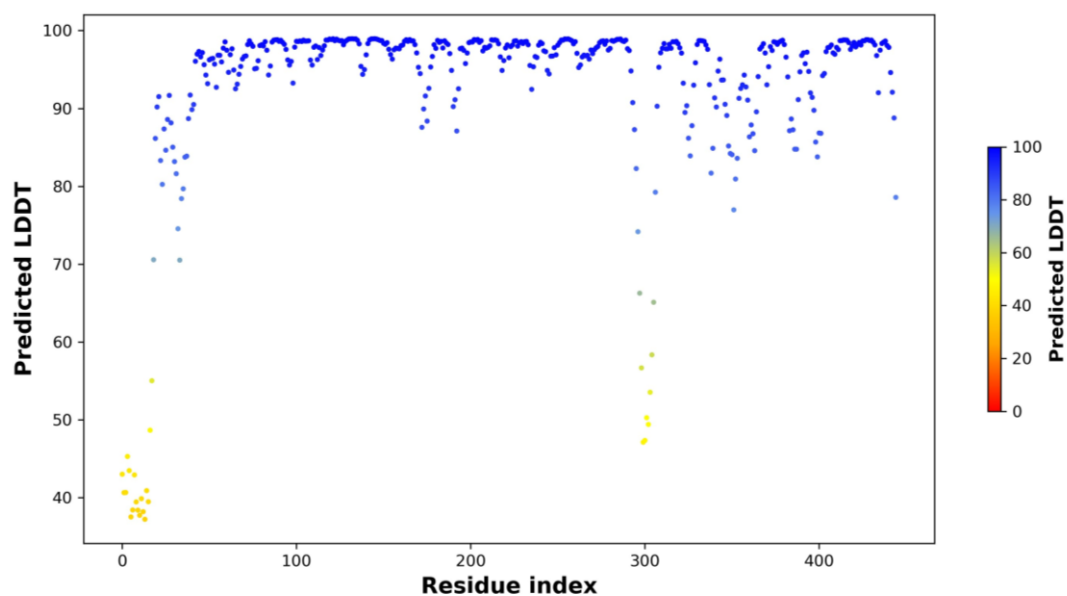


Figure S4: Plot of estimated per-residue confidence of Aga4017A 3D structure generated by AlphaFold2. Regions with a predicted local distance difference test (pLDDT) score >90 (dark blue) are expected to be highly accurate.

Chapter 6

General Discussion

The purpose of this thesis was to investigate deep-sea fish microbiomes as a source of antimicrobial metabolites, with a particular focus on identifying novel bacteriocins and characterising bacteriocinogenic strains with beneficial applications.

The Earth's marine environment is its largest ecological habitat, covering over 70 % of the planet's surface and extending to great depths of over 10,000m. The deep sea (waters below 200m) alone constitutes over 95 % of the ocean's total volume. Yet, it is estimated that only a fraction of the marine environment has been explored, with relatively little known about its microbial inhabitants compared to those on land (Danovaro et al., 2017, Greenaway et al., 2021). It is arguably one of the last great frontiers of scientific exploration on Earth.

One of the greatest threats to human health is the increasing prevalence of infections which are untreatable due to the emergence of antibiotic resistant pathogens such as methicillin-resistant *Staphylococcus aureus* (MRSA), vancomycin-resistant *Enterococcus* (VRE), *Clostridium difficile* and carbapenem-resistant Enterobacteriaceae (Frieri et al., 2017). If left unaddressed, antimicrobial resistance could force 24 million people into extreme poverty by 2030 and cause 10 million deaths per year globally by 2050 (IACG, 2019) (Resistance). It is therefore now, more important than ever, to identify sources of novel antimicrobials to combat the rise in antimicrobial resistance.

