Genome mining of marine bacteria for bioactive metabolites

DISSERTATION

zur

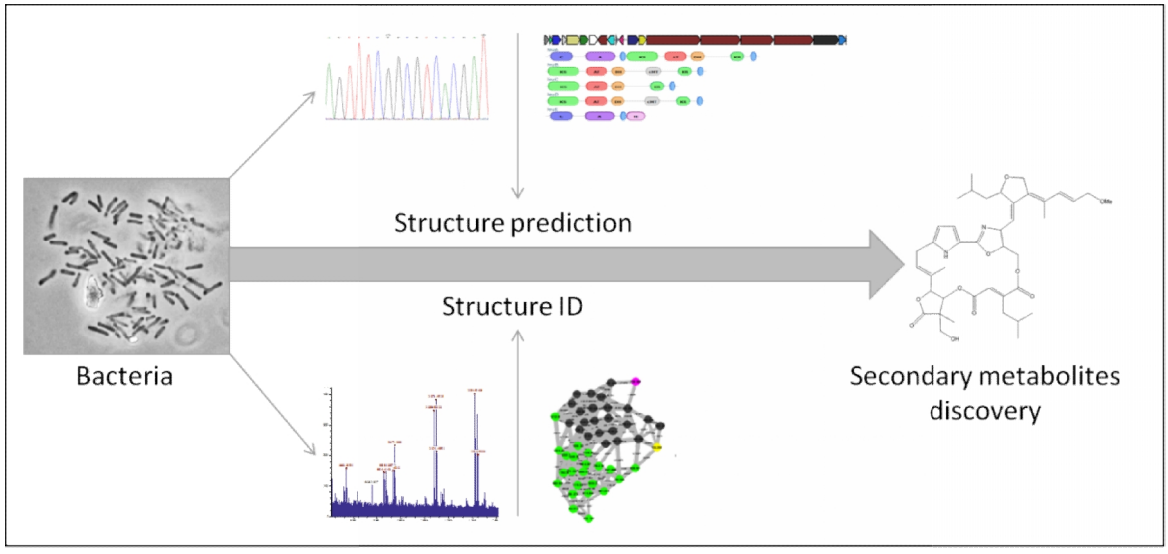
Erlangung des Doktorgrades (Dr. rer. Nat.)

der

Mathematisch-Naturwissenschaftlichen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität Bonn



vorgelegt von

JAMSHID AMIRI MOGHADDAM

aus

Kashmar, Iran

Bonn, 2018

Angefertigt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät
der Rheinischen Friedrichs-Wilhelms- Universität Bonn

1.Gutachterin: Prof. Dr. Gabriele M. König

2.Gutachter: Prof. Dr. Till F. Schäberle

Tag der Promotion: 10.01.2019

Erscheinungsjahr: 2019

In advance Publications of the Dissertation

Parts of this study have been published in advance under permission of the Faculty of Mathematics and Natural Sciences, represented by the supervisor of this study.

Amiri Moghaddam J, Crüsemann M, Alanjary M, Harms H, Dávila-Céspedes A, Blom J, Poehlein A, Ziemert N, König GM, Schäberle TF. **Analysis of the genome and metabolome of marine Myxobacteria reveals high potential for biosynthesis of novel specialized metabolites.** Scientific reports. 2018 Nov 9; 8(16600). DOI:10.1038/s41598-018-34954-y.

Amiri Moghaddam J, Dávila-Céspedes A, Kehraus S, Crüsemann M, Köse M, Müller CE, König GM. **Cyclopropane-containing fatty acids from the marine bacterium *Labrenzia* sp. 011 with antimicrobial and GPR84 activity .** Marine drugs. 2018 Oct 8; 16(10). DOI: 10.3390/md16100369.

Amiri Moghaddam J, Boehringer N, Burdziak A, Kunte HJ, Galinski EA, Schäberle TF. **Different strategies of osmoadaptation in the closely related marine Myxobacteria *Enhygromyxa salina* SWB007 and *Plesiocystis pacifica* SIR-1.** Microbiology. 2016 Feb 3. DOI: 10.1099/mic.0.000250.

Amiri Moghaddam J, Poehlein A, Fisch K, Alanjary M, Daniel R, König GM, Schäberle TF. **Draft genome sequences of the obligatory marine myxobacterial strains *Enhygromyxa salina* SWB005 and SWB007**. Genome Announc. 2018 Apr 26;6(17). pii: e00324-18. DOI: 10.1128/genomeA.00324-18.

Amiri Moghaddam J, Dávila-Céspedes A, Alanjary M, Blom J, König GM, Schäberle TF. **Draft genome sequence of the marine bacterium *labrenzia* sp. 011, a potential protective agent of mollusks.** Data in Brief. (*under review*).

König GM, Felder S, Kehraus S, Bouhired S, Harms H, Schäberle TF, **Moghaddam JA**, Mohseni MM. **The secondary metabolome of marine-derived microorganisms.** 62nd International Congress and Annual Meeting of the Society for Medicinal Plant and Natural Product Research September 2014. In Planta Med 80 (16). DOI: 10.1055/s-0034-1394477.

Acknowledgments

With this thesis I come to the end of the road of my PhD studies, which I could not have accomplished without the help and support of several people, to whom I am highly indebted.

First and foremost I would like to express my sincere gratitude to my supervisor Prof. Dr. G. M. König for giving me the opportunity to join her working group and for supporting me from the beginning to the end of this road. I thank her for entrusting me with this interesting project, which had been a real challenge for me, and which without her continuous support and encouragement would not have been possible to conclude.

Special thanks to Prof. Dr. T. F. Schäberle for supervising me during my work, editing my publications, and officiating as the second reviewer.

I really appreciate Prof. Dr. H. Wägele and Prof. Dr. Prof. Dr. Karl Wagner for their kind participation in the examination committee.

I am highly thankful to Dr. Max Crüsemann, Dr. Henrik Harms and Dr. Stefan Kehraus for helping me in mass spectrometry analysis, chemistry analysis, and their contribution in editing my publications.

Thanks are also due to Prof. Bradley Moore (Center for Marine Biotechnology & Biomedicine, Scripps Institution of Oceanography, USA), for giving me a useful research stay in his working group.

My sincere thanks goes out to our technical support staff of the Institute for Pharmaceutical Biology, University of Bonn, including Edith Neu, our former lab assistant, for teaching me how to cultivate the very difficult marine Myxobacteria , to Emilie Gorlaski and Ekaterina Eguereva for supporting me in the lab, and to Thomas Kögler for helping in all technical emergencies always with a receptive smile.

I must also acknowledge Dr. Mohammad Alanjary (Univerity of Tübingen) and Dr. Jöchen Blom (Univeristy of Gießen) for helping me in bioinformatic analysis.

To everyone past and present in the working group König, a big thank you. It was a real pleasure getting to know all of you, you made daily lab work less tedious and more pleasant. Special thanks go to Nils Böhringer and Luis J. Linares Otoya for interesting scientific discussions and being a good friends.

I would like to thank my small family for teaching me strength and patience, my lovely wife Fatima with her everyday support and my daughter Jasmin who took her first steps here in Germany and bore with me the good and bad times.

This research would not have been possible without financial assistance of the Ministry of Science, Research, and Technology of Iran for the first three years of my PhD study, and the University of Bonn and the DAAD for a short support at the end of my PhD.

I appreciated a lot the possibility to attend diverse seminars and workshops provided by BIGS Drugs and International office of the University of Bonn.

Table of Contents

Abbreviations

AAI	Average amino acid identity
ACP	Acyl carrier protein
ANI	Average nucleotide identity
APE	Aryl polyene
BGC	Biosynthetic gene cluster
BLAST	Basic local alignment search tool
bp	Base pairs
CDS	Coding sequences
CFAS	Cyclopropane fatty acid synthases
Da	Dalton
DNA	Deoxyribonucleic acid
EPS	Exopolysaccharides
HGT	Horizontal gene transfer
isDDH	*In silico* DNA-DNA hybridization
Mbp	Mega base pairs
MEP	Methylerythritol 4-phosphate
mmCoA	Methylmalonyl-CoA
mCoA	Malonyl-CoA
emCoA	Ethylmalonyl-CoA
pCoA	Propionyl-CoA
cCoA	Crotonyl-CoA
CCR	Crotonyl-CoA carboxylase/reductase
MVA	Mevalonate
NRP	Non-ribosomal peptide
NRPS	Non-ribosomal peptide synthetase
OD	Optical density
PHB	Polyhydroxybutyric acid
PK	Polyketide
PKS	Polyketide synthase
rDNA	Ribosomal DNA
RiPP	Ribosomally synthesis and post-translationally Modified peptide
SAM	S-adenosyl-L-methionine
sp./spp.	Species /species (plural)

1 Introduction

1.1 Marine bacteria as a source of secondary metabolites

Traditionally, discovery of natural products (NP) also called "secondary metabolites or specialized metabolites" has been restricted to species from terrestrial environments, and no one considered the ocean until about 1970 (Marris 2006; Dias et al. 2012; Jensen and Fenical 1994). Thereafter, marine environments, supporting 95% of the biosphere with rich and unique biodiversity, have come into the focus for discovery of novel natural products (Naughton et al. 2017).

Marine environments are unique in terms of pH, temperature, pressure, oxygen, light, nutrients, salinity, and especially rich in chlorine and bromine. Therefore, it is not surprising that marine bacteria, living in a biologically competitive environment, produce metabolites, which exhibit special activities (Blunt et al. 2004). Marine bacteria survive under harsh conditions such as cold, lightlessness and high pressure; nevertheless, they produce fascinating and structurally complex natural products. They also associate with physically unprotected soft-bodied marine organisms and protect them by production of bioactive secondary metabolites (Jensen and Fenical 1994).

The ecology of marine natural products revealed that many of these compounds are chemical weapons that have evolved to be highly potent inhibitors of physiological processes in prey, predators or competitors. Thus, marine organisms utilize their secondary metabolites for survival (Venter et al. 2004). Many bioactive natural products from marine bacterial isolates showed antimicrobial activity, such as the tauramamides from *Brevibacillus laterosporus* PNG276 (Desjardine et al. 2007), and cytotoxic activity such as salinosporamide A from *Salinispora tropica* (Feling et al. 2003). In addition, marine bacteria can generate toxic compounds, e.g. polybrominated diphenyl ethers (PBDEs) which bioaccumulate in marine animals and are likely transferred to the human food chain (Agarwal et al. 2017).

The above mentioned examples demonstrate that the vast number of marine bacteria living in the ocean environments offer great opportunities for biodiscovery (Joint et al.

2010). Compared to terrestrial bacteria, only a small number of these microorganisms have been investigated for their bioactive metabolites (Debbab et al. 2010). However, research into bioactive metabolites from marine bacteria has increased dramatically since 1990. Most of the compounds were isolated from the five bacterial phyla Proteobacteria, Actinobacteria, Cyanobacteria, Firmicutes, and Bacteroidetes (Nikapitiya 2012). In fact, bacterial taxonomy and metabolic diversity are fundamentally interrelated (Jensen and Fenical 1994). Therefore, understanding the correlation between the taxonomy and biosynthetic diversity of marine bacteria is important for the discovery of natural products.

1.2 Diversity of marine bacteria

Different ecosystems of the world's oceans host distinct bacterial communities. They are found in surface and deep water, coastal and open ocean environments, and anoxic and oxic conditions, which all greatly differ at all taxonomic levels (only <10% of bacterial types defined at 3% sequence similarity level) (Zinger et al. 2011). In a comprehensive study of the bacterial diversity from marine environments, 120,000 molecular operational taxonomic units with 3% sequence dissimilarity threshold ($OTU_{0.03}$) were identified from 509 global samples. From the identified taxonomic units, Gammaproteobacteria and Alphaproteobacteria were dominant in ocean waters, followed by a high proportion of Flavobacteria and Cyanobacteria. The dominant groups in benthic communities were Gammaproteobacteria, Deltaproteobacteria, Planctomycetes, Actinobacteria, and Acidobacteria (Figure 1.1) (Zinger et al. 2011). However, this richness is by far below the millions of bacterial types predicted by log-normal models (Curtis et al. 2002).

This enormous biodiversity of marine bacteria has boosted the discovery and biotechnological applications of marine bacterial natural products. Further, technical advances in cultivation methodologies and new isolation techniques such as I-chip allow the researchers to access a large and diverse array of previously uncultivable bacteria from marine environments (Nichols et al. 2010).

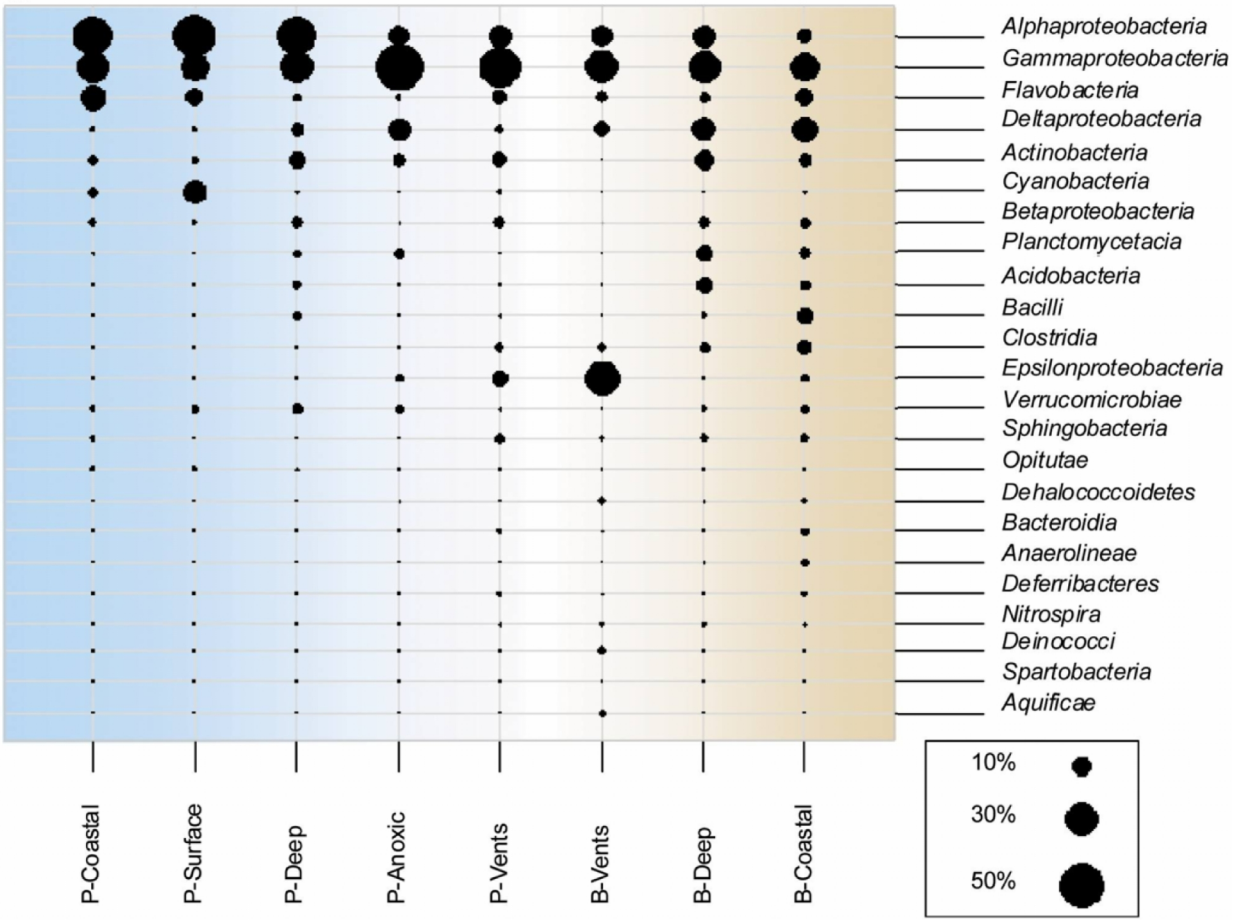


Figure 1.1 Average proportions of the main bacterial taxa per realm and ecosystem type. P= Pelagic, B= Benthic. Notice that taxonomic levels displayed here are not necessarily of the same level, but reflect the most common levels whose ecology and diversity are usually investigated in marine microbiology. (Taken from Zinger, et al. Global Patterns of Bacterial Beta-Diversity in Seafloor and Seawater Ecosystems. PLoS One. 2011;6(9):e24570).

1.3 Genomics and other -omics in secondary metabolites discovery

Bacterial secondary metabolites are biosynthesized by enzymes encoded in biosynthetic gene clusters (BGCs) in the genome of bacteria. A BGC is a set of genes which encode enzymes that collectively work together in the same pathway (Giordano et al. 2015). Advances in DNA sequencing made it possible to completely sequence bacterial genomes in a short time and at low price. The information from tens of

thousands of bacterial genomes has had a major impact on our views of the bacterial world. Today, the BGCs of secondary metabolites can be identified using genome mining software tools like antiSMASH, which represents both, a biosynthetic and an evolutionary unit of the natural products (Blin et al. 2017b). Genome mining has therefore led to a paradigm shift in natural products research, as BGCs can be identified and characterized from the genomes of many even unculturable microorganisms (Cimermancic et al. 2014).

Moreover, comparative genomics provides a unique opportunity to retrieve valuable information regarding genome structure, functional diversity, and evolution of marine microorganisms (Fernández-Gómez et al. 2012). Therefore, the genomic approach increases the chance for discovery of new metabolites by identifying the talented microbes using genome sequence analysis and subsequent characterization of the *in silico* identified BGCs (Naughton et al. 2017). Also, advances in sampling methodologies, metagenomic, direct sequencing, and single-cell genomics together with specific bioinformatic tools for natural products, facilitate the retrieval of previously inaccessible BGCs from marine sources (Giordano et al. 2015). It may be the immediate future that environmental genomic data are accessible for such biochemical pathways existing in microbes with no representative isolates presently in laboratory culture (Mühling et al. 2013).

Additionally, the genomic approach can be combined with other Meta-omic approaches such as metabolomics, proteomics, and transcriptomics, which allow culture-independent study of previously elusive microorganisms for novel natural products (Figure 1.2) (Schofield and Sherman 2013). Using high resolution mass spectrometry and molecular networking, rediscovery of known metabolites can be avoided at a very early stage of the discovery process through dereplication (Maansson et al. 2016; Yang et al. 2013), and simultaneously, discovery of novel natural products can be streamlined through the optimization of culture conditions (Crüsemann et al. 2017).

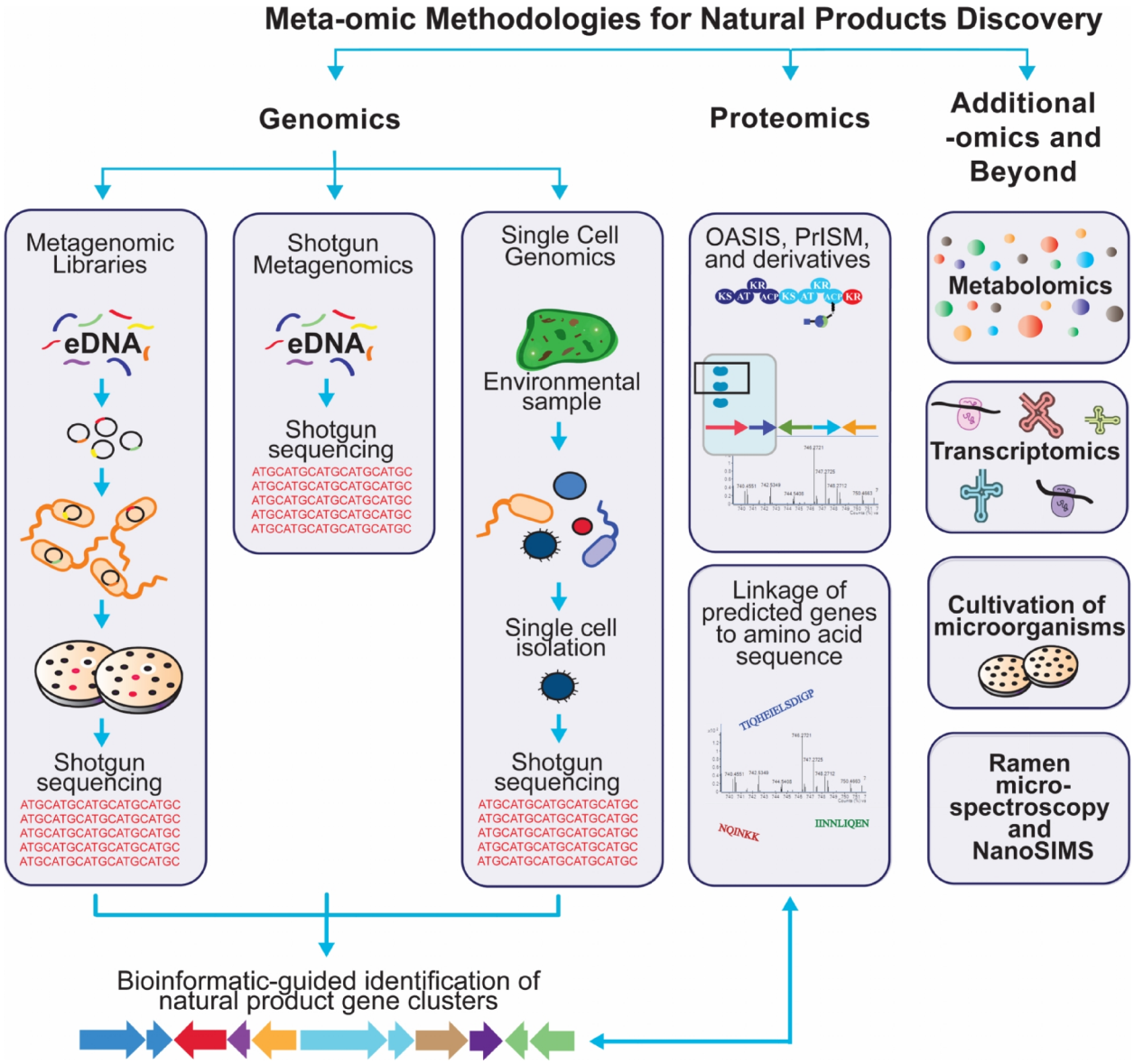


Figure 1.2 An overview of meta-omic methodologies for natural products discovery. (Taken from Schofield MM, Sherman DH. Meta-omic characterization of prokaryotic gene clusters for natural product biosynthesis. Curr Opin Biotechnol. 2013 Dec;24(6):1151-8).

1.4 Evolution of secondary metabolite biosynthesis in bacteria

Secondary metabolites confer selective advantages for bacteria, either as antibiotics or by providing a chemical language that allows communication among species, with other organisms and their environment. These selective advantages led to the evolution of

secondary metabolite biosynthesis within bacteria (Giordano et al. 2015). Many of the secondary metabolites BGCs are located in genomic islands, which are strain-specific regions of the chromosome that are generally acquired by horizontal gene transfer (HGT). The genomic islands harbor functionally adaptive traits and drive the genomic differences among closely related strains (Penn et al. 2009). The evolution of secondary metabolites was probably the result of the selection of specific traits to increase the chances to develop a compound with potent bio-molecular activity. The rates of evolutionary events such as insertions, deletions and duplications within the BGCs are much higher than in comparable primary metabolism gene clusters (Cimermancic et al. 2014). In addition, co-evolving gene clusters seem to play key roles in the evolution of larger BGCs encoding complex metabolites. Finally, BGC families evolve in specific modes, in many of which concerted evolution plays an important role (Giordano et al. 2015).

1.5 Diversity of biosynthetic gene clusters in bacteria

In the last decade, genome mining tools were established to compute BGCs from available genomes. One of the most popular tools is AntiSMASH database, which gives an exploratory overview of available biosynthetic gene cluster predictions and provides general statistics about secondary metabolite cluster counts from the search queries (Blin et al. 2017a). In AntiSMASH database, 44 different families of BGCs are clustered (until 19th July, 2018), which contains 32548 BGCs from 6200 unique species/strains (Figure 1.3).