The open ocean waters can be divided into layers or strata, each with unique defining characteristics that have a considerable influence on the diversity of its inhabitants (National Oceanic and Atmospheric Administration). The deep sea begins from approximately 200m deep and is made up of the mesopelagic (200–1000m), bathypelagic (1000–4000m), abyssal (4000–6000m) and hadal (below 6000m) zones. The deep sea is characterised by extreme environmental conditions such as high hydrostatic pressure, no sunlight and relatively low temperatures and oxygen availability. This has a significant impact on the evolution of marine organisms (Somero, 1992). Deep-sea fish are highly adapted to their environment, in terms of having specialised physical characteristics and also in their production of unusual proteins and metabolites, such as piezolytes. In such a harsh and metabolically demanding environment, it is reasonable to hypothesise that marine microbial communities too possess unique characteristics in terms of their diversity and the

bioactive compounds they produce. With regards to bacteriocins specifically, we hypothesised that marine microbiomes may possess unusual bacteriocins as defence mechanisms for competition and their survival. The human gut microbiome, for example, possesses an arsenal of antimicrobials such as bacteriocins, fatty acids and other metabolites which undoubtedly contribute to the ability to compete and persist in a complex microbial community and in this respect can be considered an innate defence system to protect from infectious organisms. We, therefore, sought to explore the deep-sea fish microbiome as a source of (novel) antimicrobials, specifically bacteriocins.

To begin, with Chapter 1 we sought to investigate the literature for the potential of marine fish gut microbiomes as a source of novel bacteriocin producing strains. To do this, we first delved into the most current research on the microbial diversity of the marine fish gut microbiome. Early studies relied on culture-dependent methods for characterisation of the gut microbiomes of fish, yet such methods may be missing a significant number of taxa, as many microorganisms are thought to be unculturable or at least not culturable by current traditional media and plating. We therefore focused on studies that utilised culture-independent methods, such as metagenomics, and next-generation sequencing technologies. Following on from this, we carried out a thorough review of existing research on bacteriocins that have been isolated from marine fish microbiota in terms of their functionality, molecular diversity, and their applications. We concluded that accurate characterisation of the diversity of the marine fish gut microbiome, by utilising culture-independent methods, is essential in establishing its potential as a reservoir of bioactive compounds. Bypassing the need to culture the microbiota has facilitated the identification of uncommon, and sometimes novel, taxa. Furthermore, several of the cited studies reported not being able to identify operational taxonomic units (OTUs) to the genus level, emphasising the novelty of this microbial niche.

In terms of bacteriocin potential specifically, many studies have reported antimicrobial activity in marine fish gut isolates, however, most have not definitively identified the molecular source responsible for this activity. We have reported that 9 novel bacteriocins or bacteriocin-like inhibitory substances have been isolated and

characterised from marine fish gut microbiota, to date. The majority so far have come from bacilli and lactic acid bacteria, similar to reports regarding bacteriocin production in the human gut (Zheng et al., 2015). We also noted that the range of activity of these bacteriocins was quite variable. For example, Formicin, a novel class 1 bacteriocin (lanthipeptide) from *Bacillus paralicheniformis* APC 1576 isolated from the gut of a mackerel, was capable of targeting both Gram-positive and Gram-negative organisms. In particular, it was effective at targeting pathogenic strains such as *Clostridium difficile* and *Listeria monocytogenes*, the former of which is on the WHO's list of 'priority pathogens' – those which are amongst the currently top reported antimicrobial resistant pathogens and, as such, are of greatest concern (Asokan et al., 2019). Another bacteriocin piscicocin, VIa and VIb, had a narrow spectrum of activity, exclusively targeting *Listeria* spp.. This highlights that the scope for finding a wide range of bacteriocins from marine fish microbiota is high. In this chapter we also discuss some of the unique challenges and prospects of isolating bacteriocin from marine fish microbiota. For example, the difficulty in emulating *in situ* marine conditions can limit the metabolism of these bacteriocinogenic strains under laboratory conditions. Though, genomics-based methods and other emerging technologies are being used to great effect to combat these challenges.

Earlier work from our laboratory had already surveyed the microbiomes of deep-sea fish from both a compositional and functional perspective and also had isolated a number of strains with potential to produce antimicrobials (Collins, 2019). This study was possible given the accessibility to the Spanish research vessel, *Vizconde de Eza*, which harvested fish at different geographical locations across the Atlantic Ocean as part of a groundfish survey by the Spanish Institute of Oceanography (IEO). This study formed the genesis of work in Chapter 2 where we screened bacterial isolates from the skin and gut of various deep-sea fish species for antimicrobial producers using *in vitro* and *in silico* methods. Bacterial isolates were first tested for activity against the common indicator strain, *Lactobacillus delbrueckii* subsp. *bulgaricus* LMG 6091, using the soft agar colony overlay method. From this, 65 bioactive isolates were selected for taxonomic analysis and assessment of their ability to inhibit a range of target organisms. The first findings from this study were that the majority of the bioactive

isolates came from the skin of the fish and that the predominant taxa were of the genus *Pseudoalteromonas*. This is consistent with previous reports that *Pseudoalteromonas* strains, particularly those found in association with eukaryotic hosts, are capable of producing diverse bioactive secondary metabolites (Carola and Staffan, 1999).

Subsequent *in vitro* screening showed that the isolates were capable of inhibiting foodborne pathogens, such as *Listeria monocytogenes* and *Enterococcus* species.

Selected isolates were chosen for genomic screening for antimicrobial and bacteriocin genes. Across 36 genomes from various taxa, we detected over 140 biosynthetic gene clusters (BGCs) representing 16 different classes of secondary metabolites. The most frequently detected classes were RiPP-like, NRPS, siderophores, terpenes and betalactones. Based on core peptide analysis, we also identified four distinct gene clusters encoding potentially novel bacteriocins including cerecidin-like class I (lanthipeptide) bacteriocin peptides, an enterocin AS-48-like circular bacteriocin peptide and two lactococcin 972-like peptide sequences. The identification of the cerecidin-like peptides in *Planococcus* is significant, firstly, in that such a finding has not been reported before and, secondly, these prepeptides shared only ~50 % identity to the cerecidin peptides (cerA1–6 and cerA7) of *B. cereus*, indicating that they are potentially novel. Crucially, the *Planococcus* cerecidin-like BGC appeared to be complete and contained crucial genes related to lanthipeptide production, such as a LanM-like modification gene, a protease, and immunity and transport genes. Also, for the same reasons, the identification of a potentially novel enterocin AS-48-like bacteriocin on a complete BGC in *Arthrobacter* sp. APC 3897 was also a significant finding. Crucially, however, we detected negligible antimicrobial activity in the cell-free supernatant of these isolates, suggesting that the antimicrobial molecules may be cell bound or that their production is tightly regulated, possibly requiring coculturing in order to induce their production.

A significant finding from Chapter 2 was that many of the antimicrobial-producing isolates had relatively low 16S rRNA gene sequence similarity to species in the relevant databases, thus suggesting that they are novel. This led us to Chapters 3 and 4, in which we sought to describe two of these isolates as novel species. We also noted that relatively low frequency of antimicrobial resistance genes in the 36

genomes. A similar finding was also reported by Collins et al., (2019) and is likely a result of a low interaction between these deep-sea isolates and terrestrial antibiotics, and therefore a low level of acquired resistance. It is also possible that these isolates possess alternative systems relating to the production of and resistance to antibiotics.

One surprising result was that the 36 genomes characterised from these isolates were not represented in the 111 Metagenome Assembled Genomes (MAGs) which was an output from the above study by Collins et al., (2019). This suggests that the culturable isolates from this study were proportionally underrepresented in the deep-sea MAGs and may highlight the difficulties, and perhaps biases, in culturing many of these marine species as mentioned above. Indeed, many of these strains grew slowly on marine agar and were psychrophilic in nature.