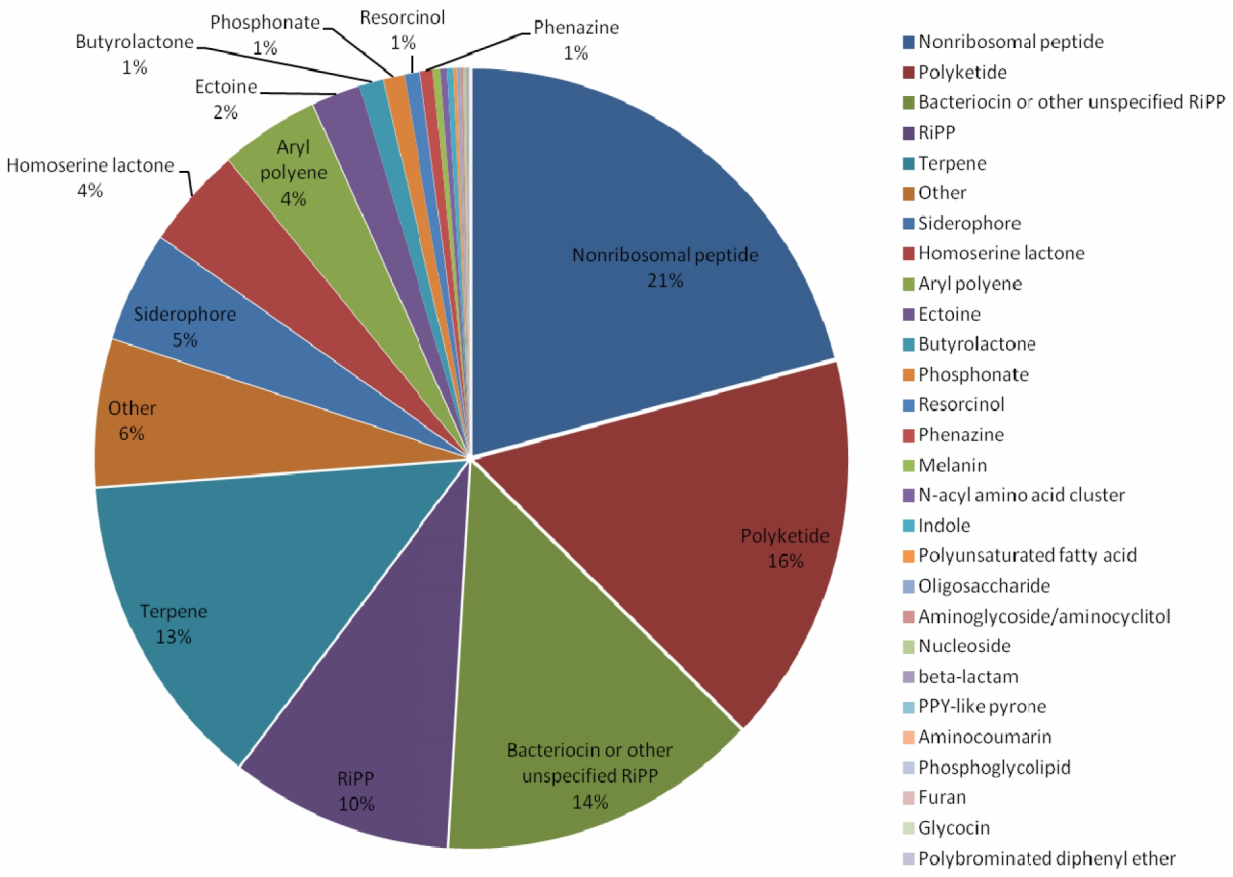


Figure 1.3 Distribution of biosynthetic gene clusters in bacteria; AntiSMASH database. (19th July 2018, https://antismash-db.secondarymetabolites.org/#!/stats). Polyketide BGCs include Type I PKS, Type II PKS, Type III PKS, Trans-AT polyketide, and other types of polyketides clusters. RiPP (ribosomally synthesis and post-translationally modified peptide) BGCs include Lanthipeptide, Thiopeptide, Lasso peptide, Sactipeptide, Ladderane, Cyanobactin, Microviridin, Head-to-tail cyclised (subtilosin-like), Proteusin, Linaridin, Pheganomycin-like ligase, Microcin, and Bottromycin-like clusters (Arnison et al. 2013). Other BGCs containing secondary metabolite related genes that do not fit into any of the other categories.

A larger database, IMG/M (Integrated Microbial Genomes and Microbiomes), includes already 61568 genomes and 808236 BGCs in its public database (19th July, 2018). The genomes available in this database belong to Proteobacteria (46.6%), Firmicutes (25%), Actinobacteria (12.6%), and Bacteroides (3.7%) (Figure 1.4) (Chen et al. 2017). Interestingly, the percentage of the BGCs derived is not consistent with the percentage of the genomes from the respective Phylum. Therefore, 44.2 % of the BGCs belong to Proteobacteria, 30.6% Actinobacteria, 19.5% Firmicutes, and 2.8% Bacteroides.

Comparing the number of the genomes and BGCs, it can be concluded that the Actinobacteria harbor more BGCs as compared to other Phyla.

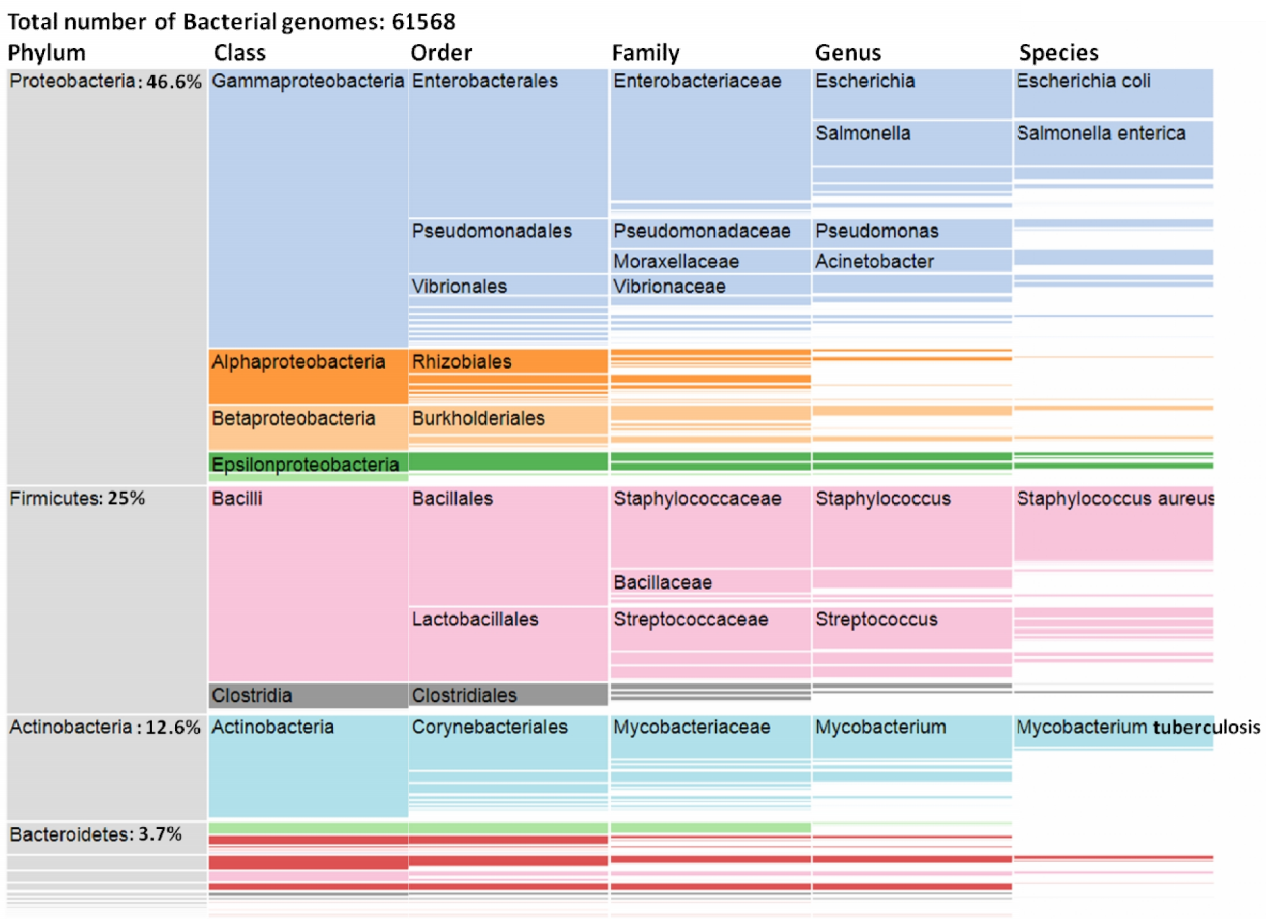


Figure 1.4 Bacterial genome sequences deposited in the IMG/M database. (19th July 2018, acquired from https://img.jgi.doe.gov/cgi-bin/m/main.cgi?section=TreeFile&page=domain&domain=bacteria). Similar color codes contributed to the subgroups of the same class of bacteria. White regions are unclassified under the respective taxonomic rank.

Further, IMG-ABC (Integrated Microbial Genomes system and Atlas of Biosynthetic gene Clusters) predicts and annotates BGCs across all microbial genomes using a combination of Clusterfinder algorithm and AntiSMASH (Hadjithomas et al. 2017). Similar to AntiSMASH database, NRPS and PKS BGCs are the most common BGC types followed by bacteriocin, terpene, and siderophore BGCs (Figure 1.5). Applying

clusterfinder algorithm revealed 56% of BGCs are annotated as putative and could not be assigned by AntiSMASH to any known BGC class (Figure 1.5).

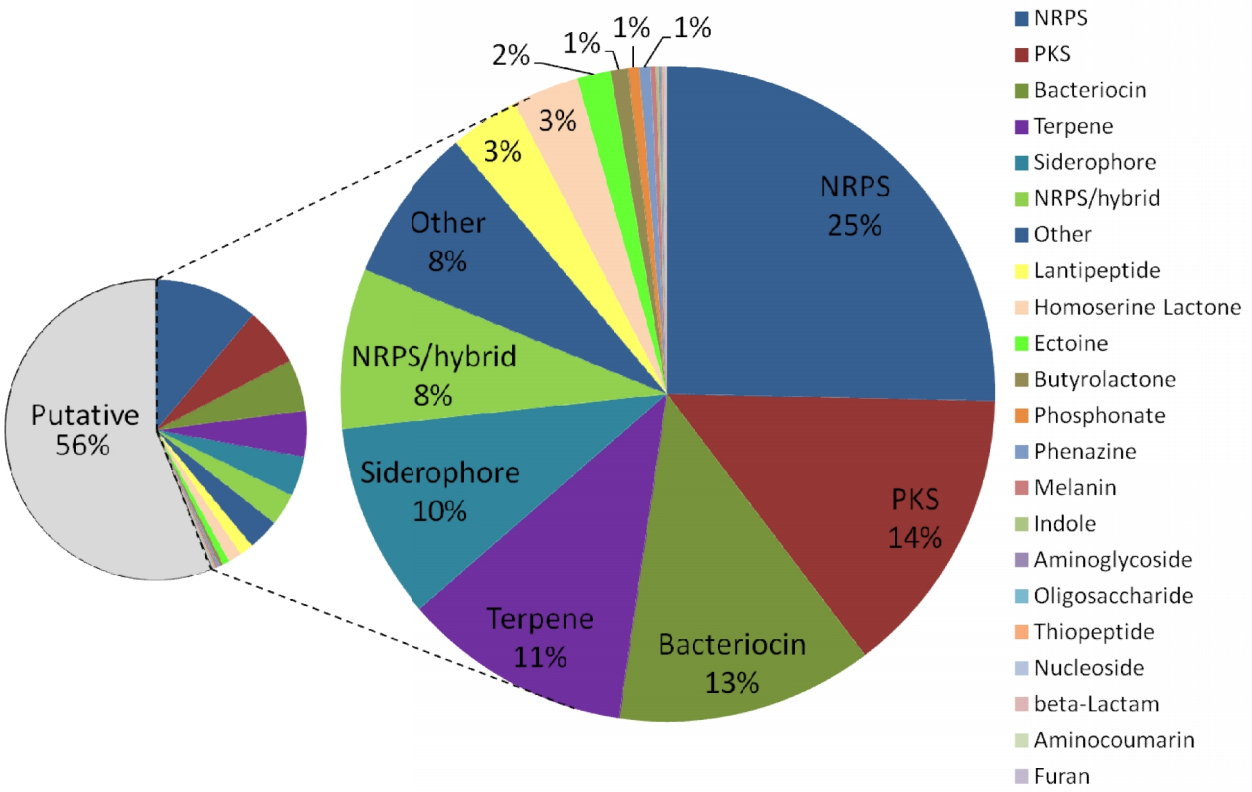


Figure 1.5 Distribution of biosynthetic gene clusters in bacteria; acquired from IMG-ABC database. (19th July 2018, https://img.jgi.doe.gov/cgi-bin/abc/main.cgi?section=BiosyntheticStats&page=breakdownByBCType).

Although BGCs have been discovered for thousands of bacterial strains, the diversity and similarity of BGCs are of importance for natural product discovery. To date, 1,500 BGCs have been experimentally characterized for their biosynthetic pathways and their products (Medema 2018). However, the genes of many gene cluster families still await characterization and even with conservative assumptions, the total number of bacterial BGC families present in the biosphere may be estimated to be ~6,000 (Cimermancic et al. 2014). Therefore, a 'global map' of biosynthesis would enable BGCs to be systematically selected for characterization by searching for, e.g., biosynthetic novelty, presence in undermined taxa, or patterns of phylogenetic distribution that indicate

functional importance (Cimermancic et al. 2014). A global phylogenomic analysis of BGCs in 1,154 complete genome sequences revealed that most known families of secondary metabolites are unique to a small set of organisms. However, BGCs of O-antigens, capsular polysaccharides, carotenoids and NRPS-independent siderophores are taxonomically widespread, and represent the most important families of molecules produced by microbes (Figure 1.6) (Cimermancic et al. 2014). The largest family in the sequence databases are aryl polyene (APE) BGCs. APEs are widely distributed throughout the Gram-negative bacteria and analogous to Gram-positive carotenoids which protect the producer from oxidative stress (Cimermancic et al. 2014).

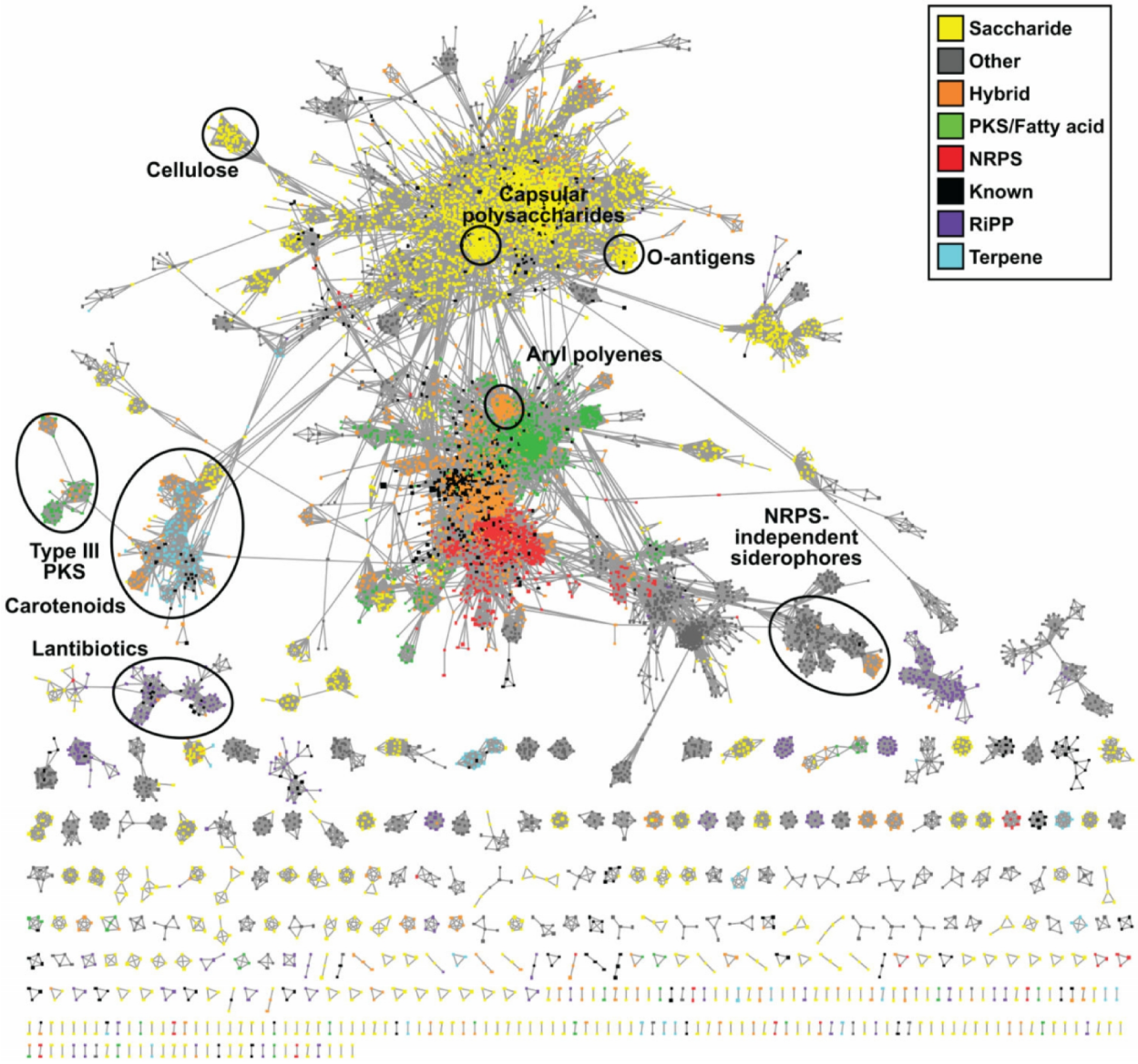


Figure 1.6 Similarity network of known and putative BGCs, detected in a set of 1,154 complete genome sequences from JGI-IMG database. 905 BGC families with distinct core genetic components were found within 11,456 identified BGCs. (Taken from Cimermancic et al. Cell. 2014 Jul 17;158(2):412-421).

1.6 Biosynthetic machinery of the main secondary metabolites in bacteria

BGCs are mainly involved in secondary metabolites production, as well as nutrient acquisition, toxin degradation, antimicrobial resistance, and vitamin biosynthesis (Cimermancic et al. 2014; Wisecaver et al. 2017; Slot 2017). The connection of microbial BGCs to small molecule metabolites that they encode is central to the

discovery and characterization of new metabolic pathways with ecological and pharmacological potential (Trautman and Crawford 2016). In the previous section, it was shown that the major BGCs are PKS, NRPS, terpene and RiPPs (more than 80 percent of identified BGCs). Here, the biosynthetic machinery of the main secondary metabolites in bacteria will be discussed briefly.

1.6.1 **Polyketide synthases (PKS)**

Polyketides are a large class of structurally diverse secondary metabolites, which are biosynthesized through the decarboxylative condensation of malonyl-CoA derived extender units (Robinson 1991). The enzymes that catalyze these condensations are referred to as polyketide synthases (PKSs) (Chan et al. 2009).

The vast diversity of polyketides is mainly derived from their substrate building blocks. Substrates functioning as starter units and extender units contribute significantly to the chemical complexity and structural diversity exhibited by this class of natural products (Ray and Moore 2016). Polyketides are built in a modular fashion that exists in most of the PKS BGCs, where each module is responsible for the incorporation of one building block into the growing product chain. Each module is further subdivided into domains, enzymatic units that are responsible for loading, condensation and further modification of the extender unit (Staunton and Weissman 2001).

PKS are classified into different types, based on their assembly lines architecture and mode of action (Hertweck 2009). The PKS assembly lines are divided into three major groups: type I, type II and type III. Type I and type II PKSs refer to that of the previously classified enzymes of fatty acid biosynthesis (Weissman 2009). Type I PKSs can generally be divided into two groups, modular and iterative. Type I applies to linearly arranged catalytic domains within large multifunctional enzymes, while in type II PKS enzymatic activities are typically present on individual proteins. Type III PKSs lack multiple catalytic domains, and a single enzyme condenses a starter unit with a series of extender units to generate a poly-β-keto chain (Chan et al. 2009). Besides the enzyme structures, PKSs are also classified based on the mechanism of synthesis.

Depending on whether a module is used only once or repeatedly, PKSs are termed as either modular (non-iterative) or iterative (Table 1.1) (Hertweck 2009).

Table 1.1 Overview on the different types of PKSs (adapted from Hertweck, 2009).

Type	Protein structure	Synthesis mechanism	Found in
PKS I	*Single* protein with *multiple modules*	Modular	Bacteria
PKS I	*Single* protein with *one module*	Iterative	Mainly fungi
PKS II	*Multiple* proteins each with *mono-functional* active site	Iterative	Bacteria
PKS III	*One* protein with *multiple modules*	Iterative	Plants, bacteria & fungi

1.6.2 Non-Ribosomal Peptide Synthetases (NRPS)

Nonribosomal peptides (NRPs) are a class of peptidic secondary metabolites that are not synthesized by ribosomes. Similar to polyketides, NRPs are biosynthesized in assembly-line fashion by multimodular enzymes called nonribosomal peptide synthetases (NRPSs). The dedicated protein modules are responsible for incorporating the amino acid building blocks into the growing NRP chain. Each module consists of several domains with defined functions (Niquille et al. 2018). Unlike ribosomes, NRPSs can accept numerous possible monomers in their assembly line (>520 monomers), which makes them highly diverse. Additional diversifications arise during the chain assembly, chain termination and post assembly-line tailoring reactions (Grünewald and Marahiel 2013). NRPSs are classified in three different groups: linear NRPS, iterative NRPS and non-linear NRPS (Mootz et al. 2002). Linear NRPS (type A) shows the same modular character as modular type I PKS, that means each module extends the peptide chain by one amino acid. Hence, the number of modules is equal to the number of amino acids in the corresponding peptide. In iterative NRPS (type B), modules or domains are used more than once, as observed for the enterobactin biosynthesis (Gehring et al. 1998). The non-linear NRPSs (type C) deviate from the canonical domain arrangement, present in type A and B NRPSs. Examples for type C NRPS are manifested in yersiniabactin, vibriobactin and bleomycin biosynthesis (Suo et al. 2001; Marshall et al. 2002; Shen et al. 2001).

1.6.3 **Terpene biosynthesis**

Terpenoids (also known as isoprenoids) are a group of secondary metabolites, that is essential in all living organisms. In most bacteria, the common precursors of all terpenoids are produced by the MEP (methylerythritol 4-phosphate) pathway (Figure 1.7). However, a few bacteria use the MVA (mevalonate) pathway instead of the MEP pathway, whereas others possess both pathways, and some lack both the MVA and the MEP pathways (Figure 1.7) (Pérez-Gil and Rodríguez-Concepción 2013).

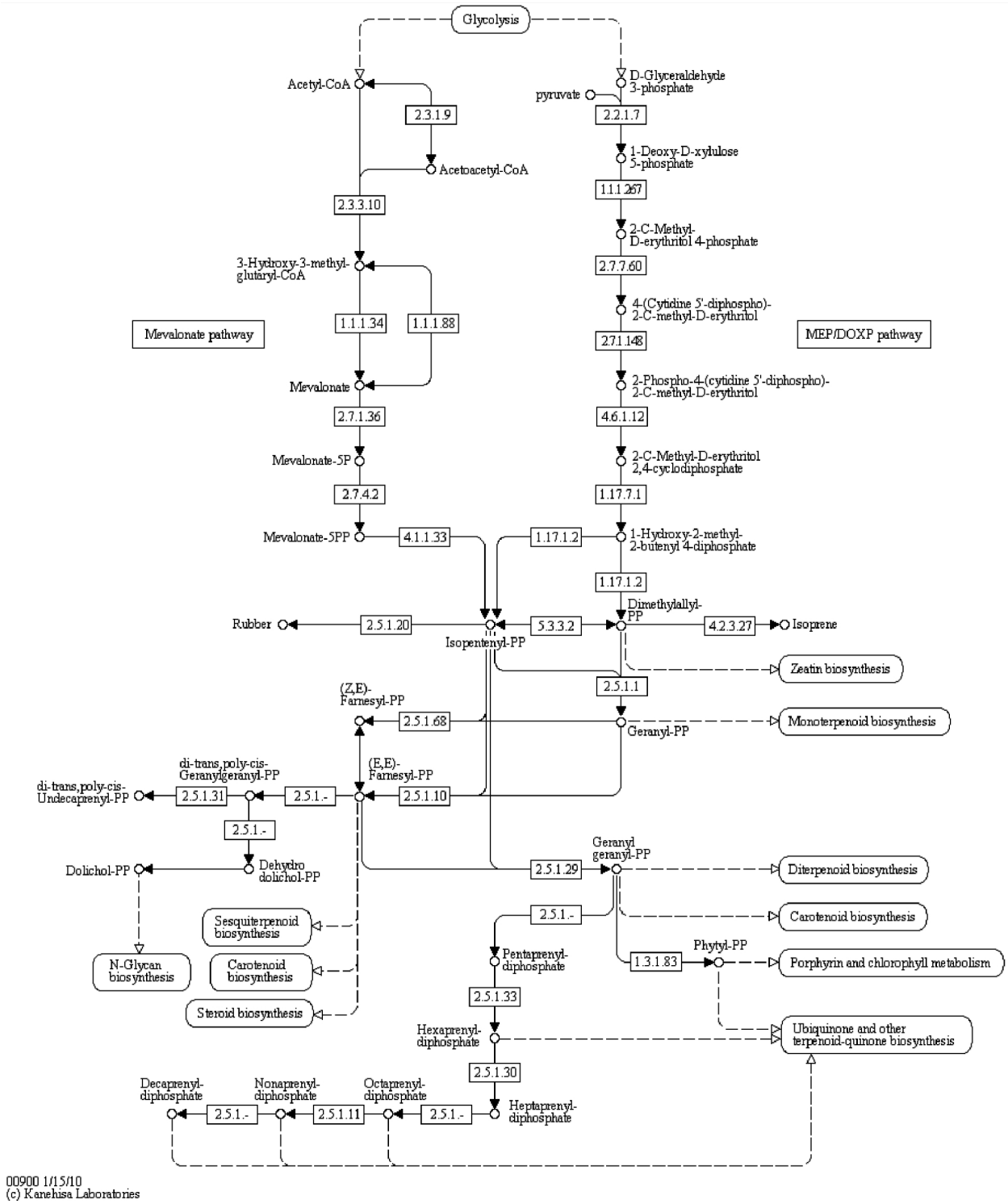


Figure 1.7 Biosynthetic pathways of terpenoid backbone structures in prokaryotes. Taken from KEGG map https://www.genome.jp/kegg-bin/show_pathway?map00900.

Numbers in the boxes represent the enzyme commission number (EC number), based on the chemical reactions they catalyze.