It is not surprising that the screening of strains from deep-sea fish microbiomes yielded some really novel microorganisms. One of them belonging to the genus *Winogradskyella*, the identification and characterisation of which is described in Chapter 3. It was initially identified as a strain that killed *Lactobacillus bulgaricus* but lost its bioactivity after subsequent culturing *in vitro* (Chapter 2). This apparent loss of antimicrobial activity has been noted before in our lab group in previous works and can be due to various reasons, such as suboptimal culturing conditions and mutations. The APC strain 3343 was isolated from the intestine of a deep-sea loosejaw dragon fish, *Malacosteus niger*. *Winogradskyella* are most commonly found in association with algae and invertebrates. It is therefore curious as to how a *Winogradskyella* strain could be found in the gut of a deep-sea fish. We proposed that it was most likely acquired through diet. We had reported previously that this strain encoded a putative sactipeptide gene (Chapter 2). Phylogenetic analysis, based on 16S rRNA and genomic sequences, and biochemical and phenotypic characterisation clearly distinguished APC 3343 from its nearest relatives amongst the type strains of *Winogradskyella* species. APC 3343 was therefore deemed the type strain of a novel species, *Winogradskyella bathintestinalis* (in English meaning 'of the deep intestine').

We also describe some key characteristic features of the genome of strain APC 3343 such as its low thermotolerance (it was incapable of growth above 30 °C), wide NaCl range for growth (of 0.5-7 % NaCl), fastidious nature, and unique cellular fatty acid profile compared to the type strains – this difference in cell wall composition is likely a reflection of cellular adaptations to being from very different environments. We also noted that APC 3343 formed yellow/orange -pigmented colonies, a characteristic of *Winogradskyella* species and usually the result of carotenoid production - this corresponds to our previous finding in Chapter 2 that many of these deep-sea strains, including APC 3343, encode terpenes (from which carotenoids are derived). Carotenoids themselves are interesting metabolites from a biotechnological point of view, as they have a wide variety of applications in the food, cosmetics and pharmaceutical industries (Moyo et al., 2022). β -carotene, for example, is perhaps one of the most well-known carotenoids, giving red, orange and yellow fruits and vegetables their colour, and is an important precursor of vitamin A, which can only be acquired by humans through diet. There is a growing demand for natural sources of β -carotene as alternatives to its chemical synthesis.

Another really novel species and type strain, *Planococcus notacanthi* APC 4016, is described in chapter 4. APC 4016 was isolated from the skin of a snub-nosed spiny eel, *Notacanthus chemnitzii*. This strain had shown consistent antimicrobial activity in our screening study against several Gram-positive target strains, including *Micrococcus luteus* and *Arthrobacter* sp. APC 3897 and, interestingly, was most active against *Planococcus* sp. APC 3906, a closely related strain from the same isolation source. This maybe not so surprising as often bacteriocins produced by particular organisms are narrow spectrum but particularly active against genetically related organisms that compete with them in a similar ecological niche. APC 4016 was subsequently found to encode one of the putative novel cerecidin-like lanthipeptide clusters. 16S rRNA gene sequence analysis showed that APC 4016 was closely related to members of the genus *Planococcus/Planomicrobium*. Its nearest related type strain was *Planomicrobium okeanokoites* IFO 12536^T, to which it shared 98.98 % 16S rRNA gene identity, which is above the accepted threshold of 98.6 % for species (Chun et al., 2018). However, the overall genomic similarity of these two species was low

and they could be clearly differentiated in the biochemical and phenotypic assays. APC 4016 was therefore deemed to be the type strain of a novel species, *Planococcus notacanthi* (in English the species name pertains to the fish from which it was isolated, *Notacanthus chemnitzii*). Some notable characteristics of this strain include its ability to grow under a wide range of growth conditions, such as temperatures of 4–40 °C and in up to 12 % NaCl, its sensitivity to all the tested antibiotics, and its production of orange-pigmented colonies – which, as above, corresponds to our genome analysis of APC 4016 in that it encoded two terpene genes but no antimicrobial resistance gene markers.