1.6.4 Bacteriocins and ribosomally synthesized and post-translationally modified peptide (RiPP) biosynthesis

Small molecules produced by the ribosomal machinery are natural products composed of the canonical 20 proteinogenic amino acids. However, they can be extensively post-translationally modified, and these modifications lead to products with many features resembling those of NRPs (McIntosh et al. 2009). Bacteriocins are peptidic toxins produced by bacteria to inhibit the growth of similar or closely related bacterial strain(s) (Cotter et al. 2013). If these peptides contain the modified amino acid lanthionine as part of their structure, they are called lantibiotics. However, efforts to reorganize the nomenclature of the family of RiPP natural products have led to the differentiation of lantipeptides from bacteriocins, based on their biosynthetic genes (Arnison et al. 2013). Historically, RiPPs have been subdivided based on, either the producing organisms (*e.g.* microcins produced by Gram-negative bacteria), or their biological activities (*e.g.* bacteriocins) (Arnison et al. 2013). Despite the fact that RiPPs cover a diverse range of structural classes, they all follow a simple biosynthetic logic: a precursor peptide consisting of an *N*-terminal leader sequence and a *C*-terminal core sequence, encoded by a single gene is translated, the leader sequence is removed by a series of transporters, peptidases or a combination of both, and the remaining active peptide moiety is further processed by other enzymes, often encoded by genes within close proximity to the precursor gene (Letzel et al. 2014). The core peptides are naturally hypervariable for subsets of the RiPP classes and many biosynthetic pathways utilize common enzymes for a subset of the post-translational modifications. The explosion in sequence information has revealed many different RiPP BGCs, including bacteriocins, lanthipeptides, thiopeptides, lassopeptides, etc. (Figure 1.8)

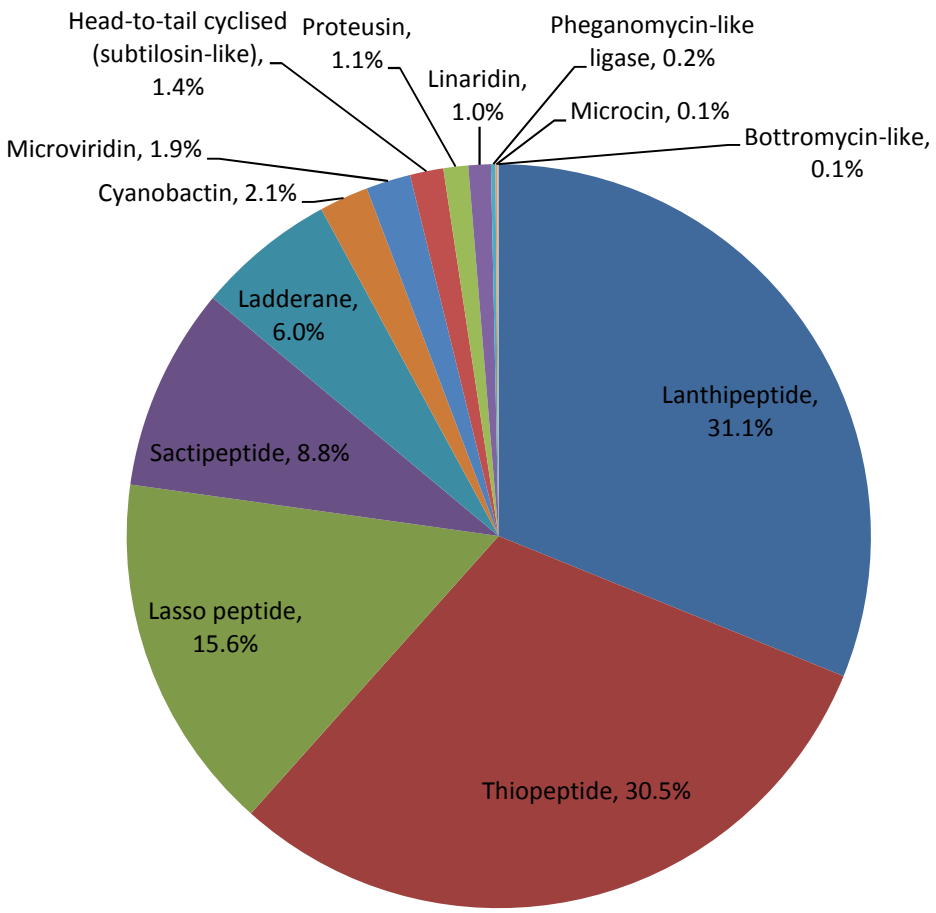


Figure 1.8 Distribution of RiPP BGCs in AntiSMASH database. (19th July 2018, https://antismash-db.secondarymetabolites.org/#!/query). In total 3600 BGCs were detected.

1.7 Myxobacteria (Deltaproteobacteria)

The phylum Proteobacteria is a large and extremely diverse group of Gram-negative bacteria divided into six bacterial groups (alpha, beta, gamma, delta, epsilon, and zeta). They include many pathogens, nitrogen fixers, and photosynthesizers (Williams and Kelly 2013). Myxobacteria belong to the Deltaproteobacteria, which are unicellular with rod-shaped vegetative cells and are able to glide over solid surfaces by secreting slime. They all show social behavior features by developing fruiting bodies and releasing myxospores in times of nutrient deficiency (Mohr et al. 2018). Myxospores are nonmotile, desiccation-resistant and can be disseminated by wind or animals and germinate under more favorable environmental conditions. Myxobacteria are united in

the order of Myxococcales, which is further divided into three suborders: Cystobacterineae, Sorangiineae, and Nannocystineae. They are known to possess the largest genomes amongst bacteria (9-16 Mbp) with a DNA of high GC content i.e. 66-72 mol% (Reichenbach, 1999).

1.7.1 **Marine Myxobacteria**

Myxobacteria have been first discovered from soil in 1809 and were thought to be occurring exclusively in terrestrial environments until Iizuka et al reported in 1998 the isolation of Myxobacteria from the marine environment (Dawid 2000; Iizuka et al. 1998). The reason for this late discovery was the difficulties that were encountered during the isolation and cultivation processes for marine Myxobacteria (Dávila-Céspedes et al. 2016). Fudou and Iizuka discovered Myxobacteria from the coasts of Japan, namely species from the genus *Enhygromyxa*, *Haliangium*, and *Plesiocystis,* which need seawater conditions in order to grow, and thus are considered as obligatory marine Myxobacteria (Iizuka et al. 1998; Fudou et al. 2002b; Iizuka et al. 2003a; Iizuka et al. 2003b). Later, by the group of König four closely related strains of *E. salina* were isolated from samples collected at the West Coast of the USA, German coasts and the Netherlands (Schäberle et al. 2010; Felder et al. 2013a).

The first Myxobacteria isolated from marine environments were thought to be members of halotolerant terrestrial Myxobacteria, whose myxospores had been washed into the ocean (Li et al. 2002). However, obligatory marine Myxobacteria are unable to grow in the absence of sodium chloride. They are phylogenetically distant and morphogenetically different from their terrestrial counterparts. They have been classified as a novel myxobacterial group (Zhang et al. 2005). 16S rDNA gene sequences from four deep-sea sediments collected at depths from 853 to 4675 m and a hydrothermal vent at a depth of 204 m were separated from terrestrial Myxobacteria at high levels of classification (Figure 1.9) (Jiang et al. 2010). This discovery indicated that marine Myxobacteria are phylogeographically separated from their terrestrial relatives, likely because of geographic separation and environmental selection.

On the other hand, halotolerant Myxobacteria like *Myxococcus fulvus* strain HW-1 can grow with and without sodium chloride in a wide range of salt water concentrations. The optimal seawater concentration for growth of HW-1 was 0 to 80% (salinity, 0.1 to 2.9%). The fruiting body structure was complete only on agar prepared with low concentrations of seawater or salts. The cells became shorter as the seawater concentration increased and rudimentary structures or even simple cell mounds appeared as the seawater concentration increased. In fact, halotolerant Myxobacteria have nearly the same 16S rRNA gene sequences as terrestrial Myxobacteria (Li et al. 2002), which was shown in the case *M. fulvus* (Figure 1.9). It is suggested that the halotolerant Myxobacteria are the result of degenerative adaptation of soil Myxobacteria to the marine environment, while halophilic Myxobacteria form a different evolutionary group that is indigenous to the ocean (Zhang et al. 2005).

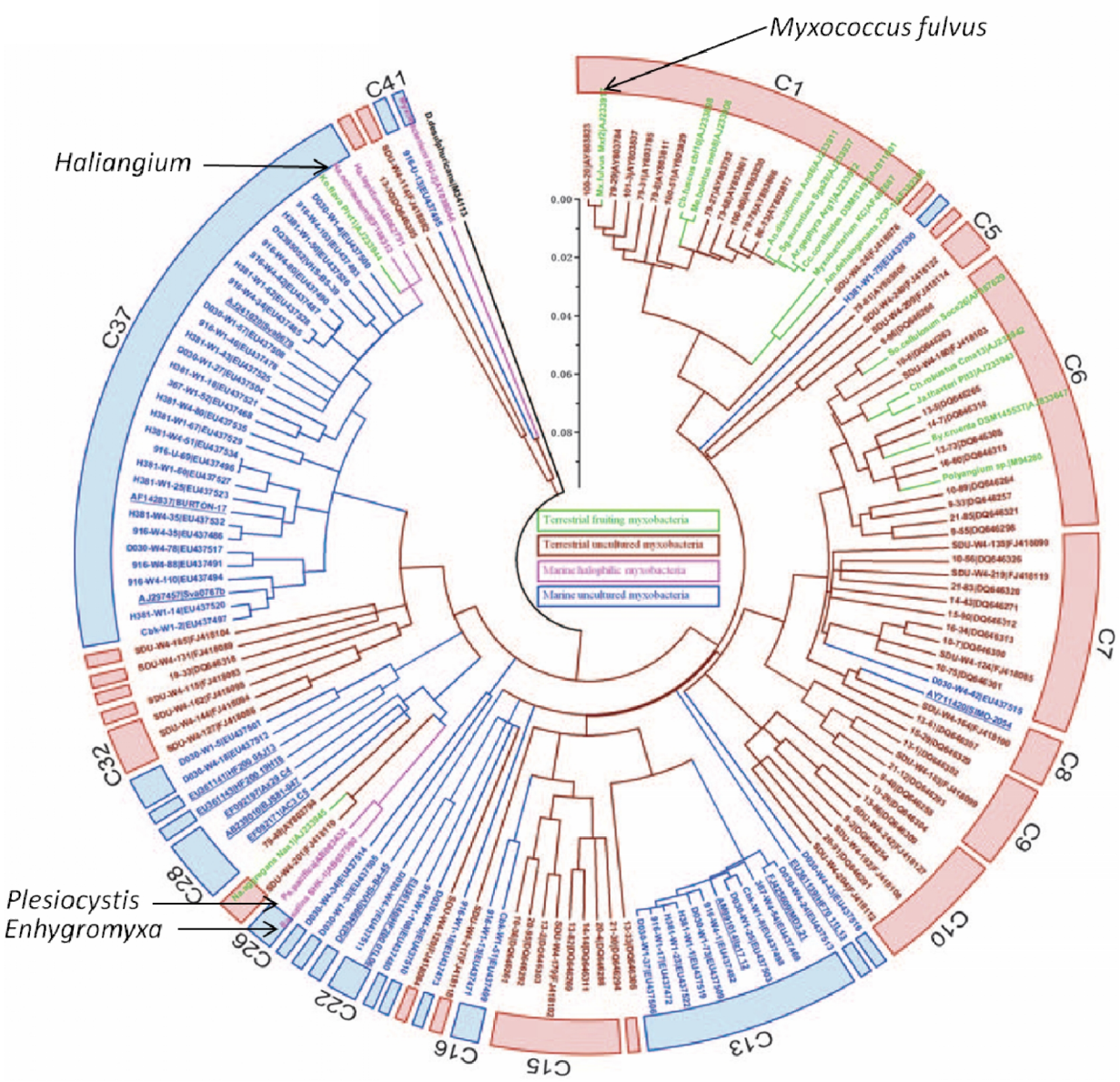


Figure 1.9 Comprehensive phylogenetic analyses of Myxobacteria using the available terrestrial and marine cultured and uncultured myxobacterial 16S rRNA gene sequences. C1 and C6 groups include the terrestrial Myxobacteria of suborder Cystobacterineae and Sorangineae, respectively. Halotolerant myxobacterium *Myxococcus fulvus* is located in C1. C26-C37 include most of the marine Myxobacteria . Marine species, *Plesiocystis* and *Enhygromyxa* are located in C26, and *Haliangium* is located in C37. (Jiang et al. 2010).

1.7.2 *Haliangium ochraceum*

Haliangium ochraceum sp. nov. (initially termed SMP-2, DSM 14365^T) and *H. tepidum* sp. nov. (initially termed SMP-10, DSM 14436^T) were isolated from seaweed and sea grass, respectively in Miura, Japan (Fudou et al. 2002b). The species were proposed to be of true marine origin according to their salt requirement for growth (2 to 3% NaCl (w/v)). The closest relative to both strains were the terrestrial Myxobacteria *Kofleria flava* (DSM 14601) with less than 94% identity in 16S rDNA sequences (Ivanova et al. 2010).

1.7.3 *Plesiocystis pacifica*

Two strains of *Plesiocystis pacifica* (SIR-1T and SHI-1) were retrieved from a sand sample of a Japanese coastal area (Iizuka et al. 2003a). These strains require 2–3% NaCl (w/v) and a pH of 7.4 for optimal growth on yeast medium with artificial seawater solution at 28 °C. The closest relative to *P. pacifica strain* was reported to be *Nannocystis exedens* DSM 71T, with 89.3% sequence identity to *P. pacifica* SIR-1T and 89.4% to *P. pacifica* SHI-1. Regarding GC content, *P. pacifica* strains, SIR-1T and SHI-1 were reported to have 69.3 and 70.0 mol %, respectively. Such a high GC content is a distinctive trait of all Myxobacteria. To date, no reports on biological testing of extracts or of any metabolites isolated from these organisms have been published.

1.7.4 *Enhygromyxa salina*

All *Enhygromyxa* species isolated to date are halophilic and considered as truly marine Myxobacteria. They are likely to have a world-wide distribution, as judged from the varied geographical occurrences (Felder et al. 2013a). *Enhygromyxa* strains were reported to have around 2% optimal NaCl (w/v) requirement for growth on yeast medium. The optimal temperature for growth was determined as 28–30 °C. In terms of gliding motility, fruiting body and myxospore formation, these bacteria show the typical features of terrestrial Myxobacteria (Figure 1.10).

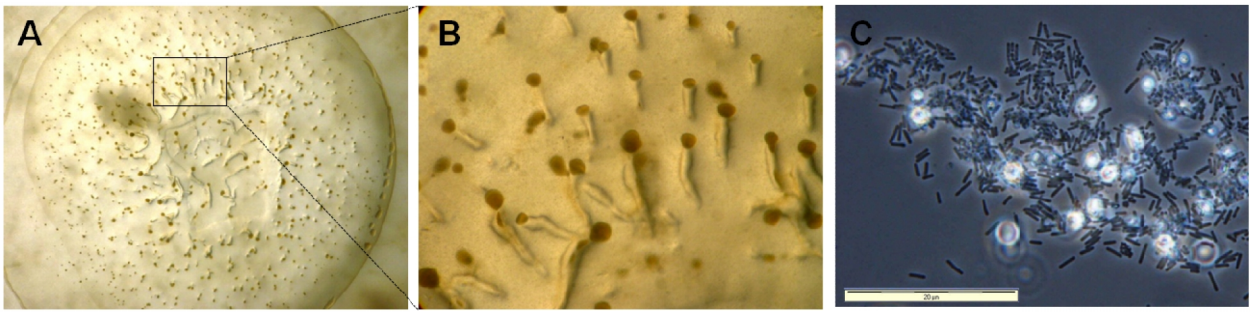


Figure 1.10 A and B: Fruiting bodies of *E. salina* SWB007 on ASW VY/4 medium. C: Microscopic image of the vegetative cell.

1.7.5 Secondary metabolites of Myxobacteria

Myxobacteria are not only fascinating because of their extraordinary lifestyle, they are also well known as producers of a variety of bioactive compounds such as polyketides, linear and cyclic peptides, and heterocyclic molecules. Terrestrial Myxobacteria have been well investigated over the past three decades, which resulted in more than 100 natural product scaffolds and approximately 600 structural derivatives with a broad range of biological activities (Plaza and Müller 2014). However, to date, only 10 obligatory marine myxobacterial strains have been isolated that can be classified into seven groups of natural products including enhygrolides, enhygromic acid, haliamide, haliangicins, salimabromide, salimyxins, and triterpenoid sterols (Tomura et al. 2017; Dávila-Céspedes et al. 2016; Sun et al. 2016b; Wei et al. 2016). The lack of more marine myxobacterial isolates and natural products is mainly due to the difficulties in isolation and cultivation of these bacteria (Felder et al. 2013b).

1.7.6 Biosynthetic potential of Myxobacteria for secondary metabolites

The immense diverse spectrum of myxobacterial compounds makes Myxobacteria an important source of novel classes of secondary metabolites (Gerth et al., 2003). Most of the structurally diverse myxobacterial natural products are synthesized by large multienzymes, i.e. PKSs, NRPSs or hybrids of both (Staunton and Weissman, 2001; Wenzel and Müller, 2009). Interestingly, more than half of the isolated myxobacterial

compounds contain elements of both, PKs and NRPs, and are therefore denoted as hybrid PK/NRP metabolites. Other bacterial producers of secondary metabolites, like the actinomycetes, synthesize prevalently pure PK or NRP compounds (Weissman and Müller, 2009). Indeed, AntiSMASH database (https://antismash-db.secondarymetabolites.org/#/query) shows that the NRPS and PKS BGCs are the majority of the BGCs in the order Myxococcales (Figure 1.11). From 178 NRPS BGCs identified, 114 BGCs are hybrids with PKS BGCs, mostly type I PKS (Figure 1.12A).

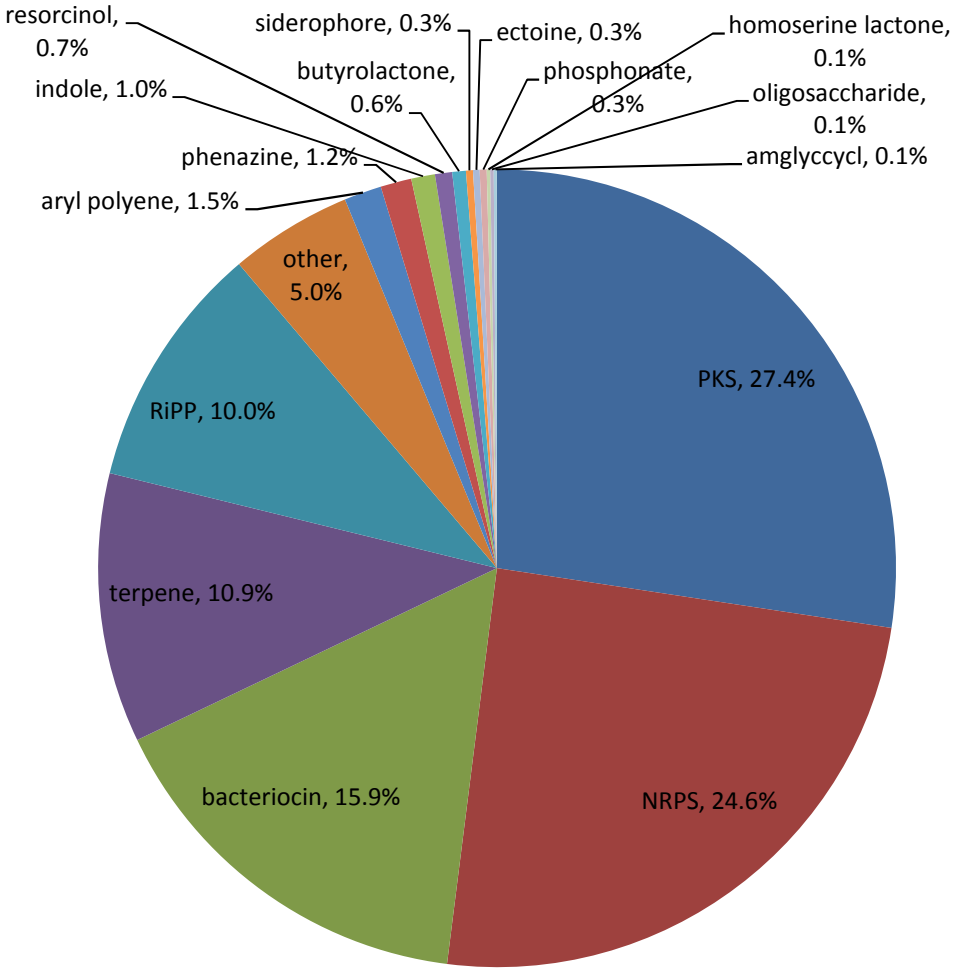


Figure 1.11 Distribution of BGCs of the order Myxococcales; AntiSMASH database. (19th July, 2018, https://antismash-db.secondarymetabolites.org/#!/query). In total 556 BGCswere detected. Family *Nannocystaceae* is not included. Polyketide BGCs include t1pks= type I PKS, t2pks= type II PKS, t3pks= type III PKS, transatpks= Trans-AT PKS, and otherks= other types of PKS cluster. RiPP (ribosomally synthesis and post-translationally modified peptide) BGCs include Lanthipeptide, Thiopeptide, Lasso peptide, Ladderane, and Microviridin, Proteusin clusters (Arnison et al. 2013).

Amglyccycl= Aminoglycoside/aminocyclitol cluster. Other BGCs containing a secondary metabolite- related protein that does not fit into any other category.

The majority of the PKS BGCs are type I PKS, but there are several type III PKS, trans-AT PKS, and other types of PKS BGCs detected (Figure 1.12B). RiPP and terpene BGCs are on the second majority of the BGCs in Myxococcales. The majority of the RiPP BGCs are bacteriocins and lantipeptides, which sometimes make hybrid BGCs with PKS and NRPS BGCs (Figure 1.12C). Terpene BGCs are mostly present as single BGCs, however in few cases make hybrids with PKS, NRPS, and RiPP BGCs (Figure 1.12D). On the other hand, some types of BGCs such as siderophores and ectoine are generally rare in terrestrial Myxococcales, excluding the family *Nannocystaceae* (Figure 1.11).

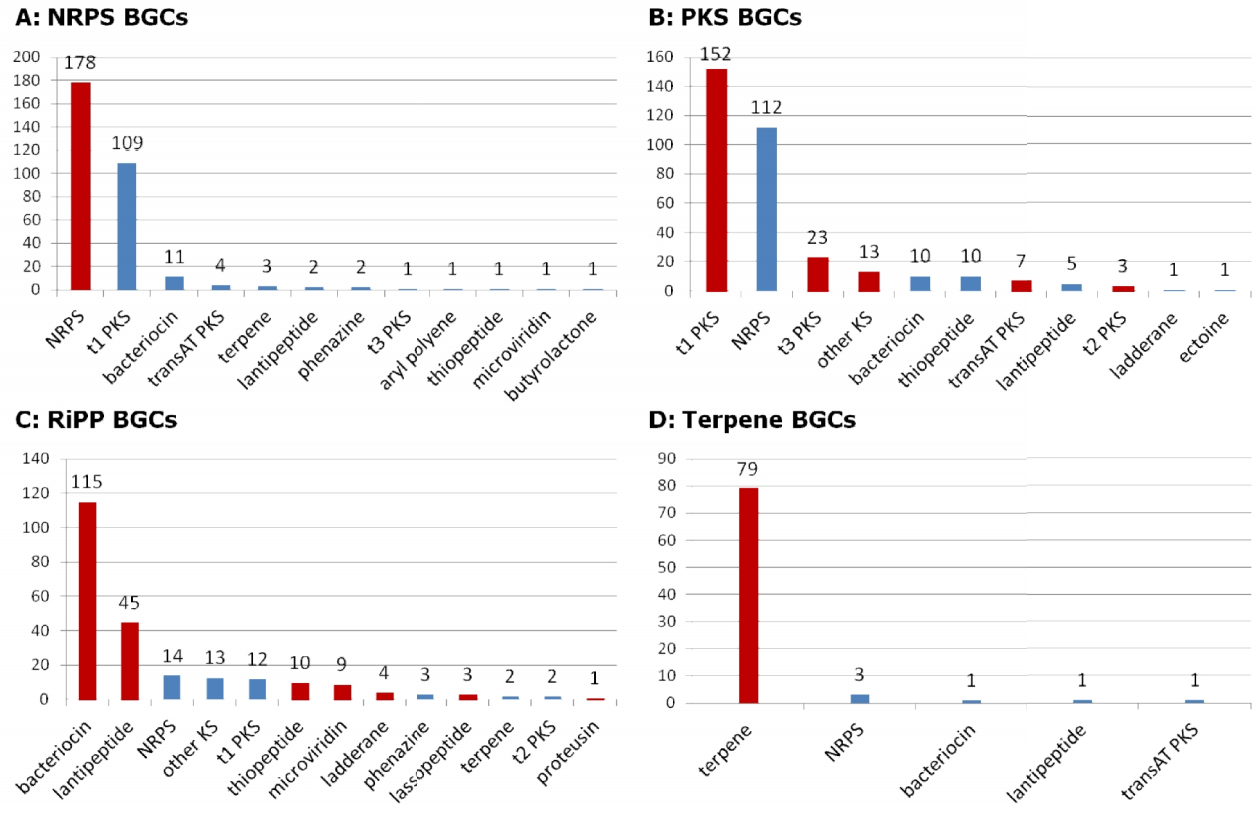


Figure 1.12 Total number of NRPS (A), PKS (B), RiPP (C), and Terpene (D) BGCs in Myxococcales and number of hybrid systems of each BGC family type; AntiSMASH database. (19th July, 2018, https://antismash-db.secondarymetabolites.org/#!/query). Main BGCs are shown in dark red and hybrid BGCs are shown in blue. t1pks= type I

PKS, t2pks= type II PKS, t3pks= type III PKS, transatpks= Trans-AT PKS, and otherks= other types of PKS cluster. Family Nannocystaceae is not included.

1.8 Rhodobacteraceae (Alphaproteobacteria)

Bacteria of the family Rhodobacteraceae are Gram negative and belong to the order Rhodobacterales, phylum Alphaproteobacteria. Rhodobacterales are especially widespread and abundant in marine environments. They are recognized as an abundant and metabolically versatile bacterial group in the world's oceans (Fu et al. 2013). The large diversity of physiological attributes of this family allowed them to exploit their environment and rapidly colonize surfaces, such as oysters shells. They may produce antibacterial components, which in turn shape the microbiome by inhibiting the growth of other bacteria (Arfken et al. 2017; Dang et al. 2008). Despite the assumed ecological importance of *Rhodobacterales*, their distribution patterns and their response to environmental drivers are poorly understood. However, it is known that distinct distribution patterns exist (Fu et al. 2013).