A significant output of this thesis has been the application of deep-sea isolates and their DNA as controls for determining relative and/or absolute quantification of bacteria in human metagenomic sequencing experiments. Commercial kits are available for such applications, however, their application in infant samples with a potentially very low microbial load is limited. In particular, we chose two strains, *Pseudoalteromonas* sp. APC 3896 and *Planococcus* sp. APC 3900, on the basis of one being Gram-negative and the other Gram-positive, respectively, and also on their absence in human metagenomic databases. Furthermore, strain APC 3900 (originally isolated from the skin of a *Notacanthus chemnitzii* specimen) shared 100 % 16S rRNA gene sequence identity and 98.3 % average nucleotide identity (from genomic sequence analysis) to APC 4016, indicating they are the same novel species; therefore, APC 3900 is even less likely to be found in human samples. Strain APC 3896 (originally isolated from the skin of a *Bathysaurus ferox* specimen) was most closely related to *Pseudoalteromonas tetraodonis* GFC (99.93 % 16S identity). *P. tetraodonis* strains produce the highly potent tetrodotoxin neurotoxin, which usually accumulates in aquatic organisms that host the producing bacteria making them highly toxic (SIMIDU et al., 1990). A well-known example of this is the pufferfish (family *Tetraodontidae*). Therefore, the presence of *P. tetraodonis* strains in infant samples is highly unlikely. Indeed, in subsequent sequencing experiments, to date, we have never detected them in control samples. Initially these experiments evaluated the utility of using either spiked samples with cells of the two strains or spiked pure DNA from the deep-sea strains into DNA isolated from the samples. The

results showed that the latter approach proved better, presumably due to the relative ease of lysing Gram-negative over Gram-positive bacteria. Interestingly, it was estimated that strains APC 3900 and APC 3896 contained 9 and 8 copies of the 16S rRNA gene, respectively, which needed to be taken into account in calculating the copy numbers per 1ng of extracted DNA and in order to optimise the required concentration of spike-in DNA. These results are outlined in the study detailed in the Appendix to Chapter 4 where we have outlined our experimental approach of preparing deep-sea spike-in DNA, applying it to mother/infant sample DNA and its subsequent detection using both 16S rDNA high-throughput sequencing (HTS) and quantitative PCR (qPCR).

These studies are still ongoing, and require optimisation, particularly with regards to achieving consistent and replicable spike-in % relative abundance output values. However, the initial results are very promising in that genomic DNA from these novel deep-sea strains can be spiked into mother/infant samples and subsequently detected using these methods. We aim to routinely use them in all future sequencing to be done on the MiMIC project.

Another really surprising finding in this thesis was the identification of isolates which could degrade the agar used in the standard preparation of the solid media. In this case concave pits around individual colonies on solid agar were observed. In Chapter 5, we describe the identification of an agarolytic (agar-degrading) isolate and describe the subsequent molecular characterisation and cloning of one β -agarase and one putative B agarase. The producing strain, APC 4017, was found to belong to the genus *Pseudoalteromonas* and could degrade solid agar at low temperatures. Agarases that are active at low temperatures have potential applications in temperature-sensitive processes, such as in DNA gel extraction, and in facilitating more environmentally friendly (in terms of heat and carbon production) reactions. Current commercial agarases used in agarose gel extractions, for example, require temperatures of at least 40 °C for activity and low-melt agarose in the liquid phase. The genes for the two putative β -agarases in our study, *aga4017A* and *aga4017B*, were found in a large gene cluster of at least 10 genes – many of which are homologues of genes associated with carbohydrate hydrolysis. Previous reports (as