1.8.1 *Labrenzia* spp. (Rhodobacteraceae)

The genus *Labrenzia* currently harbors a couple of species, which were previously classified in the genus *Stappia* (Biebl et al. 2007). They are Gram-negative rods, and motile by means of one or several polarly inserted flagella. Colonies are white to creamy, but may become pink if incubated in the dark under appropriate conditions. They require NaCl for growth, and their optimum salinity range is 1–10 %. Optimum pH is 7.0–8.5. They may be able to reduce nitrate to nitrite or to N_2. Growth is chemoheterotrophic and non-fermentative under aerobic or anaerobic conditions. Their fatty acids comprise 16 : 0, 18 : 1 *ω*7 *c*, 18 : 0, 11-methyl 18 : 1 *ω*6 *t*, 20 : 1 *ω*7 *c* and the hydroxy fatty acids 3-OH 14 : 0, 3-OH 18 : 0 and 3-OH 20 : 0, all of which are amide-linked. DNA G+C content is 56–60 mol% (Biebl et al. 2007). Several type strains were described including *L. alexandrii, L. aggregata, L. marina, L. suaedae,* and *L. alba*.

1.8.2 Secondary metabolites of *Labrenzia* spp. and their special metabolic pathways

Secondary metabolites from *Labrenzia* spp. are almost unknown. A polyketide pederine-analog was discovered from one member of the genus *Labrenzia*, i.e. *Labrenzia* sp. PHM005, which showed cytotoxic activity against different cancer cell lines *in vitro* (Schleissner et al. 2017). Different unculturable symbiotic bacteria have been proposed as the real producers of the pederine-analog compounds, such as those found in insects, lichen, and marine sponges, and their trans-AT polyketide synthase gene clusters have been identified (Piel et al. 2004).

Sulfated exopolysaccharides (EPS) are produced by a bacterium designated as *Labrenzia* sp. PRIM-30, isolated from deep seawater of Cochin, India. The average molecular weight of the EPS was 269 kDa, contained 4.76% (w/w) sulfate groups and was composed of glucose, arabinose, galacturonic acid and mannose in a ratio 14.4:1.2:1:0.6. (P et al. 2014). EPSs are extracellular polymeric substances which modify the surface physicochemical properties making the surface suitable or unsuitable for recruitment of successive colonizers (Dang et al. 2008). These biopolymers produced by bacteria, have several industrial applications such as antioxidant and prophylactic agents in pharmaceutics (Guo et al. 2010), texture modifiers and emulsifiers in food industry (Duboc and Mollet 2001), as well as the active principle in formulations in cosmeceutical industry (Thibodeau 2005). They also have several biomedical applications such as antiproliferation of cancer cells, anticoagulation, and wound healing (Ruiz-Ruiz et al. 2011; Haroun-Bouhedja et al. 2000). EPSs can also be used in the bioremediation sector for enhanced oil recovery and removal of toxic contaminants (Satpute et al. 2010; Huang et al. 2012).

Labrenzia sp. BM1 showed quorum quenching activity by degrading a wide range of *N*-acylhomoserine lactones, namely *N*-(3-hexanoyl)-l-homoserine lactone (C6-HSL), *N*-(3-oxohexanoyl)-l-homoserine lactone (3-oxo-C6-HSL) and *N*-(3-hydroxyhexanoyl)-l-homoserine lactone (3-hydroxy-C6-HSL) (Ghani et al. 2014).

L. aggregata LZB033 produces dimethylsulfoniopropionate (DMSP), which is one of the Earth's most abundant organosulfur molecules, a signalling molecule, a key nutrient for marine microorganisms and the major precursor for gaseous dimethyl sulfide (DMS) (Curson et al. 2017).

Labrenzia spp. and their close relatives are reported to be marine denitrifiers and habitual colonizers of the pseudobenthic environment provided by *Trichodesmium* spp.. *L. alexandrii* DFL 11 was reported to form the carbon biopolymer polyhydroxyalkanoate (PHA). Aerobic anoxygenic phototrophic bacteria (AAPB) such as *Labrenzia* spp. are unique players in carbon cycling in the ocean. The cellular carbon storage is an important mechanism regulating the nutrition status of AAPB. These reports reflect the potential of these organisms to produce bioactive substances for biotechnological applications.

1.8.3 Biosynthetic potential of *Labrenzia* species for secondary metabolites

Todate, no BGC was described from the genus *Labrenzia*. However, the IMG/ABC database includes of 16 *Labrenzia* genomes, and predicted several BGCs from each genome. NRPS, terpene, bacteriocin, and type I PKS gene clusters are common and conserved in this group of bacteria. Hybrid NRPS/trans-AT PKS and other BGCs are also predicted in some of the genomes (Figure 1.13).

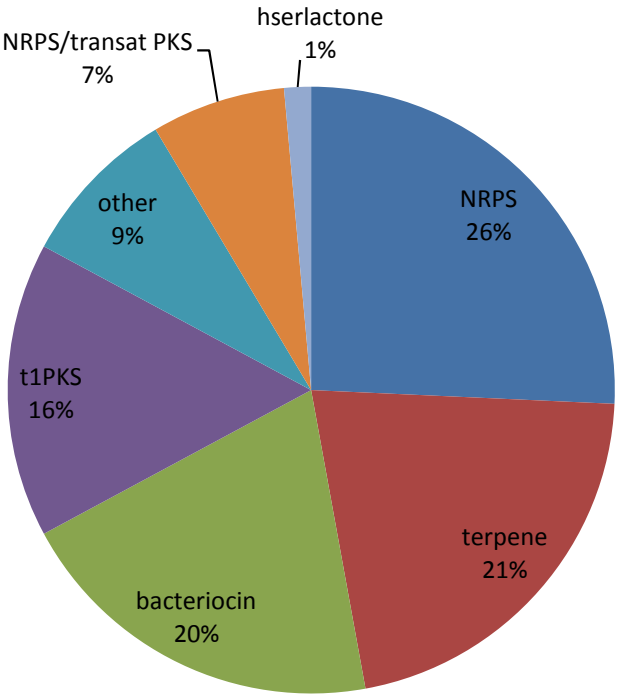


Figure 1.13 Distribution of BGCs of the genus *Labrenzia* (IMG/ABC database). 16 genomes and 70 BGCs in total. t1pks= type I PKS and transatpks= Trans-AT PKS. Other= cluster containing a secondary metabolite-related protein that does not fit into any other category. Hserlactone= Homoserine lactone cluster.

2 Scope of the study

Obligatory marine bacteria are hardly investigated concerning their bioactive metabolites. Using an approach based on genomics, including a range of bioinformatic tools, the metabolic potential of such microorganisms was intended to be investigated. Additionally, mass spectrometric technique and molecular networking were envisaged to serve in natural products discovery from marine bacteria.

The focus of the current study was i) on rare marine myxobacteria (strains of *Enhygromyxa salina*, *Plesiocystis pacifica* and *Haliangium ochraceum),* and ii) on bacteria of the genus Labrenzia, which were suggested to play an ecological role as protective agent against the roseovarius oyster disease.

The following questions were to be addressed:

What are the genomic and metabolomic similarities and differences between marine myxobacteria, especially with regard to the biosynthesis of specialized metabolites?

What are the strategies and which biosynthetic pathways of marine myxobacteria (Enhygromyxa salina SWB007 and Plesiocystis pacifica DSM 14875) serve the production of osmolytes to cope with the osmotic pressure of the saline environment?

Which biosynthetic gene clusters for bioactive metabolites are in the genomes of Labrenzia spp., and can they explain the production of antimicrobial compounds by Labrenzia sp. 011?

3 Materials and Methods

3.1 Materials

3.1.1 Chemicals and solvents

Table 3.1 Chemical substances and solutions used in this work

Substance	Manufacturer
1-aminoadamantane (ADAM)	Merck KGaA (Darmstadt, Germany)
9-fluorenylmethoxycarbonyl (FMOC)	Merck KGaA (Darmstadt, Germany)
Aceton-d 99,8%	Deutero (Kastellaun, Germany)
Acetone	Roth Chemie GmbH (Karlsruhe, Germany)
Acetonitrile	VWR International GmbH (Darmstadt, Germany)
Agar	Fulka Chemie GmbH (Buchs,Switzerland)
Amberlite® XAD16N	Dow Chemical Company (USA)
Boric acid	Roth Chemie GmbH (Karlsruhe, Germany)
$CaCl_2$ x 2 H_2O	Merck KGaA (Darmstadt, Germany)
Celite® 545	Imerys Minerals California, Inc. (USA)
Chloroform	Roth Chemie GmbH (Karlsruhe, Germany)
$CoCl_2$	Roth Chemie GmbH (Karlsruhe, Germany)
$CuSO_4$	Merck KGaA (Darmstadt, Germany)
Dichloromethane (DCM)	Roth Chemie GmbH (Karlsruhe, Germany)
DMSO	Roth Chemie GmbH (Karlsruhe, Germany)
dNTP	Promega GmbH (Mannheim, Germany)
Ethanol 99,8% p.a.	Roth Chemie GmbH (Karlsruhe, Germany)
Ethidium bromide	Roth Chemie GmbH (Karlsruhe, Germany)
Ethyl acetate	Roth Chemie GmbH (Karlsruhe, Germany)

Gel Loading Dye	Fermentas GmbH (St. Leon Rot, Germany)
Glycerol	Roth Chemie GmbH (Karlsruhe, Germany)
GoTaq®Flexi Buffer (10x)	Promega GmbH (Mannheim, Germany)
Hexane	Roth Chemie GmbH (Karlsruhe, Germany)
Imidazole	Roth Chemie GmbH (Karlsruhe, Germany)
Isoamyl alcohol	Roth Chemie GmbH (Karlsruhe, Germany)
Isopropanol	Roth Chemie GmbH (Karlsruhe, Germany)
KI	Roth Chemie GmbH (Karlsruhe, Germany)
LiCl	Roth Chemie GmbH (Karlsruhe, Germany)
$MgCl_2$ x 6 H_2O	Merck KGaA (Darmstadt, Germany)
$MgSO_4$ x 7 H_2O	Merck KGaA (Darmstadt, Germany)
$MnCl_2$ x 4 H_2O	Merck KGaA (Darmstadt, Germany)
N,N,N',N'-tetraethylendiamin (TEMED)	Roth Chemie GmbH (Karlsruhe, Germany)
Na_2-EDTA	Roth Chemie GmbH (Karlsruhe, Germany)
Na_2MoO_4 x 2 H_2O	Merck KGaA (Darmstadt, Germany)
Na_2SO_4	Roth Chemie GmbH (Karlsruhe, Germany)
NaCl	Merck KGaA (Darmstadt, Germany)
$NaHCO_3$	Roth Chemie GmbH (Karlsruhe, Germany)
NaOH	Merck KGaA (Darmstadt, Germany)
Ni-NTA agarose	Qiagen GmbH (Hilden, Germany)
Norvalin	Merck KGaA (Darmstadt, Germany)
peqGOLD Agarose	PEQLAB Biotechnologie GmbH (Erlangen, Germany)
Petroleum ether	Roth Chemie GmbH (Karlsruhe, Germany)
Potassium bromide	Roth Chemie GmbH (Karlsruhe, Germany)
Potassium Cloride	Roth Chemie GmbH (Karlsruhe, Germany)

SDS	Roth Chemie GmbH (Karlsruhe, Germany)
Sepabeads® SP207	Supelco (USA)
Silica gel 60	Merck KGaA (Darmstadt, Germany)
Sodium acetate	Merck KGaA (Darmstadt, Germany)
$SrCl_2$ x 6 H_2O	Merck KGaA (Darmstadt, Germany)
Tetrahydrofuran	Merck KGaA (Darmstadt, Germany)
Tris	Roth Chemie GmbH (Karlsruhe, Germany)
Tris-HCl	Roth Chemie GmbH (Karlsruhe, Germany)
Tryptone/Peptone from Caseine	Roth Chemie GmbH (Karlsruhe, Germany)
Yeast extract	Fluka Chemie GmbH (Buchs, Switzerland)
$ZnCl_2$	Merck KGaA (Darmstadt, Germany)

Table 3.2 Technical equipment and other material used in this project

Equipment	**Manufacturer**
Amicon Ultra centrifugal filters	Millipore GmbH (Schwalbach, Germany)
Autoclave	Varioklav®, H+P Labortechnik AG (Oberschleißheim, Germany)
Branson Sonifier 250	G. Heinemann Ultraschall- und Labortechnik (Schwäbisch Gmünd, Germany)
Biometra T3000 Thermocycler	Biometra GmbH (Göttingen, Germany)
BioRad PowerPac™ 300	Bio-Rad Laboratories GmbH (Hercules, U.S.A.)
Centrifuge Heraeus Biofuge fresco	Thermo Fisher Scientific (Waltham, U.S.A.)
Centrifuge Heraeus Contifuge Stratos	Thermo Fisher Scientific (Waltham, U.S.A.)
Centrifuge Heraeus Fresco 17	Thermo Fisher Scientific (Waltham, U.S.A.)
Centrifuge tubes (15/50 ml)	TPP AG (Trasadingen, Germany)

CopyControl™ Induction solution	Epicentre (Madison, U.S.A.)
Eppendorf Centrifuge 5415 D	Eppendorf (Hamburg, Germany)
Eppendorf tubes 0.5, 1.5, 2 ml	Eppendorf (Hamburg, Germany)
Centrifuge tubes 15/50 ml	TPP AG (Trasadingen, Germany)
Gel chambers Horizon 58 and Horizon 11.14	Life technologies (Karlsruhe, Germany)
Incubator	Memmert GmbH + Co. KG (Schwalbach, Germany)
Inolab pH meter	WTW GmbH (Weilheim, Germany)
Intas iX Imager	Intas Science Imaging Instruments GmbH (Göttingen, Germany)
Kodak DC290	Kodak GmbH (Stuttgart, Germany)
Laminar Airflow Clean Bench BSB 4A (Hera Safe, Class II)	Heraeus (Hanau, Germany)
Magnetic stirrer (IKA® RH basic)	IKA® Werke GmbH & Co. KG (Staufen, Germany)
Milli-Q® Water System	Millipore (Eschborn, Germany)
Multitron incubation shaker	IKA® Werke GmbH & Co. KG (Staufen, Germany)
Nalgene cryogenic vials	Nalgene Nunc International (Rochester, U.S.A.)
Parafilm®	Pechiney Plastic Packaging Company (Chicago, U.S.A.)
Scale (Satorius BL 3100)	Satorius AG (Göttingen, Germany)
Scale (Satorius BP 221S)	Satorius AG (Göttingen, Germany)
Sterile filter (0.2 µm)	Renner GmbH (Dannstadt, Germany)
Thermomixer Eppendorf	Eppendorf (Hamburg, Germany)
Transferpette®-8	Brand GmbH + Co. KG (Wertheim, Germany)
UV mini 1240 UV/VIS spectrometer	Shimadzu (Kyoto, Japan)
UV cuvettes	Ratiolab GmbH (Dreieich, Germany)
Water bath (Haake DC10)	Thermo Haake GmbH (Karlsruhe, Germany)

3.1.2 Enzymes

The enzymes used in the scope of this study are listed in **Table 3.3**. Appropriate enzyme reaction buffers, which provide optimal reaction conditions as for polymerase, restriction and ligation were purchased together with the enzymes.

Table 3.3 Enzymes used in the scope of this study

Enzyme	Manufacturer
GoTaq® Flexi DNA Polymerase (5 u/µl)	Promega (Mannheim, Germany)
Lysozyme	Roth (Karlsruhe, Germany)
Proteinase K	Roth (Karlsruhe, Germany)
Restriction enzymes	Fermentas GmbH (St. Leon Rot, Germany), Promega (Mannheim, Germany)
RNase (DNase free)	Roth Promega (Mannheim, Germany)
Fast-Link™ DNA Ligase	Epicentre (Madison, U.S.A.)
T_4-DNA-Ligase	Fermentas GmbH (St. Leon Rot, Germany)

All enzymes were applied following the respective manufacturer's recommendations for use. Lysozyme and Proteinase K were used as stock solutions with concentrations of 100 mg/ml and 20 mg/ml, respectively. Restriction enzymes were purchased together with the appropriate reaction buffers and were used according to the provided company's protocols.

3.1.3 Molecular weight marker

The following DNA standards listed in **Table 3.4** were used for gel electrophoresis.

Table 3.4 Molecular weight marker applied for size estimation

Molecular weight marker	Manufacturer
Gene Ruler™ DNA Ladder Mix	Fermentas GmbH (St. Leon Rot, Germany)
Lambda Mix Marker	Fermentas GmbH (St. Leon Rot, Germany)
PageRuler Unstained Protein Ladder	Fermentas GmbH (St. Leon Rot, Germany)

3.1.4 Molecular biological kits

Commercial molecular kits that were utilized in this work are listed in **Table 3.5**. The application of the kits was performed according to the manufacturer's instruction manuals.

Table 3.5 Molecular biological kits

Molecular biological kit	Manufacturer
QIAquick PCR Purification Kit	Qiagen GmbH (Hilden, Germany)
pGEM®-T Vector System I	Promega (Mannheim, Germany)
PureYield™Miniprep System	Promega (Mannheim, Germany)
QIAGEN Plasmid Midi Kit	Qiagen GmbH (Hilden, Germany)
QIAquick Gel Extraction Kit	Qiagen GmbH (Hilden, Germany)
QIAprep®Spin Miniprep Kit	Qiagen GmbH (Hilden, Germany)
DNA Clean and Concenttator™-5	Zymo Research Europe (Freiburg, Germany)
Wizard® SV Gel and PCR Clean-Up System	Promega (Mannheim, Germany)
GenElute™ Bacterial Genomic DNA Kit	Sigma-Aldrich (USA)

3.1.5 Water

For the preparation of culture media, demineralized water was provided by a reverse osmosis system (IMB, Germany). A Milli-Q® Academic Water Purification System (Millipore GmbH, Germany) was used to generate ultra-pure water prepared

from demineralized water. Milli-Q® water was used for all applications, if not specified otherwise. For PCR application, autoclaved Milli-Q® water was applied.

3.1.6 Culture Media

Culture media were prepared with demineralized water prior to steam sterilization. The following table shows the composition of the used media (**Table 3.6**).

Table 3.6 List of media and their composition

ASW VY/4:	
25ml yeast suspension(10%)	
add up to 1L 75% artificial sea water (ASW)	
	adjust to pH=7.5
	after sterilization add:
	1ml/L trace element solution
	1ml/L vitamin B_{12} solution

ASW-Coli medium:	
30ml *E. coli* BKA13 suspension (10%)	
add up to 1L 75% artificial sea water (ASW)	
	adjust to pH=7.5
	after sterilization add:
	1ml/L trace element solution
	1ml/L vitamin B_{12} solution

Minimal medium MM63:
1.98 g $(NH_4)_2SO_4$
13.61 g KH_2PO_4

4.21 g KOH	
0.25 g $MgSO_4$	
	Add 700 mL Milli-Q water
Add 1 mL $FeSO_4$ solution	(0.0011 g / mL FeSO4)
	adjust to pH=7.2
Add 5 g NaCl	
	Add up to 900 mL with Milli-Q water
	After sterilization and cooling down to room temperature add:
	100 mL autoclaved glucose solution (5 g / 100 mL)

Luria-Bertani (LB)-Medium	
10 g tryptone	
5 g yeast extract	
10 g NaCl	
ad 1 L H_2O	adjust to pH=7.5

LB-Agar	
10 g tryptone	
5 g yeast extract	
10 g NaCl	
15 g Agar	
ad 1 L H_2O	adjust to pH=7.5

Trace element solution	
20 mg $ZnCl_2$	5 mg LiCl
100 mg $MnCl_2$ x 4 H_2O	20 mg KBr
10 mg H_3BO_3	20 mg KI
10 mg $CuSO_4$	10 mg Na_2MoO_4 x 2 H_2O
20 mg $CoCl_2$	5.2 g EDTA-Na_2 x 2 H_2O
5 mg $SnCl_2$ x 2 H_2O	add 1 L H_2O; solution was sterile filtrated

Cyanocobalamine solution	
50 mg cyanocobalamine	
add 100 ml H_2O	solution was sterile filtrated

3.1.7 Antibiotics

The antibiotics listed in **Table 3.7** were used as additives in culture media for the purpose of clone selection. The antibiotics were used as antibiotic stock solutions and were solved either in autoclaved Milli-Q water.

Table 3.7 List of antibiotics used as additive in culture media

Antibiotics	**Manufacturer**
Ampicillin	Roth Chemie GmbH (Karlsruhe, Germany)
Kanamycin	Sigma-Aldrich Co. LLC (St. Louis, MO, USA)

3.1.8 Buffers and solutions

All buffers and solutions (**Table 3.8**) were prepared with Milli-Q water (3.1.5). Stock solutions not containing any organic solvents or thermo labile components were usually steam sterilized prior to use.

Table 3.8 Buffers and solutions

Buffers for DNA extraction:	
TE buffer	**SET buffer**
10 mM Tris-HCl 1 mM EDTA pH 8.0	75 mM NaCl, 25 mM EDTA 10 mM Tris-HCl pH 7.5
RNase A stock solution (DNase-free)	
50 mg lyophilised RNase A 5 ml Tris-HCl (10 mM, pH 7.5) 15 mM NaCl storage at –20 °C	
Buffers for plasmid purification:	
Buffer P1	**Buffer P2**
50 mM Tris-HCl 10 mM EDTA 100 µg/mL RNase A pH 8	200 mM NaOH 1% SDS
Buffer P3	**Buffer QBT (Equilibration buffer)**
3M potassium acetate pH 5.5	750 mM NaCl 50 mM MOPS (pH 7.0) →including 15 % isopropanol and 15 ml of 10% Triton X-100 solution (v/v)
Buffer QC (Wash buffer)	**Buffer QF (Elution Buffer)**
1 M NaCl 50 mM MOPS (pH 7.0) →including 15 % isopropanol	1.25 M NaCl 4M Tris (pH 8.5) →including 15 % isopropanol
Buffers for gel electrophoresis:	
10 x TBE buffer	**Modified 50 x TAE buffer**
0.89 M Tris base 20 mM EDTA	2 M Tris-acetate 50 mM EDTA (pH 8)

0.87 M H_3BO_3
purified water ad 1000 ml
6 x Gel loading buffer
40 g sucrose
0.25 g bromophenol blue
ad 100 ml purified water

3.1.9 Software and Databases

Basic Local Alignment Search Tool (BLAST) is provided by the National Centre for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/pubmed). This software was applied to analyze nucleotide data for sequence similarities using blastx (nucleotide sequence is translated and compared to the amino acid sequences database). For further analysis of the amino acid sequences blastp (protein database using a protein query) was used. In order to analyze 16S rDNA, the database was searched using blastn (nucleotide query vs. nucleotide databases).

Entrez Nucleotide Database is maintained by the National Centre for Biotechnology Information (NCBI) (http://www.ncbi.nlm.nih.gov/gene). This source is linked to numerous databases, e.g. GenBank, NCBI Reference Sequences (RefSeq), and the Protein Data Bank of the Research Collaboratory for Structural Bioinformatics (RCSB PDB). Therefore, it lends itself for detailed analysis of nucleotide and amino acid sequences and putative anticipation on their possible functions in secondary metabolites biosynthesis.

Global Natural Products Social Molecular Networking (GNPS) web-platform provides public data set deposition and/or retrieval through the Mass Spectrometry Interactive Virtual Environment (MassIVE) data repository. The GNPS analysis infrastructure further enables online dereplication, automated molecular networking analysis, and crowdsourced MS/MS spectrum curation.

SPAdes – St. Petersburg genome assembler – is an assembly toolkit containing various assembly pipelines. SPAdes version 3.10.1 was released under GPLv2 on March 1, 2017.

CheckM provides a set of tools for assessing the quality of genomes recovered from isolates, single cells, or metagenomes. It provides robust estimates of genome completeness and contamination by using collocated sets of genes that are ubiquitous and single-copy within a phylogenetic lineage.

Trimmomatic performs a variety of useful trimming tasks for illumina paired-end and single ended data. The selection of trimming steps and their associated parameters are supplied on the command line.

QUAST stands for Quality Assessment Tool. The tool evaluates genome assemblies by computing various metrics. This document provides instructions for the general QUAST tool for genome assemblies, MetaQUAST, the extension for metagenomic datasets, and Icarus, interactive visualizer for these tools.

Mauve Contig Mover (MCM) can be used to order a draft genome relative to a related reference genome. The Mauve Contig Mover can ease a comparative study between draft and reference sequences by ordering draft contigs according to the reference genome.

Rapid Annotation using Subsystem Technology (RAST) is a fully-automated service for annotating complete or nearly complete bacterial and archaeal genomes (http://rast.nmpdr.org/). It provides high quality genome annotations for these genomes across the whole phylogenetic tree.

EDGAR is designed to automatically perform genome comparisons in a high throughput approach. EDGAR provides novel analysis features and greatly simplifies the comparative analysis of related genomes. Visualization features, like synteny plots or Venn diagrams, are offered to the scientific community through a web-based and thus platform independent user interface where the precomputed data sets can be browsed.

GGDC web service reports digital DDH for a universal and accurate delineation of prokaryotic (sub-)species without inheriting the pitfalls of classic DDH, and also calculates differences in genomic G+C content.

CGView Comparison Tool (CCT) is a package for visually comparing bacterial, plasmid, chloroplast, or mitochondrial sequences of interest to existing genomes or sequence collections.

Biosynthetic Genes Similarity Clustering and Prospecting Engine (BiG-SCAPE). This program defines a distance metric between Gene Clusters using a combination of three indices (Jaccard Index of domain types, Domain Sequence Similarity the Adjacency Index).