mentioned in Chapter 5) have found agarase genes on large 'polysaccharide-utilisation loci' which can be over 100 kb in length and involve multiple agarases of different families, as well as other enzymes. These loci are dedicated to the complete (or near complete) degradation of agar and its derivatives. It is possible that strain APC 4017 encodes more agarases that were not detected in our annotations. The cooperative degradation of agarose by multiple bacteria has also been reported. For example, Yun et al., (2021) investigated agarose degradation in two human gut isolates, *Bacteroides plebeius* and *Bifidobacterium longum* ssp. *Infantis*. They showed that *B. plebeius* and *B. longum* degraded and metabolised agarose and the resulting derivatives consequently in turn, ultimately resulting in the production of 3,6-anhydro-L-galactose (a sugar monomer). The coordinated metabolism of otherwise indigestible nutrients may be an efficient means of survival in a competitive environment such as the human gut, or even the deep-sea. Amino acid sequence analysis showed that Aga4017A and Aga4017B had structural similarities and key characteristics corresponding to GH16 and GH50 family β -agarases, respectively. Significantly, Aga4017B showed relatively low sequence similarity to characterised β -agarases, sharing only 51.81 % identity to its nearest characterised homolog from the PDB database. This suggests that Aga4017B may represent a novel uncharacterised GH50 β -agarase or may be a related carbohydrate-binding protein with similar functionality to GH50 β -agarases. *E. coli* TOP10 transformants carrying the recombinant plasmid pCR2.1-TOPO_aga4017A vector were positive for agar-degradation. The pCR2.1-TOPO cloning system with TOP10 cells is not typically used for expression, therefore the observed agarolytic activity could be due to plasmid leakage. Nonetheless, it shows that Aga4017A is indeed an agar-degrading enzyme. Elucidation of the predicted 3D model of Aga4017A illustrated that it was a globular protein organised into three distinct modules, which when correlated back to the predicted conserved domains of the protein sequence are: the signal peptide sequence, GH16 β -agarase catalytic domain and the carbohydrate-binding domain. Furthermore, this study represents one of the first to characterise a β -agarase protein sequence from a deep-sea fish microbiome isolate. Another potential application of these agarases could be in giving a nutritional advantage to probiotic cultures into which they are cloned. A possibility here would be to clone a β -agarase into a GRAS

or commensal gut organism, such as *Lactobacillus*, which would then be tested in mice fed agar or its derivatives (or even seaweed). The utilisation of the agar by the lactobacilli may then give them a competitive advantage over other commensal microbes in the microbiome which could enhance persistence and functional activity. Furthermore, products of agarose degradation, such as neoagaro-oligosaccharides, have been shown to promote the growth of gut-benefiting bacteria, such as bifidobacteria, thus having prebiotic effects (Hu et al., 2006). In this case the possession of agarose could be viewed as a probiotic feature in of itself. Further work is therefore required in developing recombinant expression vector(s) encoding the novel Aga4017A agarase gene, or even both Aga4017A and the putative β -agarase, Aga4017B.

Taken together, this thesis shows that not only is the microbiome of deep-sea fish a source of cultivable antimicrobial bacteria but that they encode a wide range of antimicrobial BGCs, including potentially novel bacteriocins. The identification of novel species further highlights this unique microbial niche as a source of novelty (and reinforces its potential as a source for production of secondary metabolites such as carotenoids). It is also shown that, as well as for sourcing novel antimicrobials, the deep-sea fish microbiome is a source of other bioactive functionalities which may have applications in the exploitation of microbiome research.