Antibiotic Resistant Target Seeker (ARTS) is provided by Ziemert lab in Tübingen and explores secondary metabolite gene clusters and uncover new resistant targets. It also screens secondary metabolite gene clusters for known targets, find unique metabolism and resistance factors in SM gene clusters a, and discover leads to novel antibiotic targets.

Universal Protein Resource (UniProt) is a comprehensive resource for protein sequence and annotation data. The UniProt databases are the UniProt Knowledgebase (UniProtKB), the UniProt Reference Clusters (UniRef), and the UniProt Archive (UniParc).

Clustal Omega is the latest addition to the Clustal family which is located in UniProt database. It offers a significant increase in scalability over previous versions, allowing hundreds of thousands of sequences to be aligned in only a few hours.

MultiGeneBlast is an open source tool for identification of homologs of multigene modules such as operons and gene clusters. It is based on a reformatting of the FASTA headers of NCBI GenBank protein entries, using which it can track down their source nucleotide and coordinates.

Clone Manager 9 is a purchased program with a set of tools for enzyme operations, cloning simulation, graphic map drawing, primer design and analysis and

sequence alignments. Clone Manager 9 was mainly used for primer design and the simulation of cloning procedures.

AntiSMASH (antibiotics & Secondary Metabolite Analysis Shell) is a free online tool that allows the rapid genome-wide identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genomes (http://antismash.secondarymetabolites.org/). It integrates and cross-links with a large number of in silico secondary metabolite analysis tools that have been published earlier. AntiSMASH is powered by several open source tools: NCBI BLAST+, HMMer 3, Muscle 3, Glimmer 3, FastTree, TreeGraph 2, Indigo-depict, PySVG and JQuery SVG.

Kodak 1D software version 3.5.4 and iX Imager software: These programs were provided together with the respective gel documentation system (Kodak Scientific Imaging Systems and Intas Gel iX imager) and were applied to edit digital photos of electrophoresis.

3.1.10 Strains and isolates

The marine myxobacterial strains *Enhygromyxa salina* SWB005, SWB006 and SWB007 were obtained from the strain collection of the Institute for Pharmaceutical Biology, University of Bonn, Bonn, Germany. These strains have been isolated from marine sediments, which originated from beach area of Santa Barbara, U.S. (*E. salina* SWB005), of Borkum, Germany (*E. salina* SWB006) and Prerow, Germany (*E. salina* SWB007)(Schäberle et al. 2010; Felder et al. 2013a). *Enhygromyxa salina* DSM 15201 (formerly named *E. salina* SMP-6) and *Plesiocystis pacifica* DSM 14875 (type strain, formerly named *P. pacifica* SIR-1) were obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ). Those strains have been isolated from coastal sands (*E. salina* DSM 15201) and semi-dried seagrass (*P. pacifica* DSM 14875) of Japanese coasts(Iizuka et al. 2003a).

Labrenzia sp. 011 was isolated from marine sediment of the coastal area of Kronsgaard, Germany. The sample was air-dried for 24 h and portions thereof were placed on artificial seawater-cycloheximide (ASW-WCX) agar plates. After 5 days of incubation at 30°C, creamy yellowish colonies were transferred to Petri dishes containing marine agar (Difco 2216) for purification through several passages until an axenic culture was obtained (Figure 9.9). Based on 16S rDNA alignments, the isolate shares 99% identity with its closest relatives, namely *Labrenzia alba* strain 5OM6 (AN: NR_042378.1) and *L. aggregata* strain RMAR6-6 (AN: CP019630.1). The bacterium was thus identified as *Labrenzia* sp. strain 011. The strain *Labrenzia* sp. 011 is stored in the German Collection of Microorganisms and Cell Cultures (DSMZ) with the accession number DSM 107099.

E. coli BKA13, kindly provided by Prof. Erwin Galinski (Institute of Microbiology and Biotechnology, University of Bonn, Bonn, Germany), is a derivative of *E. coli* MKH13 (Haardt et al. 1995) and cannot synthesize the compatible solute trehalose (Δ*otsB*, *otsA´*). Like its parental strain, *E. coli* BKA13 is missing all transporters for compatible solute accumulation and the genes for the conversion of choline into glycine-betaine; genotype:Δ(putPA)101, Δ(proP)2,Δ(proU)608, betTIBA.

3.2 Methods

3.2.1 Cultivation, extraction, and isolation of secondary metabolites (marine Myxobacteria)

All Myxobacteria were grown in ASW-VY/4 medium (1 L contains 75% artificial sea water (ASW), 25 mL of a 10% yeast suspension, trace elements solution and vitamin B_{12} filled up to the final volume with milli-Q water. Two 100 mL precultures, containing visible fruiting bodies, were used to inoculate 1 L ASW-VY/4 medium, respectively. The cultures were shaken on a rotary shaker at 140 rpm for 14 days at 30 °C. Adsorber resin Sepabeads® SP207 (Supelco, 20 g/L) was added to the cultures 48 hours before extraction. Bacterial pellet and adsorber resin were separated from the medium with a filter (pore size 2) and extracted with approx. 500 mL acetone until the

organic phase became uncolored. After the organic solvent was evaporated under vacuum conditions, the residue was redissolved in 100 mL aqueous methanol (60%) and extracted seven times with 100 mL dichloromethane. Crude lipophilic dichlormethane extracts were thus obtained. The extracts were further separated via RP_{18} Solid-Phase-Extraction (SPE) utilizing Bakerbond SPE Silica 1000 mg/ 6 mL columns and reduced pressure in a Bakerbound vacuum chamber. Thereby, a stepwise elution process with respectively 30 mL of petroleum ether, dichloromethane, acetone ethyl acetate, and methanol was employed.

For the isolation of enhygrolide A the myxobacterial strain *E. salina* SWB006 was cultivated in a 30 L stirred bioreactor using 20 L ASW-VY/4 medium containing 10 g/L of the adsorber resin Amberlite® XAD16N (Dow Chemical Company). The culture was grown at 28 °C, an airflow of 5 L/min and stirring at 200 rpm. After 120 h, the biomass and the adsorber resin were harvested by centrifugation and extracted with acetone and methanol until the organic phase got uncolored. The acetone phase was lyophilized and the residual 824 mg crude acetone extract was solved in acetone and adsorbed at 40 g Celite® 545 material.

This material was fractionated on a 12 g NP Silica 40 µm Reveleris® Flash cartridge by automatized Chromatography Systems REVELERIS® X2 Flash with integrated evaporative light scattering (ELSD)/ UV-Vis detection. A stepwise gradient solvent system of increasing polarity and a flow rate of 30 mL/min was used, starting with 100% hexane for 4.0 min to 100% CH2Cl2 within 6.0 min and hold for 3.0 min at 100% CH2Cl2. The gradient was changed then within 13.0 min to 100% EtOAc. Finally, the gradient was changed within 5.0 min to 20% MeOH and the cartridge was washed for additional 15 min under these conditions. According to the measured ELSD and UV signals at λ= 290, 320, and 350 nm the crude extract was separated into 18 fractions. Fraction ten, tR: 13–14 min yielded 2.0 mg of Enhygrolide A. The structure was confirmed by comparison of ^{1}H- and ^{13}C-NMR and HRESI-MS data with literature values (Felder et al. 2013b).

Enhygrolide A : white powder; ^{1}H and ^{13}C NMR data (see Table 9.1), HRESI-MS *m/z* = 357.1464 $[M + H]^+$ (calcd. for $C_{22}H_{22}NaO_3$, *m/z* 357.1461, 3.35 Δppm).

HPLC-MS/MS analysis

Samples were analyzed by HPLC-MS/MS on a micrOTOF-Q mass spectrometer (Bruker) with ESI-source coupled with a HPLC Dionex Ultimate 3000 (Thermo Scientific) using a Zorbax Eclipse Plus C18 1.8 µm column, 2.1x50 mm (Agilent). The column temperature was 45° Cs. MS data were acquired over a range from 100-3000 m/z in positive mode. Auto MS/MS fragmentation was achieved with rising collision energy (35-50 keV over a gradient from 500-2000 m/z) with a frequency of 4 Hz for all ions over a threshold of 100. UHPLC starting conditions with 90 % H_2O containing 0.1% acetic acid as mobile phase were kept isocratic for 0.5 min. Followed by a gradient solvent system to 100% acetonitrile (0.1% acetic acid) within 4 min. 2 µl of sample solution was injected to a flow of 0.8 ml/min. All MS/MS data were converted to .mzXML format and transferred to the GNPS server (gnps.ucsd.edu) (Wang et al., 2016) and molecular networking was performed based on the GNPS data analysis workflow using the spectral clustering algorithm(Guthals et al. 2012).

3.2.2 Molecular networking (marine Myxobacteria)

For the molecular network analysis, all nodes that contained ions from blank medium were removed. A molecular network was created by the online workflow at GNPS (Wang et al. 2016) using the spectra with a minimum of four fragment ions and by merging all identical spectra into nodes, representing parent masses. Compounds with similar fragmentation patterns are connected by edges, displaying molecular families with similar structural features. The data was filtered by removing all MS/MS peaks within +/- 17 Da of the precursor m/z. MS/MS spectra were window filtered by choosing only the top 6 peaks in the +/- 50 Da window throughout the spectrum. The resulting data were then clustered by MS-Cluster with a parent mass tolerance of 0.02 Da and a MS/MS fragment ion tolerance of 0.02 Da to create consensus spectra. Further, consensus spectra that contained less than 2 spectra were discarded. A network was then created where edges were filtered to have a cosine score above 0.5 and more than 4 matched peaks. Further edges between two nodes were kept into the network if and only if each of the nodes appeared in each other's respective top 10 most similar nodes. The spectra in the network were then searched against GNPS' spectral

libraries. The library spectra were filtered in the same manner as the input data including analog search. All matches kept between network spectra and library spectra were required to have a score above 0.5 and at least 4 matched peaks. The network was visualized via Cytoscape 3.6.1. The molecular network file is available at NDEx site (http://doi.org/10.18119/N9988C).

3.2.3 Media and growth conditions for organic osmolytes (marine Myxobacteria)

To prepare a medium for the marine Myxobacteria without any background of organic osmolytes, first an *E. coli* BKA13 culture was prepared. Therefore, *E. coli* BKA13 cells were grown in the minimal medium MM63 including 0.5% NaCl (Grammann et al. 2002; Larsen et al. 1987) for 20 h, shaking at 37° C. The OD_{600} was measured and the culture was precipitated by centrifugation for 10 min at 8873 rcf. The cell pellet was washed twice with tap water and subsequently resuspended in tap water and used for the preparation of ASW-Coli medium as described below. The amount of tap water used for resuspension was calculated by following equation: (OD_{600}×initial culture volume)/50.

E. salina SWB007 and *P. pacifica* DSM 14875 were grown in ASW-Coli medium containing 75% artificial sea water (ASW), 3% *E. coli* BKA13 suspension in Milli-Q water. To determine the salt tolerance range of *E. salina* SWB007, 100 ml Erlenmeyer flasks containing 30 ml ASW-Coli medium with different NaCl concentrations (from 0% NaCl to 5% NaCl, in 0.5% steps) were prepared. The cultures were inoculated with 100 µl of dispersed fresh fruiting bodies. The dispersed fruiting bodies suspension was prepared beforehand by manually collecting fruiting bodies from another liquid culture by using a pipette. Growth was determined by measuring the decrease in OD_{600}, since the *E. coli* cells served as prey and were lysed by the Myxobacteria , which resulted in clearing of the medium.

For compatible solutes identification the Myxobacteria were grown in 5 L Erlenmeyer flasks containing 1 L ASW-Coli medium with 0.5% and 3% NaCl for *E. salina* SWB007, and 1% and 3.5% NaCl for *P. pacifica*, respectively. Two independent experiments were

performed and the mean value is given in Table 5.1. Incubation was performed at 30°C at 140 rpm, using rotary shakers. After lysis of all *E. coli* cells (4 days, except for *E. salina* in 3.5% NaCl medium, which was incubated for 5 days), the fruiting bodies were collected and precipitated by centrifugation at 17000 rcf for 1min. The supernatant was discarded and the pellet was freeze-dried using a lyophiliser (Christ Beta 1-16, Martin Christ GMBH, Osterrode, Germany).

3.2.4 Extraction of organic osmolytes (marine Myxobacteria)

The extraction of cytoplasmic solutes was performed following the Bligh & Dyer extraction method, as modified by Galinski & Herzog (Galinski and Herzog 1990) . In brief, 500 µl of Bligh & Dyer solution (10/5/4 v/v/v methanol/chloroform/H_2O_{demin}) were added to 30 mg of dry material. After shaking for 5 min, 130 µl of chloroform and 130 µl of H_2O_{demin} were added and the mixture was shaken again for 5 min. Finally, the mixture was centrifuged at 17000 rcf for 3 min and the hydrophilic top layer, containing the compatible solutes, was taken off for subsequent analyses.

3.2.5 HPLC and LCMS analysis for compatible solutes and free amino acids (marine Myxobacteria)

The aqueous phase was analysed by HPLC using a refractive index detector RI-71 (Shodex) and a UV detector UV1000 (Thermo scientific) with the following HPLC conditions. Column: LiChrospher 100-NH_2 5 µm (Merck); isocratic flow of 80% acetonitrile and 20% water at 1 ml/min. The samples were diluted 1:4 with the solvent before measurement. Additionally, reference compounds were run to identify the compatible solutes. The evaluation was performed with ChromQuest5 (Thermo Quest Corporation). Mass spectrometric analysis was performed using an ESI-MS microTOF spectrometer (Bruker Daltonik, Bremen, Germany) coupled to an HPLC. The same parameters as described above were applied, and the sample was split before transfer to ESI-MS.

Free amino acids were analysed by FMOC/ADAM-HPLC using a flourescence detector (Spectra System FL2000; 254 nm exc.; 315 nm em.; Thermo separation products).

HPLC column: Merck Superspher 60 RP-8, 4µm 125x4 mm LiChrocart-System (Alltech Grom GMBH). Amino acids were derivatised with 9-fluorenylmethoxycarbonyl (FMOC) by adding 40 µl of boric acid buffer (0.5 M boric acid and 25 µM norvalin) and 80 µl FMOC-reagent (1 mM FMOC in acetone) to 40 µl of 1:50 diluted sample material. After shaking for 45 s, 100 µl 1-aminoadamantane (ADAM) reagent (50% 40 mM ADAM in boric acid buffer, 50% acetone) were added and the mixture was shaken again for 45 s. Finally, 140 µl of solvent A (80% 50 mM sodium acetate in H_2O, 20% 0.5% tetrahydrofuran in acetonitrile; (v/v)) were added. Solvent B consisted of 20% sodium acetate buffer and 80% acetonitrile. The HPLC was run with 1.25 ml/min using the following gradient; 100% to 91% A in 15 min, 91% A to 70 % A in the next 15 min, 70% A to 40 % A in the next 10 min, 40% A to 0 % A (=100% B) in the next 2 min, maintained at 100% B for the next 7 min, 100% B to 100% A in the next 2 min and maintained at 100% A for the next 2 min. The evaluation was performed with ChromQuest5 (Thermo Quest Corporation).

3.2.6 Genome sequencing and assembly

The genomic DNA of *E. salina* SWB005 and SWB007 was isolated as described before (Amiri Moghaddam et al. 2018). In brief, fruiting bodies, which appeared after several days of fermentation in ASW-VY/4 liquid medium (see cultivation for details), were harvested. DNA was isolated using the GenElute™ Bacterial Genomic DNA Kit (Sigma-Aldrich). Illumina shotgun paired-end sequencing libraries were generated and sequenced on a MiSeq instrument (Illumina, San Diego, CA, USA). Quality filtering using Trimmomatic version 0.36 (Bolger et al. 2014) resulted in 3,773,950 and 3,458,266 paired-end reads for *E. salina* SWB007 and *E. salina* SWB005, respectively. Paired-end reads were combined using the Spades assembler v3.10, yielding initial sequence contigs (Bankevich et al. 2012). After filtering contigs smaller than 500 bp, the remaining contigs were determined with Quast (Gurevich et al. 2013). Genome completeness was estimated using CheckM (Parks et al. 2015) and compared to the published genome data of *E. salina* DSM15201. The resulting genomes have been deposited at NCBI GenBank with the accession numbers PVNK00000000 (*E. salina* SWB005) and PVNL00000000 (*E. salina* SWB007)(Amiri Moghaddam et al. 2018). The

genome sequences of *E. salina* DSM 15201, *P. pacifica* DSM 14875 and *Haliangium ochraceum* DSM 14365 were obtained from NCBI GenBank, accession numbers are JMCC00000000, ABCS00000000 and CP001804 (Ivanova et al. 2010), respectively.

The genomic DNA of *Labrenzia* sp. 011 was isolated as described before (Harms et al. 2017). In brief, cell pellets were harvested from a one-week culture in marine broth liquid medium. DNA was isolated using the GenElute™ Bacterial Genomic DNA Kit (Sigma-Aldrich). Illumina shotgun paired-end sequencing libraries were generated and sequenced on a MiSeq instrument (Illumina, San Diego, CA, USA). Quality filtering using Trimmomatic version 0.36(6) resulted in 495,158 paired-end reads for *Labrenzia* sp. 011. Paired-end reads were combined using the Spades assembler v3.10, yielding initial sequence scaffolds (Bankevich et al. 2012). After filtering scaffolds smaller than 1 kb, 65 contigs remained as determined with Quast (Gurevich et al., 2013). Genome completeness was estimated using CheckM (Parks et al. 2015), using genus level marker genes, and yielded 83.2%. The available genome sequences of these *Labrenzia* strains were used (accession numbers in parentheses): *L. alexandrii* DFL-11^T (ACCU00000000), *L. aggregata* RMAR6-6 chromosome (CP019630), *L. suaedae* DSM 22153^T (FRBW00000000), *L. alba* CECT 5095^T (CXWE00000000), *Labrenzia* sp. CP4 (CP011927), *Labrenzia* sp. VG12 (CP022529), *Labrenzia* sp. DG1229 (AYYG00000000), *Labrenzia* sp. C1B10 (AXBY00000000), *Labrenzia* sp. C1B70 (AXCE00000000), *Labrenzia* sp. OB1 (JSEP00000000), *L. marina* DSM 17023^T (PPCN00000000), and *Labrenzia* sp. UBA4493 (DGNL00000000).

3.2.7 Genome alignment and annotation

To ease the comparative study of the draft genomes, Mauve Contig Mover (MCM)(Rissman et al. 2009) was used to order and/or reverse the contigs and align the other myxobacterial draft genomes relative to the *E. salina* SWB007 draft genome. FASTA files were used as input and the reordered FASTA files of the mauve output data were used for further analysis. Coding sequences of the reordered contigs were determined by using the RAST prokaryotic genome annotation server (Aziz et al. 2008). Therefore, the genetic code 11, which is used by most bacteria, was used in classic

RAST and the options “automatically fix errors”, “fix frame shifts”, “build metabolic model” and “backfill gaps” were selected. To obtain the putative pathways of terpenoid buliding blocks, KEGG maps of the terpene backbone biosynthesis and degradation pathways of leucine, isoleucine and valine were compiled using RAST as hierarchical trees. All reactions for a given cellular process with links to the KEGG map were visualized with annotated proteins, which putatively catalyze the reaction (Aziz et al. 2008).

3.2.8 Genome comparison

The EDGAR 2.2 genomic pipeline was used for genome comparison (Blom et al. 2016). Therefore, the RAST-annotated GenBank files were uploaded to EDGAR and the core genome, orthologous genes and singletons were identified. Visualization was done using a Venn diagram; core genome size and gene numbers in every subset of the dispensable genomes were indicated. To visualize the drop of the core genome size and the increase of the pan genome with the introduction of each genome, a core vs. pan plot of the genomes was generated. To compare the gene order and co-localization of genes in the different genomes, a synteny plot was generated.

Phylogenetic trees of the investigated marine Myxobacteria and Labrenzia strains were constructed based on a linear combination of multiple alignments of the nucleotide sequences of orthologous genes in the core genome. The alignments were created using MUSCLE (Edgar 2004), and the PHYLIP (Felsenstein 1989) implementation of the neighbor-joining algorithm was used to deduce the tree. For a deeper qualitative comparison between the genomes, the average amino acid identity (AAI) and average nucleotide identity (ANI) matrixes of all conserved genes in the core genome were computed by the BLAST algorithm using JSpeciesWS (Richter et al. 2016) or EDGAR and visualized as heat maps. In silico DNA-DNA hybridization (isDDH) was performed based on identities / HSP length formula using the DSMZ GGDC service tool (Meier-Kolthoff et al. 2014). The CGView Comparison Tool (CCT) was used to create a graphical map of the BLAST results comparison of the available genomes to the reference genome (Grant et al. 2012).

3.2.9 Prediction of specialized metabolites biosynthetic gene clusters and similarity network of the BGCs

Biosynthetic gene clusters (BGCs) for specialized metabolites were identified using antiSMASH v4 (Blin et al. 2017b), default parameters and incorporation of the ClusterFinder algorithm were applied. The distribution of all identified BGCs of the AntiSMASH analysis was visualized in a circular chord diagram using Circos table viewer, whereby the putative BGCs were not considered (Krzywinski et al. 2009). A similarity network of the BGCs among different genomes of marine Myxobacteria was obtained using a modified Pfam domain similarity metric implemented in BigScape (Yeong M. 2016; Ceniceros et al. 2017). Various cutoffs ranging from 0.5 to 0.85 were used to produce different network versions. Based on manual inspection of connections to known BGCs at various versions of the network and regional alignments were done by EDGAR, a cut-off of 0.75 was chosen for the final analysis (Ceniceros et al. 2017). Additional screening for resistance markers and potential antibiotic targets was performed using the ARTS webserver (Alanjary et al. 2017) and clusters positive for known resistance markers and duplicated essential genes were subsequently annotated in the final similarity network using Cytoscape 3.6.1. The similarity network file is available at NDEx site (Pratt et al. 2015) (http://doi.org/10.18119/N9F30V). The fraction of the genomes with a shared BGC that is devoted to specialized metabolism was aligned using EDGAR regional alignment to enable comparison of the similar gene clusters (Blom et al. 2016).

3.2.10 Bioinformatics analyses for organic osmolytes in marine Myxobacteria

The genome of *P. pacifica* DSM 14875 and a draft genome of *E. salina* SWB007 were analyzed using the bioinformatics program antiSMASH3.0 (Weber et al. 2015). In the genomic data of *E. salina* SWB007 a gene locus was identified harboring genes related to betaine, ectoine, and hydroxyectoine biosynthesis. The gene cluster was subsequently annotated using the RAST annotation program (version 2.0) (Overbeek et

al. 2014) and the UniProt Basic Local Alignment Search Tool (BLAST). The contig was submitted to GenBank (accession number: KU237243).

3.2.11 Prediction of CFAS coding regions in *Labrenzia* spp.

Coding sequences of all analyzed genomes were determined using the RAST prokaryotic genome annotation server (Aziz et al. 2008). Coding regions for cyclopropane fatty acyl-phospholipid synthase (CFAS) were detected using RAST genome browser (Aziz et al. 2008) and AntiSMASH v4 with applied default parameters and incorporation of the ClusterFinder algorithm (Blin et al. 2017b).

3.2.12 Genome comparison and phylogeny of CFAS in *Labrenzia* spp.

The EDGAR 2.2 genomic pipeline was used for genome comparison (Blom et al. 2016). Therefore, the RAST-annotated GenBank files were uploaded to EDGAR and the core genome, orthologous genes and singletons were identified. The average amino acid identity (AAI) matrix of all conserved genes in the core genome were computed by the BLAST algorithm and visualized as heat map (Blom et al. 2016). The fractions of the genomes with conserved CFAS genes were searched using EDGAR regional alignment to enable comparison of the similar gene regions (Blom et al. 2016) and mapped using MultiGeneBlast (Medema et al. 2013). Multiple sequence alignment and phylogenetic analysis of the CFAS genes were done using Clustal Omega program located in UniProt database (Sievers et al. 2011).

4 Results I: Genome mining of marine Myxobacteria for bioactive metabolites

4.1 Draft genome sequences of the obligatory marine Myxobacteria *Enhygromyxa salina* SWB005 and *E. salina* SWB007

4.1.1 Background

Enhygromyxa salina is a marine myxobacterium, and strains of this species have been isolated from different locations around the globe (Iizuka et al. 2003a; Schäberle et al. 2010). To date only a few obligatory marine Myxobacteria are known and only one genome was available, i.e. from the strain *E. salina* DSM 15201. Like other Myxobacteria the genomes are large, in a range of 8-10 Mbp, and harbor many BGCs coding for the production of specialized metabolites. The here analyzed strains are producers of the natural products salimabromide, salimyxins, and enhygrolides (Felder et al. 2013a; Felder et al. 2013b; Dávila-Céspedes et al. 2016). *E. salina* SWB005 and *E. salina* SWB007 were isolated from marine sediments from the coast of Santa Barbara, California, USA (SWB005) and Prerow, Germany (SWB007), respectively (Schäberle et al. 2010; Felder et al. 2013a).

4.1.2 Genome sequences of new marine Myxobacteria *Enhygromyxa salina* SWB005 and SWB007

Isolated DNA from fruiting bodies of both *E. salina* strains were sequenced. The assembled sequence resulted in 192 contigs (>500 bp) with 73-fold average coverage for *E. salina* SWB007 and in 312 contigs (>500 bp) with 75-fold average coverage for *E. salina* SWB005. The assemblies were validated and the read coverage determined with QualiMap version 2.1 (García-Alcalde et al. 2012). The resulting draft genomes are 10,602,813 bp (*E. salina* SWB007) and 9,010,436 bp (*E. salina* SWB005) in length, while the GC contents are 68.1% and 69.5% (difference of 1.4%), respectively. Automatic annotation and identification of rRNA and tRNA genes was performed using

the software tool Prokka (Seemann 2014). This yielded 2 rRNA genes, 78 tRNA genes, 3,682 protein-encoding genes with function prediction, and 3,265 genes coding for hypothetical proteins for strain SWB005 and 4 rRNA genes, 57 tRNA genes, 4,253 protein-encoding genes with function prediction, and 3,987 genes coding for hypothetical proteins for strain SWB007, respectively.

In silico DNA-DNA hybridization (isDDH) was performed based on the identities/high-scoring segment pairs (HSP) length formula (Meier-Kolthoff et al. 2014), and produced a value of 26.10% (+23.8, -28.6%) which failed the isDDH cutoff of ≥70% that would have determined them to be the same species. The average nucleotide identity (ANI) was calculated to be 78.7%. Therefore, the *in silico* parameters ANI, isDDH and the difference of the G+C value define these two strains as two distinct species of the genus *Enhygromyxa*.

4.1.3 Biosynthetic potential of the newly sequenced *Enhygromyxa salina* strains for bioactive metabolites

Using the antiSMASH 4.0.0 tool (Blin et al. 2017b) for the analysis of the genomes revealed 40 BGCs in *E. salina* SWB005 and 46 BGCs in *E. salina* SWB007, indicating the high potential of these bacteria for biosynthesis of specialized metabolites.

4.1.4 Accession numbers

This Whole Genome Shotgun projects have been deposited at DDBJ/ENA/GenBank under the following accessions: PVNK00000000 (*E. salina* SWB005) and PVNL00000000 (*E. salina* SWB007), the versions described in this study are version PVNK01000000, and PVNL01000000, respectively.

4.2 Comparative genomic and metabolomic analysis of marine Myxobacteria

4.2.1 Background

Microorganisms are well known for their ability to produce so called secondary or rather specialized metabolites, of which many have proven to have important medical, biotechnological, agricultural, and nutritional applications (Rutledge and Challis 2015). In recent years, genome mining has led to a paradigm shift in natural products research as BGCs, regions of the genome that encode for the production of natural products, can be identified and characterized from the genomes of many microorganisms (Cimermancic et al. 2014). A BGC represents both, a biosynthetic and an evolutionary unit, which can be identified using genome mining software tools like antiSMASH (Blin et al. 2017b). This sequence-based approach increases the chance for discovery of new metabolites by identifying the talented microbes using genome sequence analysis and subsequent characterization of the *in silico* identified BGCs (Naughton et al. 2017). The comprehensive biosynthetic potential, including silent clusters, rather than what is currently expressed and apparent in the lab, is shown. Combined with a metabolomic approach, using high resolution mass spectrometry and molecular networking, rediscovery of known metabolites can be avoided at a very early stage of the discovery process through dereplication (Maansson et al. 2016; Yang et al. 2013), and simultaneously, discovery of novel natural products can be streamlined through optimization of culture conditions (Crüsemann et al. 2017).

Marine environments, holding 95% of the earth's biosphere, have come into the focus for natural product discovery as a consequence of the emergence of antimicrobial resistance, boosted by the limitations in novel drug developments from the usual producers of terrestrial environments (Naughton et al. 2017). Myxobacteria are a group of Deltaproteobacteria, which have been first discovered from soil in 1809. These organisms were thought to be occurring exclusively in terrestrial environments until Iizuka *et al* reported in 1998 the isolation of Myxobacteria from marine environment (Dawid 2000; Iizuka et al. 1998). Terrestrial Myxobacteria have been well investigated over the past three decades, which resulted in more than 100 natural product scaffolds and approximately 600 structural derivatives with a broad range of biological activities

(Plaza and Müller 2014). However, to date, only 10 obligatory marine myxobacterial strains, which need sea-like conditions in order to grow, have been isolated and from them, only seven groups of natural products have been identified, including enhygrolides, enhygromic acid, haliamide, haliangicins, salimabromide, salimyxins, and triterpenoid sterols (Figure 4.1)(Tomura et al. 2017; Dávila-Céspedes et al. 2016; Sun et al. 2016b; Wei et al. 2016). The lack of more marine myxobacterial isolates and natural products is mainly due to the difficulties in isolation and cultivation of these bacteria (Felder et al. 2013b).

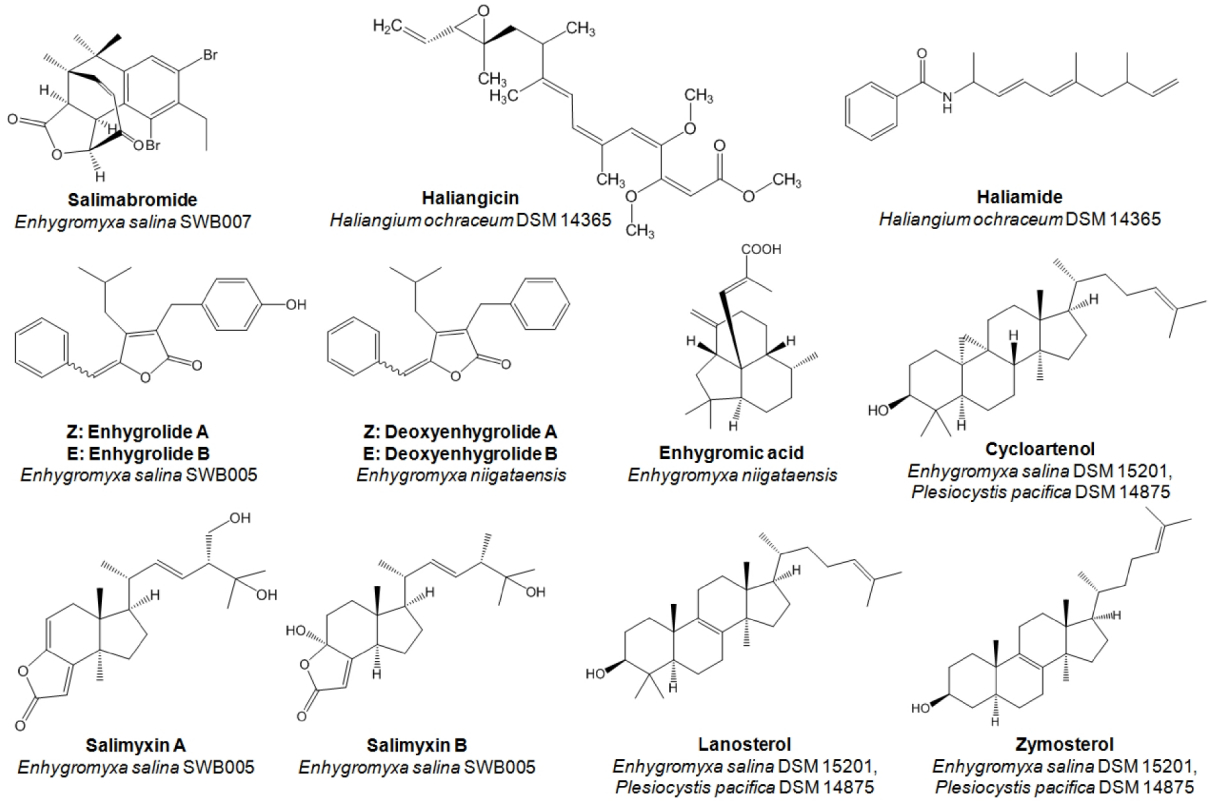


Figure 4.1 Structures of specialized metabolites isolated from marine myxobacteria

Here, we conducted comparative genomic analysis of the five marine Myxobacteria for which genomes are publicly available, thereunder two newly sequenced strains from our lab. This analysis was carried out in order to compare the similarities and differences in the biosynthesis of specialized metabolites in marine Myxobacteria. We report the distribution and similarity within the existing BGCs in the genomes, revealing the uniqueness and variability of BGCs harbored by these bacteria. Furthermore,

metabolomes of the marine myxobacterial strains were analyzed and compared using mass spectrometry networking, to evaluate if the trends from genome analysis are translatable into actual metabolite profiles.

4.2.2 General characteristics and phylogeny of marine myxobacterial genomes

Five draft genomes are currently available from obligatory marine Myxobacteria: *Plesiocystis pacifica* DSM 14875, *Haliangium ochraceum* DSM 14365, *Enhygromyxa salina* DSM 15201, and, related to the latter one, *E. salina* SWB005 and *E. salina* SWB007, of which the last two were recently sequenced from our working group (Amiri Moghaddam et al. 2018) (Table 4.1). The quality of the draft genomes differs and the number of contigs varies between 1 for *H. ochraceum* DSM14365 to 330 for *E. salina* DSM 15201. However, all strains possess large genomes ranging from 9 to 10.6 Mbp. Like in terrestrial Myxobacteria, the GC content is rather high, i.e. between 67 and 71% and the number of predicted gene coding sequences (CDS) is around 7,000-8,500, which is in accordance to the large genome size of these strains.

Table 4.1 General characteristics of available marine myxobacterial genomes

Lineage	**Bacteria; Proteobacteria; Deltaproteobacteria; Myxococcales**				
Species	*Enhygromyxa salina*			*Plesiocystis pacifica*	*Haliangium ochraceum*
Strain[a]	SWB007	SWB005	DSM 15201	DSM 14875	DSM 14365
Genome size (Mbp)	10.6	9.0	10.4	10.6	9.4
GC Content	68.1	69.5	67.4	70.7	69.5
Number of Contigs	192	312	330	237	1
CDS	8293	7054	8178	8447	7032

[a] The accession numbers are given in the Materials and Methods section

A phylogenetic tree of marine Myxobacteria was constructed based on a nucleotide sequence alignment of the core genomes (see below) (Figure 4.2). The *E. salina* strains belong to the order of Myxococcales and the *P. pacifica* DSM 15201 type strain is the closest relative to the *E. salina* clade. They are part of the *Nannocystaceae* family. However, the first isolated marine myxobacterium *H. ochraceum* DSM 14365^T belongs to the family of *Kofleriaceae* and the core genome of this strain is distinct from the other marine Myxobacteria (see below).

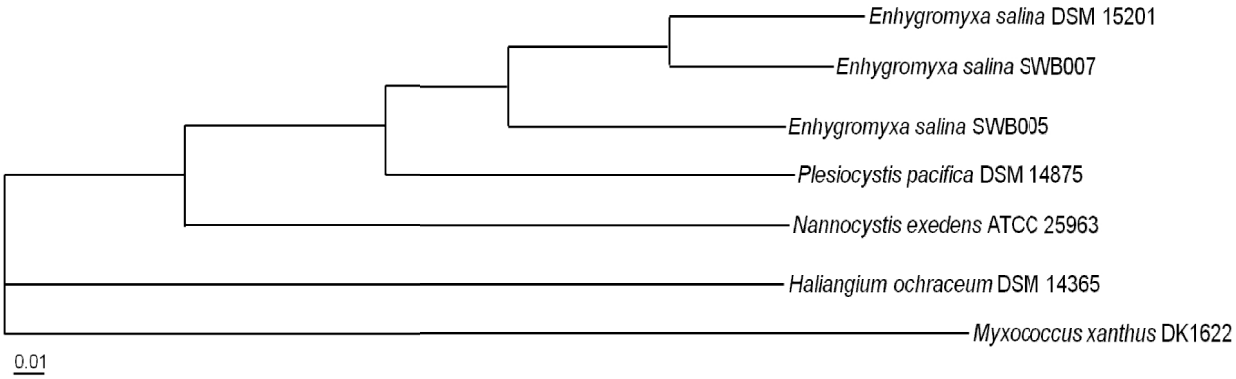


Figure 4.2 Phylogenetic tree of selected marine myxobacteria. Available genomes of marine myxobacteria were used to build the tree based on nucleotide sequence alignment of the core genomes. The closely related halophilic strain *Nannocystis exedens* ATCC 25963, as well as the terrestrial *Myxococcus xanthus* DK1622, which represents the outgroup, were included. Tree for 7 genomes, build out of a core of 645 genes per genome, 4515 in total. The core has 838246 bp per genome, 5867722 in total. The tree topology was evaluated in 500 bootstrap iterations and showed a branch conservation of 100%.

4.2.3 Genome comparison of marine myxobacterial strains

A synteny plot of the reciprocal best blast hits of all CDS within the contiguous contigs was constructed using the EDGAR pipeline. The genome of *E. salina* SWB007 was chosen as reference for synteny analysis, because (i) it is bigger in size, thereby the chance to cover genomic parts of the other strains is higher, (ii) it is of high quality, and (iii) due to the high relationship between the genera *Enhygromyxa* and *Plesiocystis*, which excludes *Haliangium* as reference.

According to the synteny plot, there are many CDS located in different positions compared to the reference genome of *E. salina* SWB007. However, there is still rather good synteny of orthologous genes within the areas that reside inside contig boundaries of *E. salina* SWB007 and the genomes of *E. salina* SWB005 and DSM 15201, as well as *P. pacifica* DSM 14875. The latter showed slightly lower synteny (Figure 4.3 A). This result indicates a low degree of genome divergence within these marine Myxobacteria. On the nucleotide level, *E. salina* SWB007 and DSM 15201 are highly similar, while the identity ratio of *E. salina* SWB005 is slightly lower and further decreases for *P. pacifica* DSM 14875 and *H. ochraceum* DSM 15365, respectively (Figure 4.3 B).

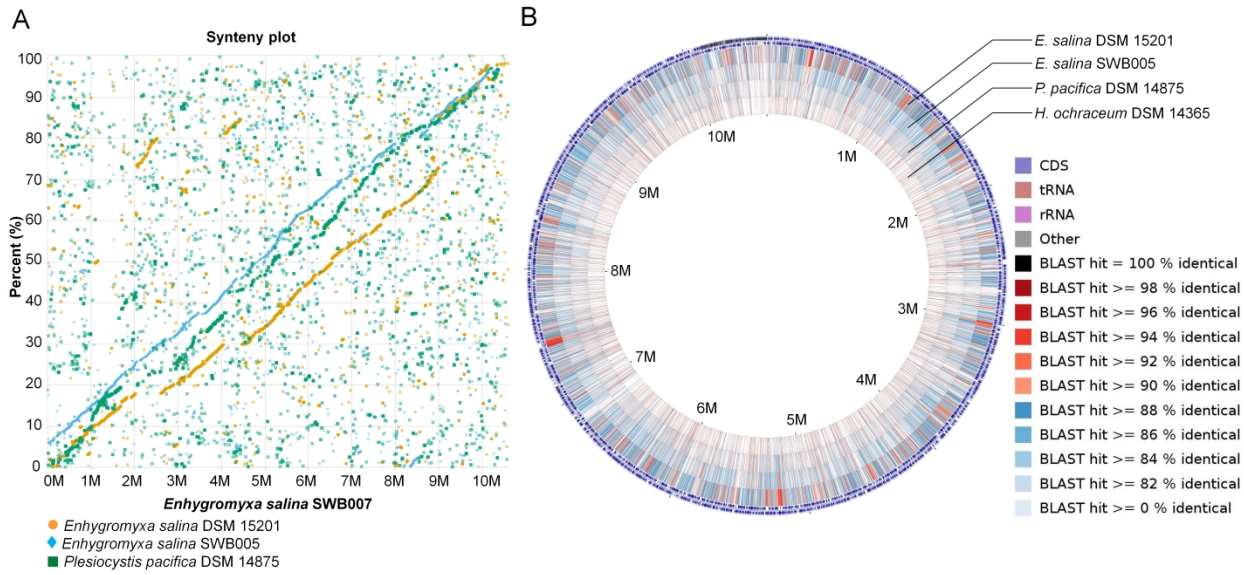


Figure 4.3 A: Synteny plot of selected genomes of marine myxobacteria within the contiguous contigs. *E. salina* SWB007 serves as reference. Positions within the reference genome are depicted on the x-axis, while the y-axis shows the relative position within the other genomes. Dots reflect the stop position of orthologous genes within the areas that reside inside contig boundaries. B: BLAST comparison of selected strains. *E. salina* SWB007 genome serves as reference and *E. salina* DSM 15201, *E. salina* SWB005, *P. pacifica* DSM 14875 and *H. ochraceum* DSM 14365 are indicated each by one ring.

Based on in silico parameters which determine if genomes belong to the same species (*i.e.* ANI value ≥96%, isDDH value ≥70%, and difference in G+C content ≤1%)(Meier-Kolthoff et al. 2014; Colston et al. 2014), both strains, *E. salina* SWB005 as well as SWB007, can be considered as distinct new species of the genus *Enhygromyxa*. The ANI value between *E. salina* SWB007 and *E. salina* DSM 15201 is 85% with an isDDH

value of 29% and a G+C difference of 0.7%. The ANI values between *E. salina* SWB005 and the other *E. salina* strains is 79% with an isDDH value of 23% and a G+C difference of more than 1% (Figure 4.4). On the amino acid level, all *E. salina* strains and *P. pacifica* DSM 14875 show 74.7-92.7% average amino acid identity (AAI) between each other (Figure 4.5). Therefore, the orthologous genes in these strains probably perform the same functional roles. However, the function of the orthologous genes in *H. ochraceum* DSM 14365 is more uncertain, since the AAI is only 48% towards other strains (Figure 4.5).

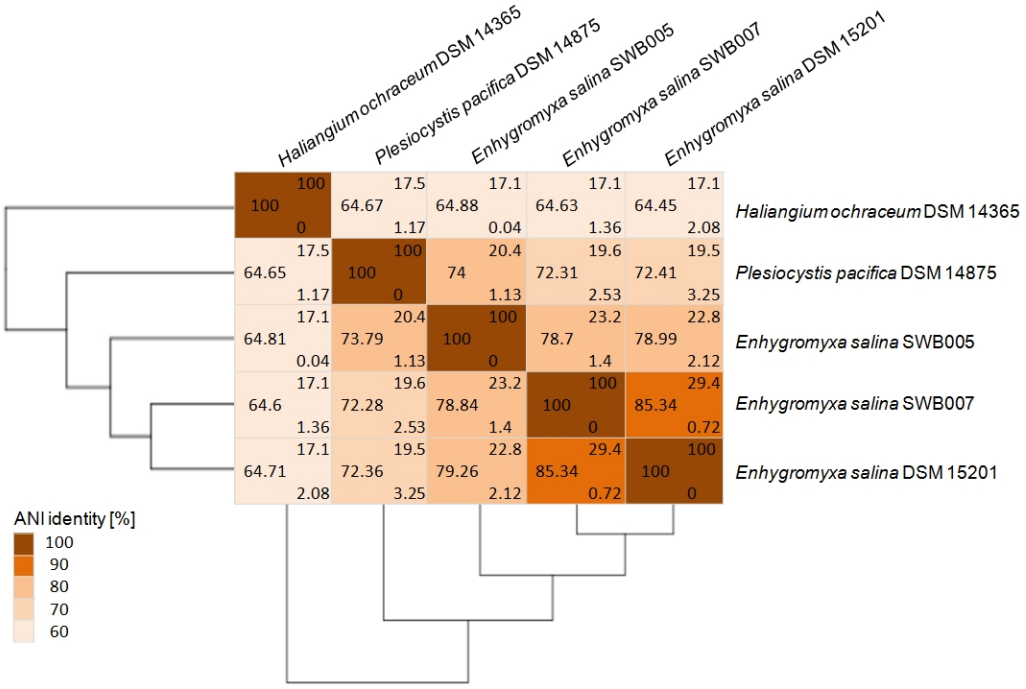


Figure 4.4 Average Nucleotides Identity (ANI) matrix of the available marine Myxobacteria genomes combined with the DNA-DNA hybridization (DDH) estimation and difference in G+C content. In each box: [a] (left-middle): ANI, [b] (right-up): DDH and [c] (right-down): difference in G+C content; all values are given in percent.

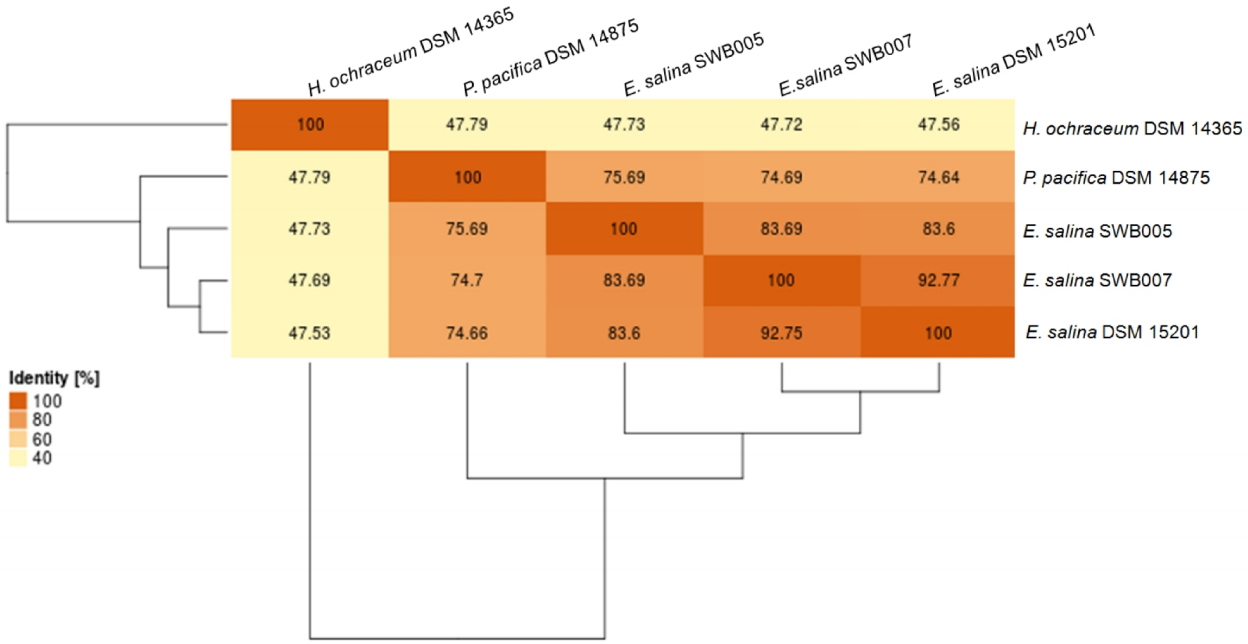


Figure 4.5 Average Amino Acids Identity (AAI) of the available genomes of marine Myxobacteria. Orthologous genes of the core genomes were analyzed for their median identity.

In order to obtain further insights into the degree of similarity between the analyzed genomes, the number of core genes, as well as of singletons was determined (Figure 4.6A). Even for the most closely related strains investigated here, *i.e.* the *E. salina* strains, more than 1600 CDS represent singletons, which is equivalent to 21-23% of each genome (Figure 4.6A). This value duplicates if the next further relative, *i.e. P. pacifica* DSM 14875 is considered, since this strain has 3365 singletons (equivalent to ~40% of the genome). *H. ochraceum* DSM14365 has 5220 (equivalent to 74% of the genome) singletons (Figure 4.6A). The core genome of these marine Myxobacteria consists of 1130 CDS. This relatively low number, equivalent to 13-16% of the CDSs per strain, is due to the inclusion of the more distantly related *H. ochraceum* DSM 14365 genome to the analysis. For comparison the core genome of six *Myxococcus* genomes, including 4 different species and three *M. xanthus* strains consists of 4,693 CDS. This accounts for 56.6–63% of the CDSs in each genome (Sharma et al. 2016). If only the *E. salina* strains are considered, they have >4600 CDSs in their core genome, and inclusion of *P. pacifica* DSM 14875 in the analysis results in a core genome of >3600 CDSs (Figure 4.6B). Hence, the pan genome

increases by about 2000 CDSs by every additional *E. salina* strain. If the other marine Myxobacteria are included, the pan genome increases further by almost 3500 CDSs of *P. pacifica* DSM 14875 and by 5000 CDSs of *H. ochraceum* DSM14365, respectively (Figure 4.6B).

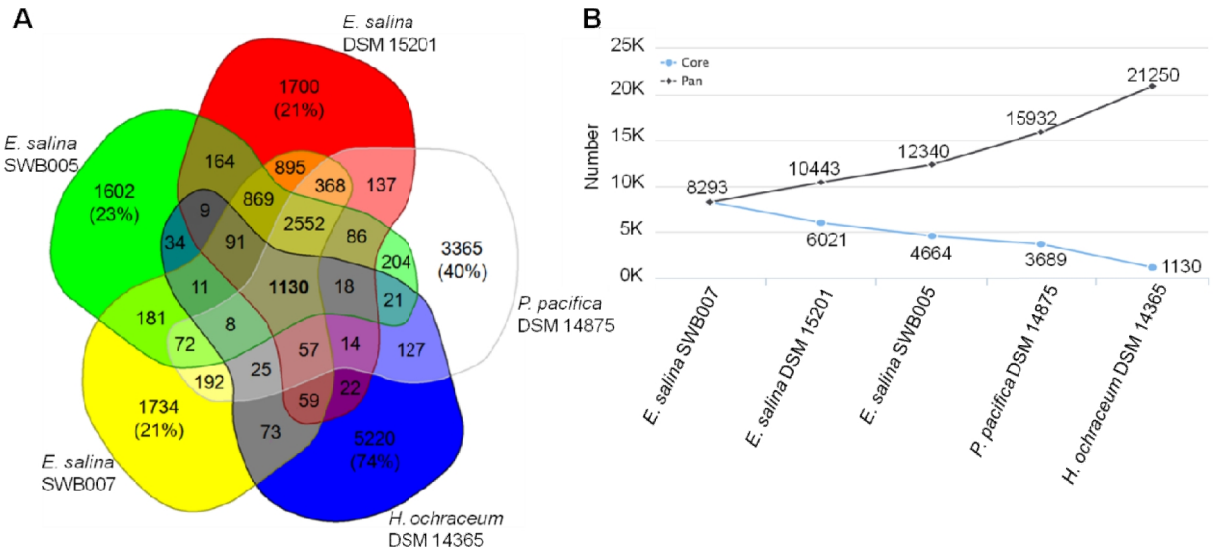


Figure 4.6 A: Venn diagram of the CDS counts in different subsets of the genomes (singletons are given as percentage of the respective genome). B: core vs. pan genome plot of the genomes.