References

- ASOKAN, G. V., RAMADHAN, T., AHMED, E. & SANAD, H. 2019. WHO global priority pathogens list: a bibliometric analysis of medline-pubmed for knowledge mobilization to infection prevention and control practices in Bahrain. *Oman Medical Journal*, 34, 184-193.
- CAROLA, H. & STAFFAN, K. 1999. Marine *Pseudoalteromonas* species are associated with higher organisms and produce biologically active extracellular agents. *FEMS Microbiology Ecology*, 30, 285-293.
- CHUN, J., OREN, A., VENTOSA, A., CHRISTENSEN, H., ARAHAL, D. R., DA COSTA, M. S., ROONEY, A. P., YI, H., XU, X.-W., DE MEYER, S. & TRUJILLO, M. E. 2018. Proposed minimal standards for the use of genome data for the taxonomy of prokaryotes. *International Journal of Systematic and Evolutionary Microbiology*, 68, 461-466.
- COLLINS, F. W. J. 2019. *An investigation into antimicrobial production in the Lactobacillus genus and the fish microbiome*. PhD Thesis, University College Cork, Cork, Ireland.
- DANOVARO, R., CORINALDESI, C., DELL'ANNO, A. & SNELGROVE, P. V. R. 2017. The deep-sea under global change. *Current Biology*, 27, R461-R465.
- FRIERI, M., KUMAR, K. & BOUTIN, A. 2017. Antibiotic resistance. *Journal of Infection and Public Health*, 10, 369-378.
- GREENAWAY, S. F., SULLIVAN, K. D., UMFRESS, S. H., BEITTEL, A. B. & WAGNER, K. D. 2021. Revised depth of the Challenger Deep from submersible transects; including a general method for precise, pressure-derived depths in the ocean. *Deep Sea Research Part I: Oceanographic Research Papers*, 178, 103644.
- HU, B., GONG, Q., WANG, Y., MA, Y., LI, J. & YU, W. 2006. Prebiotic effects of neoagaro-oligosaccharides prepared by enzymatic hydrolysis of agarose. *Anaerobe*, 12, 260-266.
- MOYO, A. C., DUFOSSÉ, L., GIUFFRIDA, D., VAN ZYL, L. J. & TRINDADE, M. 2022. Structure and biosynthesis of carotenoids produced by a novel *Planococcus* sp. isolated from South Africa. *Microbial Cell Factories*, 21, 43.
- NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION, U. S. D. O. C. *Layers of the Ocean* [Online]. Available at: <https://www.noaa.gov/jetstream/ocean/layers-of-ocean> (Accessed March 28, 2023).
- RESISTANCE, I. C. G. O. A. 2019. No Time to Wait: Securing the future from drug-resistant infections. *Report to the Secretary-General of the United Nations*. Available at: <https://www.who.int/publications/i/item/no-time-to-wait-securing-the-future-from-drug-resistant-infections> (Accessed on 5 January 2024).

- SIMIDU, U., KITA-TSUKAMOTO, K., YASUMOTO, T. & YOTSU, M. 1990. Taxonomy of four marine bacterial strains that produce tetrodotoxin. *International Journal of Systematic and Evolutionary Microbiology*, 40, 331-336.
- SOMERO, G. N. 1992. Biochemical ecology of deep-sea animals. *Experientia*, 48, 537-543.
- YUN, E. J., YU, S., PARK, N. J., CHO, Y., HAN, N. R., JIN, Y.-S. & KIM, K. H. 2021. Metabolic and enzymatic elucidation of cooperative degradation of red seaweed agarose by two human gut bacteria. *Scientific Reports*, 11, 13955.
- ZHENG, J., GÄNZLE, M. G., LIN, X. B., RUAN, L. & SUN, M. 2015. Diversity and dynamics of bacteriocins from human microbiome. *Environmental Microbiology*, 17, 2133-2143.

Acknowledgements

Giorraíonn beirt bóthar.

A pair shortens the journey.

Firstly, I would like to sincerely thank my supervisors, Prof. Paul Ross, Prof. Colin Hill, Prof. Catherine Stanton and Dr. Mary Rea for everything, particularly for their support, encouragement and patience throughout the years. I'm truly grateful for the opportunity that I have been given and for the experience, from which I've learned and gained so much.

Thank you to Fergus C., who began this expedition, for all his help, for answering all my questions, and for getting the fish. To Lorraine, Niki and Neda, thank you for all the advice and ideas, for helping me keep on top of things and for getting me through. Thank you to Des for sharing all his tips and tricks with me when I was just getting started and didn't know how to do an overlay. Thank you to Colin B., Dave and Veronica for all their help on the computer tech side of things.

Thank you also to Claire, Natasha and Rachel for making the commute to Moorepark so much easier and more enjoyable. Between Cork and Fermoy, I've been in too many labs and met too many wonderful people to thank them all by name, so I will just give a shoutout to everyone (past and present) in Labs 337 (UCC), 3.63 (UCC), 4.25 (UCC), APC One (Moorepark), APC Two (Moorepark), and Lab 4.23 (UCC). Also, I would like to thank the Food Bioscience and Food Chemistry gangs in Moorepark for all the lunchtime chats.

Thank you also to my friends, my training partners, and my coaches in the athletics club(s) for their ongoing support and advice, on and off the track. Lastly, thank you to my family for their advice (and for putting up with me), and especially to my grandparents for all the biscuits and cups of tea.

Go raibh míle maith agaibh.

Thank you all very much.