4.2.4 Presence of specialized metabolite biosynthetic gene clusters in the genomes of marine myxobacterial strain

In order to estimate the potential of the strains for the production of specialized metabolites, the genomes were screened *in silico* for the presence of BGCs putatively coding for the production of such compounds (Blin et al. 2017b). All organisms investigated have a high variety of BGCs in their genomes, *i.e.* 30-46 putative BGCs were identified in each strain (Table 4.2). These numbers even doubled, if a more general cluster finder algorithm was applied to estimate the cluster boundaries (assigning putative BGCs) based on frequencies of locally encoded protein domains detected by Pfam (Blin et al. 2017b). In terms of novel metabolites, the numbers of identified BGCs by AntiSMASH which had similarities to known BGCs from the MIBiG database (Medema et al. 2015) were counted (Table 4.2). 10-11 BGCs of each *E.*

salina strain, 5 from *P. pacifica* DSM 14875, and 17 from *H. ochraceum* DSM 14365 matched partly or completely to validated gene clusters.

Analyzing the classes of metabolites predicted from the identified BGCs, revealed that PKs (2-11 per strain), fatty acids (5-9), and terpenes (3-9) represent the majority of predicted specialized metabolites, followed by bacteriocins (4-6). NRPs (0-4) and mixed PK/NRPs (0-3) are less common (Figure 4.7). However, it should be noted that because draft genome sequences were analyzed, big BGCs such as PKS and NRPS can be split across contigs and the real number of BGCs might be overestimated.

To get additional insights into the nature of the metabolites putatively corresponding to a BGC, an analysis using the ARTS webserver was performed (Alanjary et al. 2017). This tool aims to enable prioritization of BGC, which correspond to antibacterial compounds. It is based on the fact that, to avoid suicide, an antibiotic producer harbors resistance genes often within the same BGC responsible for manufacturing of the compound. Known resistance, as well as possible resistant housekeeping genes are detected (Alanjary et al. 2017). Using this analysis, several resistance model hits were identified (Table 4.2 and Figure 4.8) suggesting that these specific BGCs code for antibacterial compounds. 7 to 13 resistance model hits were identified among the *E. salina* strains, including beta-lactamase, ABC-transporters, and other efflux systems. In *P. pacifica* DSM 14875 and *H. ochraceum* DSM 14365 only 4 hits pointing toward antibacterials were identified.

Table 4.2 Overview of antiSMASH and ARTS analysis[a]

BGCs	*E. salina* SWB007	*E. salina* SWB005	*E. salina* DSM 15201	*P. pacifica* DSM 14875	*H. ochraceum* DSM 14365
Total[b]	80	56	77	76	62
Predicted by antiSMASH	46	40	38	34	30
Predicted putative	34	16	39	42	32
MIBiG hits	11	10	10	5	17
Resistance model hits	9	7	13	4	4

[a] The number of BGCs predicted by antiSMASH and of the antibiotic-related BGCs predicted by ARTS is given. [b] Sum of the BGCs predicted by antiSMASH, also considering the putative ones.

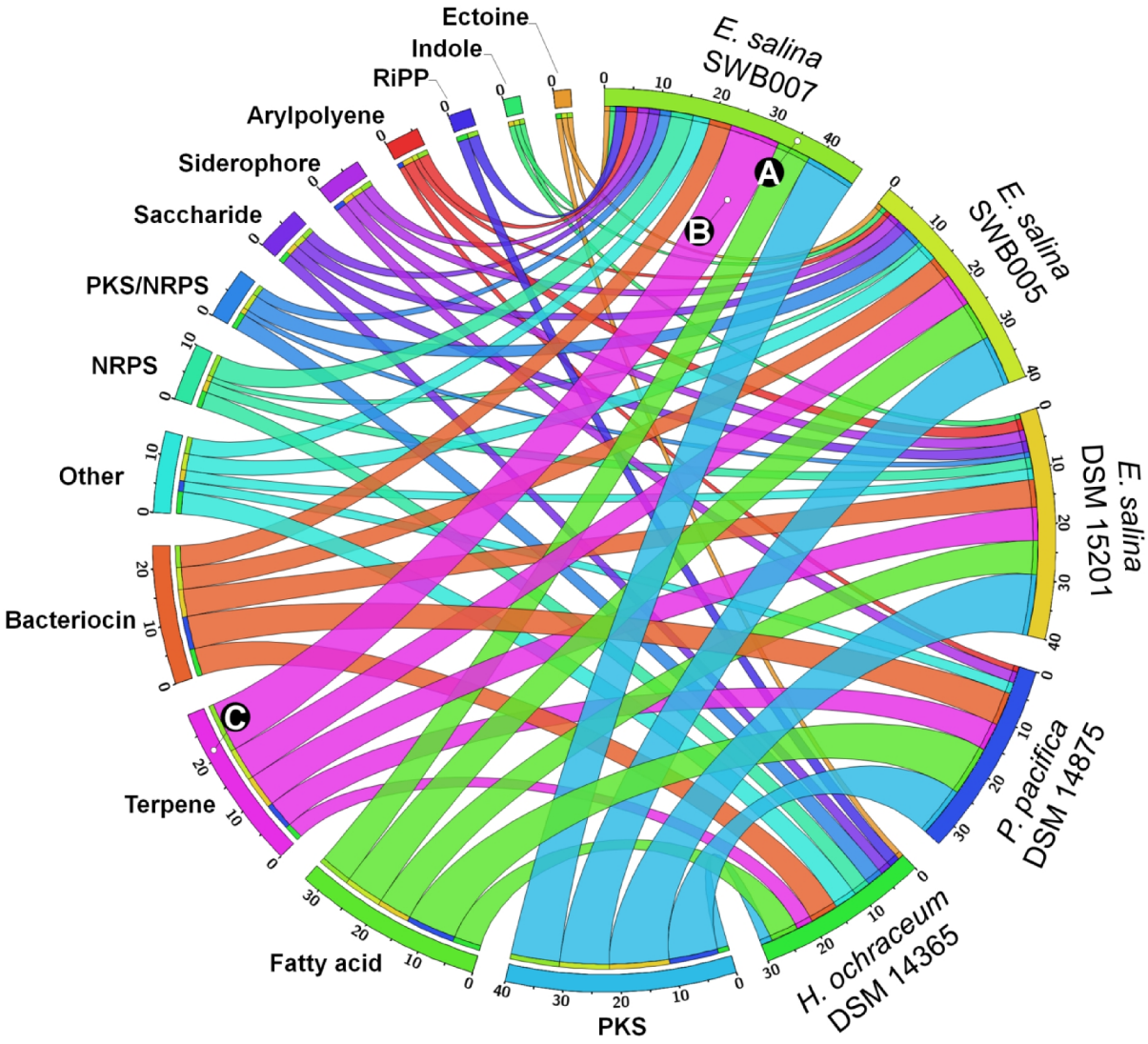


Figure 4.7 Distribution of different BGC types in the five genomes. The percentages of a specific BGC type in a given genome and the contribution of each strain´s genome in a given BGC type are given at the outer ring. A: share of each BGC type in the strain with BGC's color code in descending order. B: color code of the respective strain and number of BGCs in each strain. C: ribbon color is set to respective BGC type. D: BGC color code. E: percentage of the BGCs of each strain with color code of respective strains.

In a next step, we analyzed if BGCs encoding specialized metabolites are shared between the myxobacterial strains. A similarity network of all detected BGCs in the genomes was created based on the Pfam similarity metrics (Ceniceros et al. 2017). Out of the 351 BGCs identified, 124 (35%) can be found in at least two strains (Figure 4.8A). The closely related strains *E. salina* SWB007 and *E. salina* DSM 15201 have the biggest overlap, whereby more than two third (71%) of the BGCs show similarities (Table 4.3A). *E. salina* SWB005 and *P. pacifica* DSM 14875 share several similar gene clusters with at least one other strain in the network (19.3% and 8.9%, respectively). Conversely, *H. ochraceum* DSM 14365 has only one BGC in common with other strains. This BGC is annotated as putatively related to 3-hydroxybutyryl-CoA biosynthesis. In fact, excluding *H. ochraceum,* only seven BGCs equivalent to 9-11% of the BGCs in each strain are similar between all *E. salina* strains and *P. pacifica* (Figure 4.8B). However, 38.7% of the shared similar BGCs are categorized as putative, meaning that no corresponding metabolite class can be predicted (Table 4.3B). From the predictable BGCs, PKS clusters contribute to the highest share with 14.5%, followed by terpene (12.1%), and fatty acid (11.3%) BGCs (Table 4.3B). If only the strain specific (unique) BGCs are considered, half of them (50.7%) are classified as putative. The other half of the unique BGCs can be linked to the biosynthesis of polyketides (9.7%), fatty acids (8.8%), others (6.1%), and further less abundant ones (Figure 4.8A and Table 4.3B). BGCs coding for peptidicmetabolites, e.g. encoding NRPSs, PKSs/NRPSs and RiPPs, are mostly strain specific in the investigated strains.

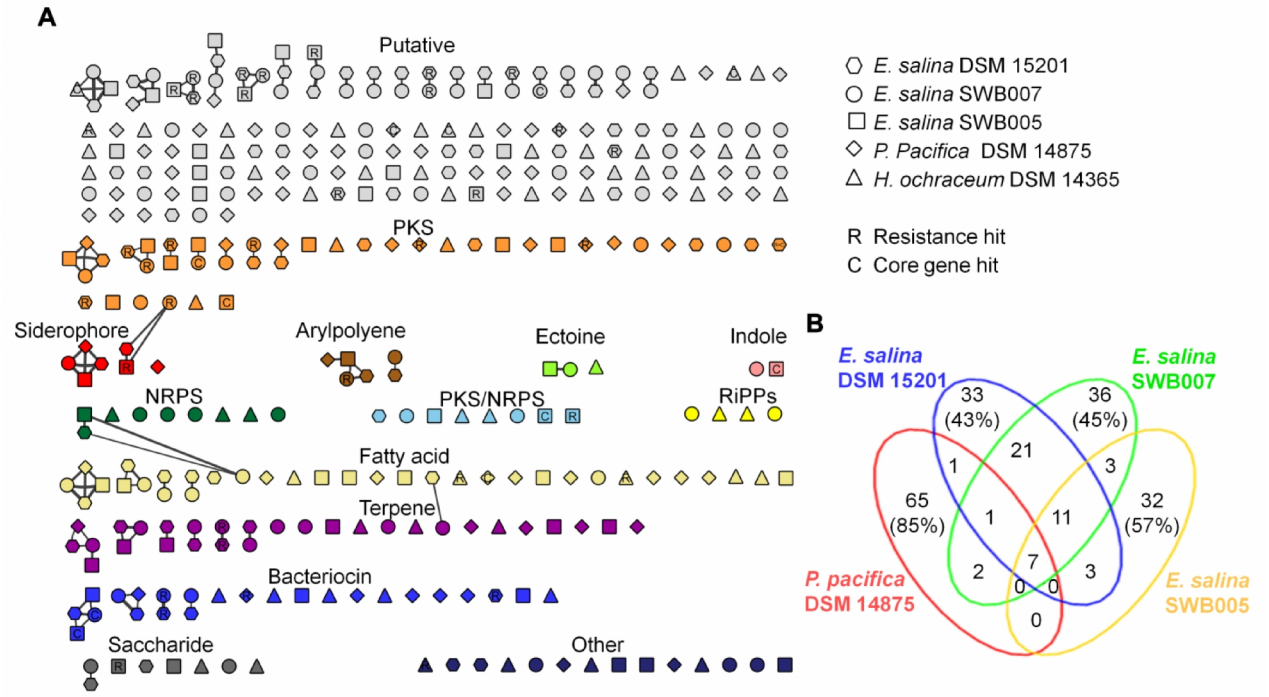


Figure 4.8 Similarity network of the predicted biosynthetic gene clusters (BGCs) in the five analyzed genomes. A: Unique and shared similar BGCs (connected by a line). ARTS hits for resistance (R) and essential core genes (C) are labeled inside the respective nodes. B: Venn diagram displaying node counts according to distribution in strains (*H. ochraceum* is excluded, since this strain has only 1 BGC which is similar to a BGC of the other strains).

Table 4.3 Contribution of the unique and shared similar BGCs in the similarity network (Figure 4.8A) based on strain (A) and cluster family (B).

A: Strain	**Shared**	**(%)**	**Unique**	**(%)**
E. salina DSM 15201	44	35.5	33	14.5
E. salina SWB007	44	35.5	36	15.9
E. salina SWB005	24	19.3	32	14.1
P. Pacifica DSM 14875	11	8.9	65	28.6
H. ochraceum DSM 14365	1	0.8	61	26.9
Total	**124**	**100**	**227**	**100**

B: Cluster family	**Shared**	**(%)**	**Unique**	**(%)**
Putative	48	38.7	115	50.7
PKS	18	14.5	22	9.7
Terpene	15	12.1	13	5.7
Fatty acid	14	11.3	20	8.8
Bacteriocin	11	8.9	13	5.7
Other	0	0	14	6.1
NRPS	2	1.6	7	3.1
Saccharide	2	1.6	6	2.7
Siderophore	6	4.8	1	0.4
PKS/NRPS	0	0	9	4
Arylpolyene	6	4.9	0	0
RiPPs	0	0	4	1.8
Indole	0	0	2	0.9
Ectoine	2	1.6	1	0.4
Total	**124**	**100**	**227**	**100**

In a next step, the biosynthetic pathways, which were predicted by the similarity network analysis to be present in the strains, were analyzed in deeper detail, e.g. if all enzymes necessary for the biosynthesis of the products and relevant precursors are encoded in the genomes.

4.2.4.1 Terpenes

Many of the shared specialized metabolite BGCs correspond to terpene biosynthesis. The *E. salina* strains harbor six to nine terpene BGCs, *P. pacifica* DSM 14875 five and *H. ochraceum* DSM 14365 only three. In silico metabolic analysis using RAST revealed that all the strains harbor the potential to generate the building blocks necessary for terpene assembly. This can be achieved either through the mevalonate pathway, or by degradation pathways, *i.e.* of valine, leucine, isoleucine and the acyclic terpene degradation pathway (Figure 4.9, 4.10, and 4.11). The precursors isopentenyl-PP, geranyl-PP, farnesyl-PP and geranylgeranyl-PP can be expected to arise from the mevalonate pathway (Figure 4.9). By degradation of valine, leucine and isoleucine, isovaleryl coenzyme A (3-methylbutanoyl-CoA), 3-hydroxy-3-methylglutaryl-CoA, acetyl-CoA, acetoacetyl-CoA and methylmalonyl-CoA can be built (Figure 4.10). In addition, a putative BGC linked to the acyclic terpene utilization pathway (atu) can contribute acetate, acetyl-CoA and 7-methyl-3-oxo-6-octenoyl-CoA, whereby the latter yields 3-methylcrotonyl-CoA after two steps of β-oxidation (Figure 4.11). Several of the identified terpene BGCs could be linked to known terpene BGCs, including geosmin, squalene, sterols and carotenoids.

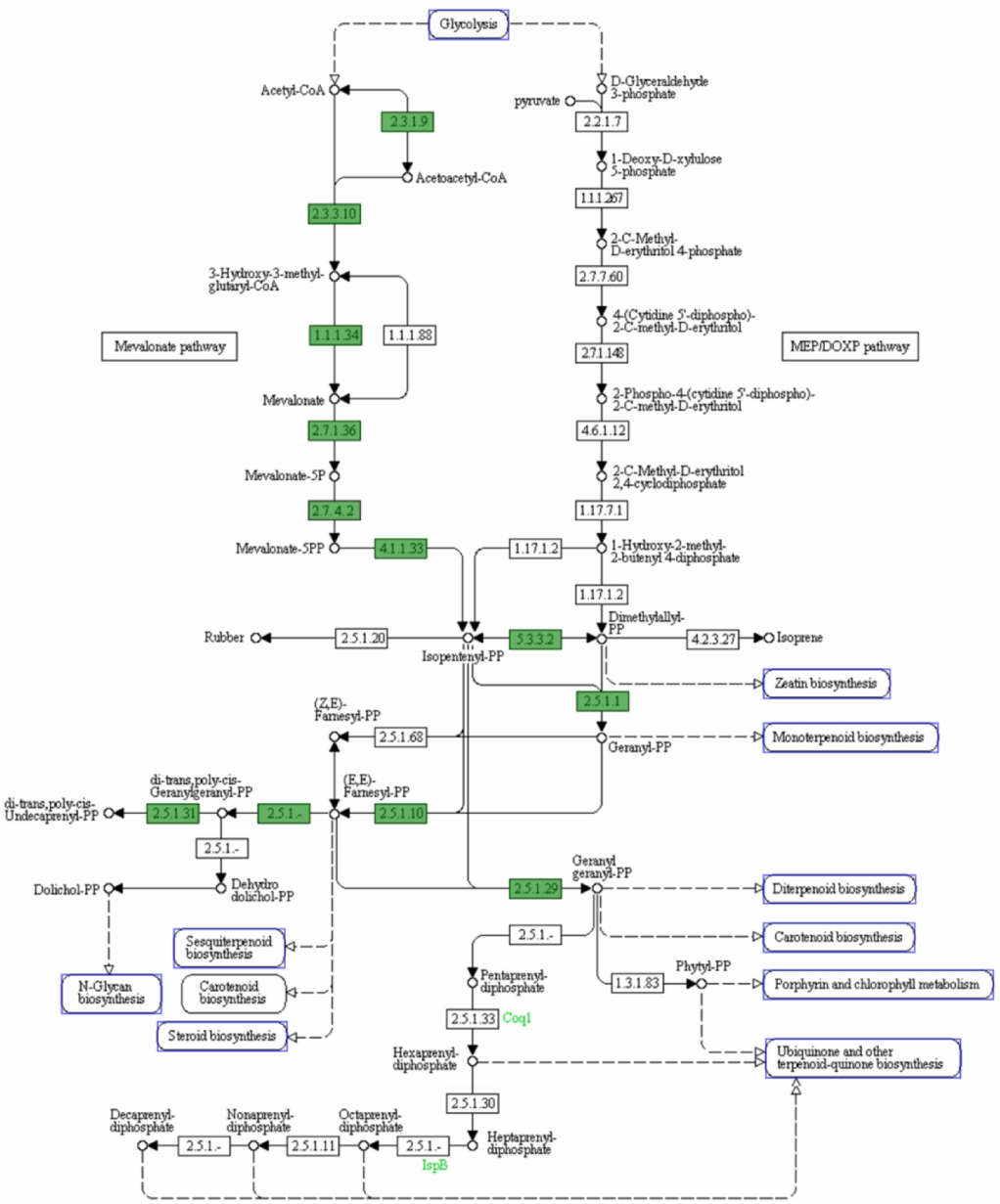


Figure 4.9 Terpenoid backbone biosynthesis pathway in *E. salina* SWB007. Green boxes indicate the presence of at least one feature for the respective reaction. (The KEGG map was compiled using RAST).

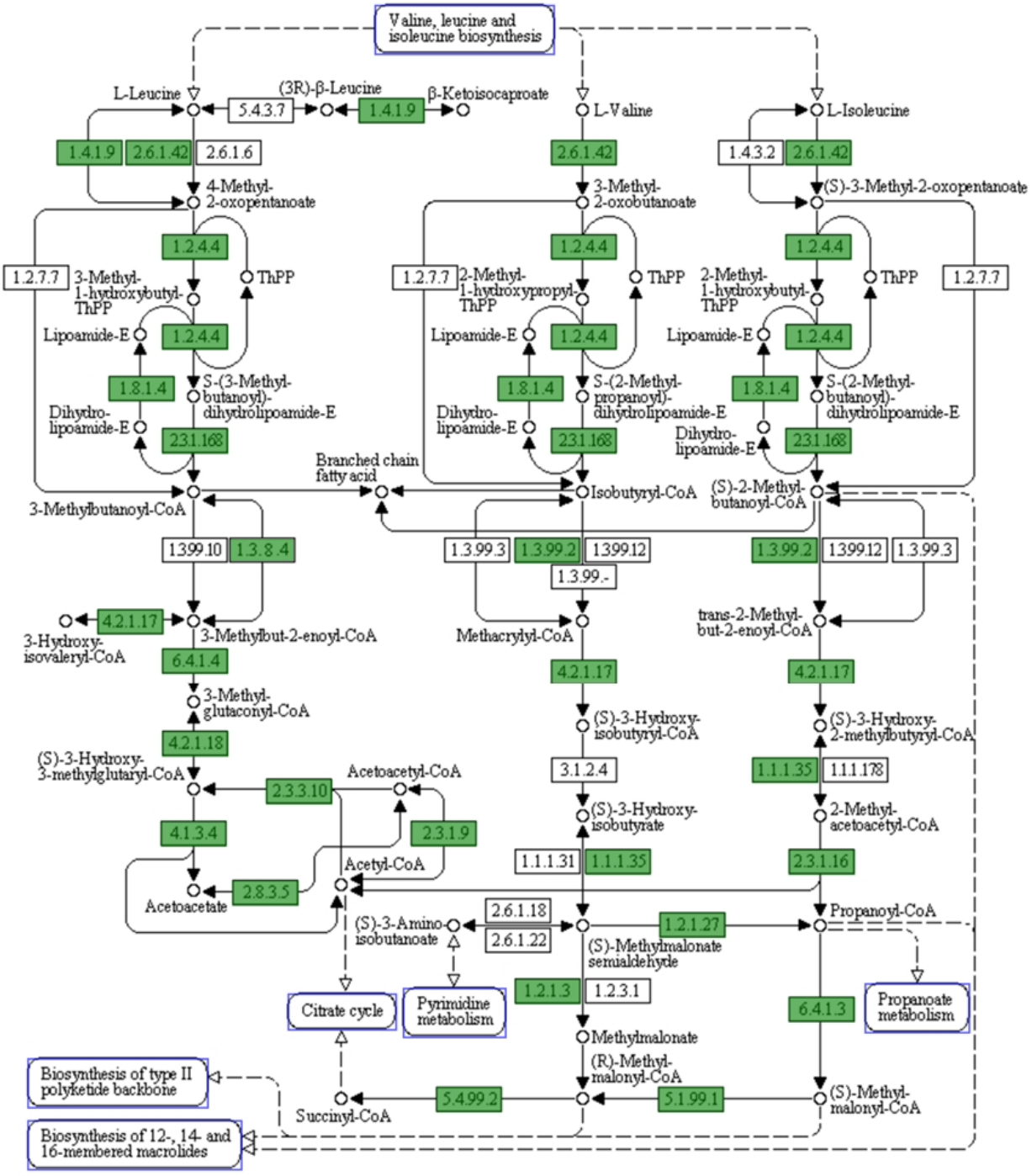


Figure 4.10 Leucine, isoleucine and valine degradation pathway of *E. salina* SWB007. Green boxes indicate the presence of the necessary encoding gene. The gene coding for EC 1.3.8.4 was identified manually by performing a BLAST search on the genome. (The KEGG map was compiled using RAST).

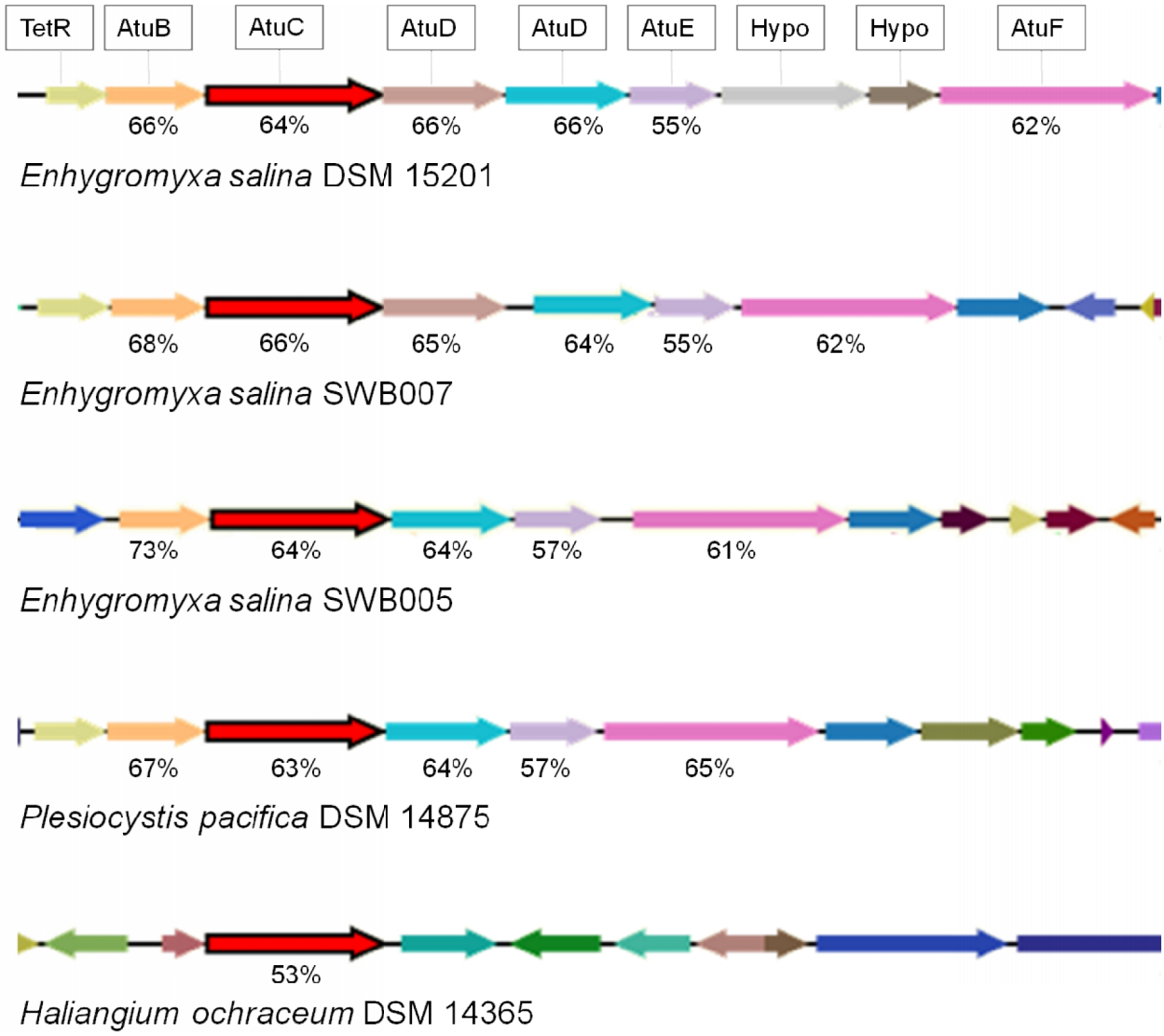


Figure 4.11 Acyclic terpene utilization (atu) gene cluster in the available genomes of marine Myxobacteria. Identities (protein level) towards *Pseudomonas citronellolis* are given below the genes. The proteins encoded by the respective gene are indicated in the black boxes. Similar genes have the same color code. Hypo = Hypothetical protein.

The predicted geosmin BGC shows high similarity to the BGC of *Nostoc punctiforme* PCC 73102 (ATCC 29133), which was investigated before (Giglio et al. 2008). Beside the gene encoding the geosmin synthase/cyclase, two genes encoding transcription regulators were also detected (Figure 4.12). Solely *P. pacifica* DSM 14875 does not harbor this cluster. The same gene cluster can be also found in the closely related halophilic myxobacterium *Nannocystis exedens* ATCC 25963 (Figure 4.12).

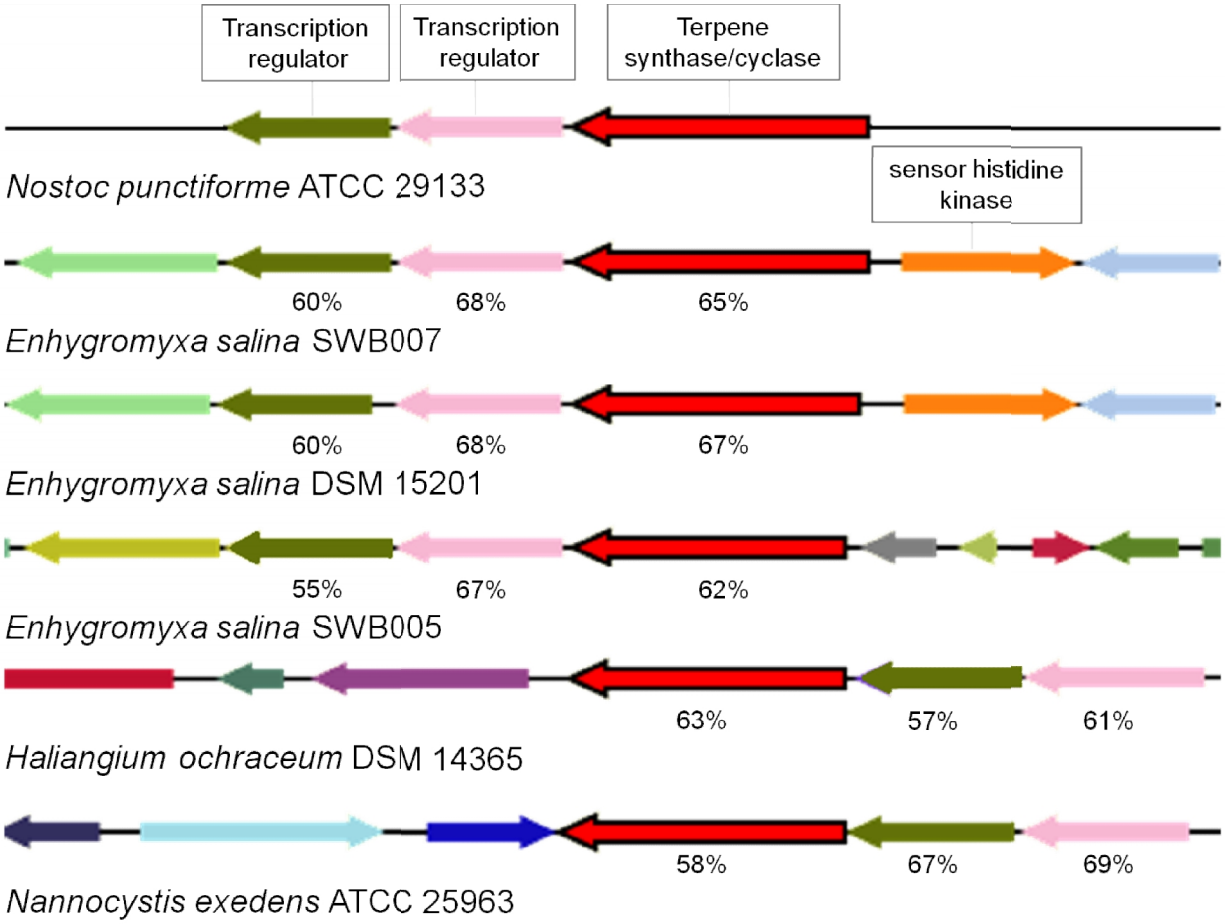


Figure 4.12 Geosmin BGC in available genomes of marine Myxobacteria as well as Nannocystis exedens ATCC 25963. Identities of the genes products to the geosmin gene products of Nostoc punctiforme ATCC 29133 are given below. The same genes have similar color code.

Interestingly, in this bacterium 2-methylisoborneol and geosmin were identified as the main volatile compounds (Dickschat et al. 2007). The squalene BGC was detected in all five investigated strains. This BCG includes two squalene synthases (HpnC and D) and a squalene-associated FAD-dependent desaturase (HpnE), necessary to convert farnesyl diphosphate (FPP) to squalene (Figure 4.13). In addition, the *E. salina* strains harbor three conserved squalene/hopene cyclases in other locations of their genomes, while *P. pacifica* DSM 14875 harbors two of them. The squalene/hopene cyclases

detected in one of the BGC conserved in all *E. salina* strains and *P. pacifica* DSM 14875 showed BLAST hits towards different described sterol synthases including lanosterol and cycloartenol synthases (Figure 4.14). *H. ochraceum* DSM 14365 does not harbor any additional squalene/hopene cyclase. Furthermore, a carotenoid BGC was found to be shared between all investigated strains. The essential genes for geranylgeranyl-CoA diphosphate synthase, a phytoene synthase, two dehydrogenases and a polyprenyltransferase are present (Tian and Hua 2010) (Figure 4.15).

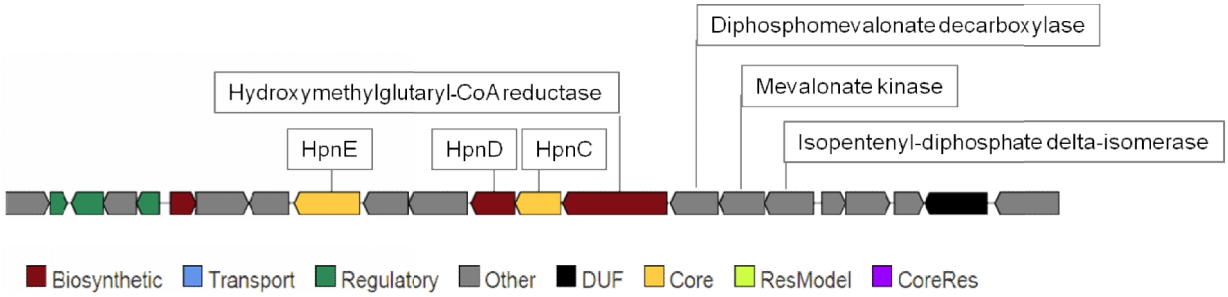


Figure 4.13 Biosynthetic gene cluster (BGC) for squalene biosynthesis in *E. salina* SWB007. The same BGC is present in all five available genomes of marine Myxobacteria .

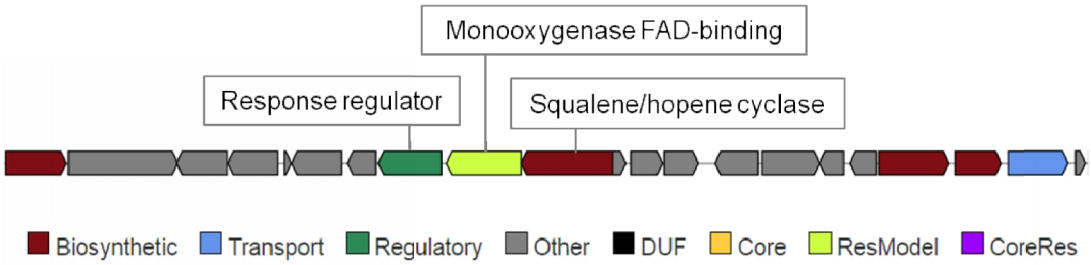


Figure 4.14 Conserved gene cluster for sterol biosynthesis in *E. salina* strains and *P. pacifica* DSM 14875. The genes of *E. salina* SWB007 are shown here as model.

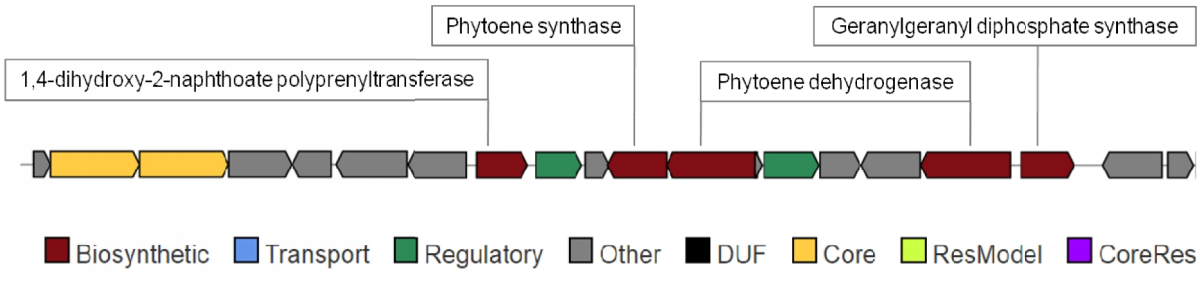


Figure 4.15 Conserved gene cluster for carotenoid biosynthesis in available genomes of marine Myxobacteria. The genes of *E. salina* SWB007 are shown here as model.

4.2.4.2 Polyketides

The biggest group of specialized metabolite BGCs is linked to polyketides, *i.e.* 11.4% of all BGCs (Figure 4.8). The total count of polyketide BGCs obtained by the *in silico* analysis of the draft genomes is 9-11 for *E. salina* strains and *P. pacifica* DSM 14875, while *H. ochraceum* DSM 14365 harbors only two. The genes coding for biosynthesis of starter and extender units for polyketide assembly were identified.

Biosynthesis of extender units for polyketides biosynthesis were investigated using in silico analysis of the genomes. Conserved regions within the genomes of all strains revealed genes coding for the biosynthesis of malonyl-CoA. Therefore, four separate proteins are encoded: 1) Biotin carboxyl carrier protein of acetyl-CoA carboxylase (BCCP), 2) Biotin carboxylase of acetyl-CoA carboxylase (BC), 3) Acetyl-coenzyme A carboxyl transferase (CT) alpha chain and 4) Acetyl-coenzyme A carboxyl transferase (CT) beta chain.

As mentioned in terpene biosynthesis part, methylmalonyl-CoA (mmCoA) can be derived from propionyl-CoA (pCoA), which itself is formed in the catabolism of isoleucine and valine (Figure 4.10). Carboxylation of pCoA to mmCoA can by catalyzed by a CT, which is present in all genomes analyzed. (Annotated as acetyl-coenzyme A carboxyl transferase alpha chain (EC 6.4.1.2)/acetyl-coenzyme A carboxyl transferase beta chain (EC 6.4.1.2)/propionyl-CoA carboxylase (PCC) beta chain (EC 6.4.1.3)).

Beside the standard extender unit malonyl-CoA (mCoA), the results indicate that the strains also possess the potential to synthesize methylmalonyl-CoA (mmCoA) and

propionyl-CoA (pCoA). The latter is formed in the catabolism of isoleucine and valine (Figure 4.10) and can serve as precursor for mmCoA. Ethylmalonyl-CoA (emCoA) can be biosynthesized through carboxylation of butyryl-CoA (bCoA). Carboxylation of bCoA is a described side activity of the propionyl-CoA carboxylase (PCC), which is part of the mmCoA biosynthesis (see above). Another pathway yielding emCoA is the conversion of crotonyl-CoA (cCoA) to emCoA by the catalytic activity of a cCoA carboxylase/reductase (CCR). A gene putatively coding for this conversion was identified in *E. salina* SWB007, *i.e.* annotated as crotonyl-CoA reductase / alcohol dehydrogenase (accession: WP_106090768), 61% identity to Leu10 and 51% identity to TgaD, which are part of leupyrrin and thuggacins BGCs in *Sorangium cellulosum* (Buntin et al. 2010; Kopp et al. 2011). It is of interest that none of the polyketide BGCs in these bacteria could be linked to any known polyketide BGC and they are just partly similar to BGCs of terrestrial Myxobacteria and streptomycetes. For example, a putative type 1 PKS BGC was shared between *E. salina* strains and *P. pacifica* DSM 14875, shows some similarities to the thuggacin BGC from *Chondromyces crocatus (Buntin et al. 2010)* (Figure 4.16). However, the corresponding metabolite to this BGC is unknown.

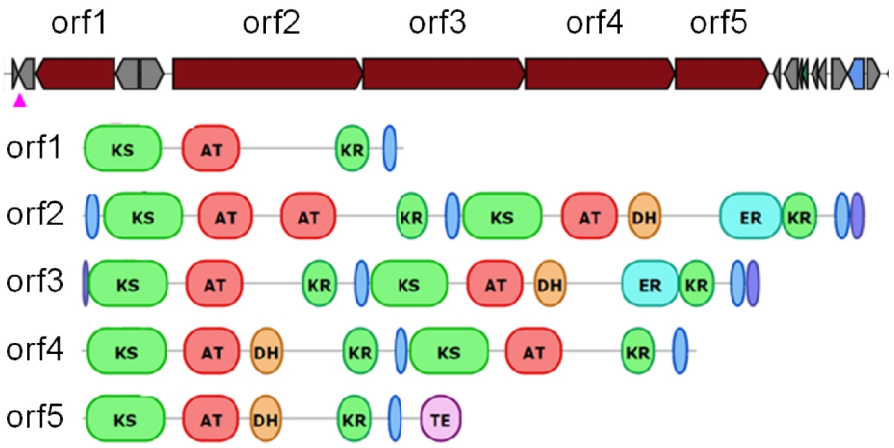


Figure 4.16 Type I PKS gene cluster found in *E. salina* strains and *P. pacifica* DSM14875. The genes of *E. salina* DSM 15201 are shown here as model.

In addition, there are some type III polyketide synthases (PKSIII) BGCs found in all strains except *Haliangium ochraceum* DSM 14365. *P. pacifica* DSM 14875 harbors one and *E. salina* DSM 15201, SWB007, and SWB005 harbor two, three and four PKSIII

BGCs, respectively. One PKSIII BGC is shared between *E. salina* strains and *P. pacifica* DSM 14875, while another PKSIII BGC is shared only between the *E. salina* strains. Furthermore, *E. salina* SWB007 carries a unique PKSIII BGC, consisting of genes encoding a PKSIII, a methyltransferase, and an oxidoreductase. In the vicinity genes for a polyprenyl synthetase and a polyprenyltransferase were detected (Figure 4.17).

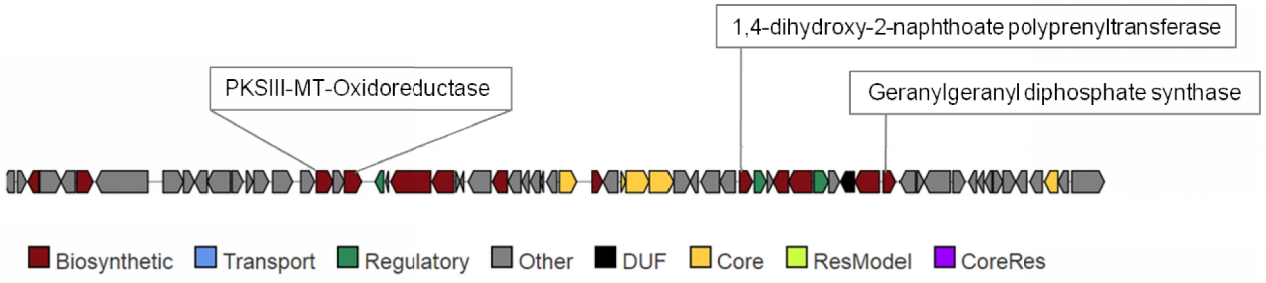
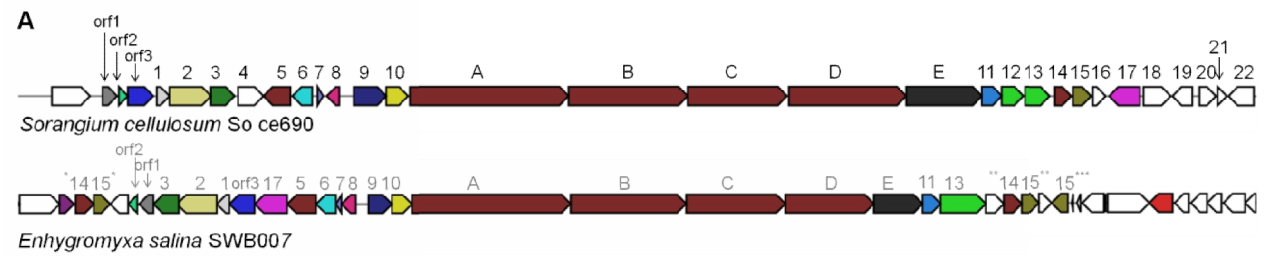


Figure 4.17 PKSIII BGC with genes associated with terpene biosynthesis in close proximity in *E. salina* SWB007.

4.2.4.3 Non-Ribosomal Peptides (NRPS) and PKS/NRPS hybrids

Almost all of NRPS and PKS/NRPS hybrid BGCs were strain specific and only identified in *E. salina* strains and *H. ochraceum* DSM 14365. In *E. salina* SWB007 a strain specific type 1 PKS/NRPS BGC was identified, showing high homology to the reported leupyrrin BGC from *Sorangium cellulosum* So ce690 (Kopp et al. 2011). In depth investigation of the gene cluster revealed that all genes necessary for leupyrrin biosynthesis are present (Figure 4.18).



Protein	Annotation	Functional class	Identity/coverage (%) in *E. salina* SWB007
orf1	Phospholipid/glycerol acyl transferase	-	52/97
orf2	hypothetical protein	-	55/90
orf3	serine/threonine protein kinase	-	48/98
leu1	type II thioesterase	Other	47/91
leu2	formate dehydrogenase	Tailoring (reduction)	62/99
leu3	hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyltransferase	-	39/97
leu4	WS/DGAT/MGAT acyltransferase	Tailoring (reduction)	-
leu5	proline adenylation protein (A)	Precursor biosynthesis	42/98 (A)
leu6	prolyl-ACP dehydrogenase	Precursor biosynthesis	53/98
leu7	acyl carrier protein	Precursor biosynthesis	52/76
leu8	Dehydrogenase (KR)	Tailoring (oxidation)	51/100 (KR)
leu9	C and PCP domain protein	Scaffold biosynthesis	34/105 (C-PCP)
leu10	alcohol dehydrogenase (crotonyl-CoA reductase / alcohol dehydrogenase)	Precursor biosynthesis	61/99 (ER)
leuA	non-ribosomal peptide synthase/polyketide synthase (C-A-PCP-KS-AT-DH-KR-ACP)	Scaffold biosynthesis	49/102 (C-A-PCP-KS-AT-DH-KR-ACP)
leuB	polyketide synthase (KS-AT-DH-cMT-KR-ACP)	Scaffold biosynthesis	40/67 (KS-DH-cMT-KR-ACP)
leuC	polyketide synthase (KS-AT-DH-KR-ACP)	Scaffold biosynthesis	48/101 (KS-AT-DH-KR-ACP)
leuD	polyketide synthase (KS-AT-DH-cMT-KR-ACP)	Scaffold biosynthesis	59/31 (KS-DH-cMT-KR-ACP)
leuE	non-ribosomal peptide synthase (C-A-PCP-TE)	Scaffold biosynthesis	36/64 (C-PCP-TE)
leu11	phospholipase	-	41/101
leu12	geranylgeranyl reductase	Other	44/46
leu13	cytochrome P450	Tailoring (oxidation)	60/52
leu14	methylase type II O-methyltransferase	Tailoring (methylation)	46/90 (14*) and 49/84(14**)
leu15	Rieske (2Fe-2S) domain-containing protein	-	65/100 (15*), 64/96 (15**) and 44/34 (15***)
leu16	TetR family transcriptional regulator	-	-
leu17	very-long-chain acyl-CoA synthetase (A)	Activation / processing	50/94 (A)
leu18	carotenoid oxygenase	-	-
leu19	HpcH/HpaI aldolase	-	-
leu20	HpcH/HpaI aldolase	-	-
leu21	dehydratase	-	-
leu22	cytochrome P450 hydroxylase	-	-

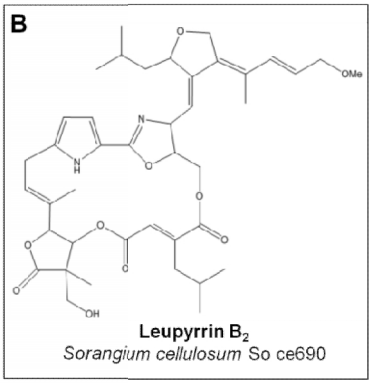


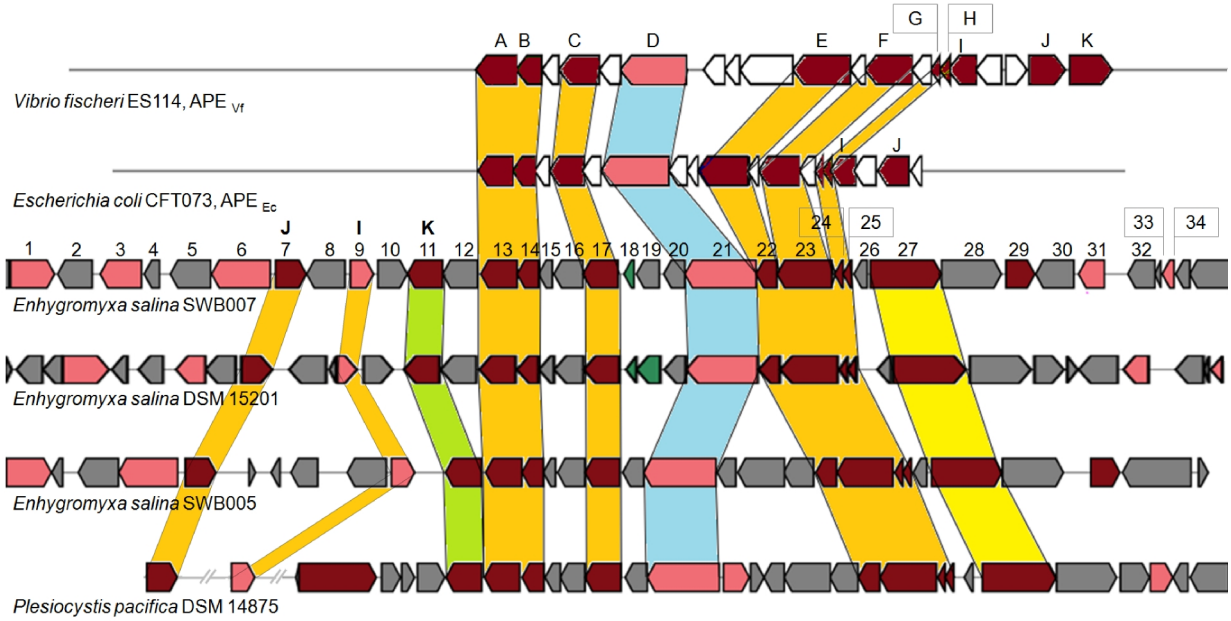
Figure 4.18 A: Leupyrrin biosynthetic gene cluster in *E. salina* SWB007 and *Sorangium cellulosum*. Proteins encoded in the leupyrrin BGC of the *Sorangium cellulosum* So ce690, their functional class and identity/coverage in *E. salina* SWB007. Same colors/numbers indicate similar annotation. B: Structure of Leupyrrin B_2 from *Sorangium cellulosum* So ce690.

4.2.4.4 Bacteriocins

Several bacteriocin BGCs were identified in each strain. Thereby, the similarity network (Figure 4.8) indicated many of the them to be similar BGCs, i.e. 11 out of 19 have at least one counterpart if the *E. salina* strains and *P. pacifica* DSM 14875 are considered.

4.2.4.5 Arylpolyenes

Arylpolyene (APE) BGCs were detected only in *E. salina* strains and *P. pacifica* DSM 14875. One of them is well conserved within all with homologous gene clusters from different marine photobacterium strains and resembles the APE BGC of *Escherichia coli* CFT073 and of *Vibrio fischeri* ES114 (100% of the biosynthetic genes show similarity, Figure 4.19)(Cimermancic et al. 2014). Another APE BGC was only found in *E. salina* SWB007 and *E. salina* DSM 15201. However, the latter one did not show high similarity to any known BGCs.



APE $_{Vf, Ec}$ Annotation of biosynthetic genes		Protein	Annotation in available marine myxobacteria
		orf1	pyridine nucleotide-disulfide oxidoreductase
		orf3	pseudouridine synthase
		orf6	1-deoxy-D-xylulose-5-phosphate synthase
J	O-methyltransferase	orf7	biotin synthase
I	acyltransferase	orf9	acyltransferase
K	halogenase	orf11	halogenase
A	Beta-ketoacyl synthase	orf13	Beta-ketoacyl synthase
B	short-chain dehydrogenase/reductase SDR	orf14	short-chain dehydrogenase/reductase SDR
		orf15	beta-hydroxyacyl-[acyl carrier protein] dehydratase FabZ
		orf16	Beta-ketoacyl synthase
C	Beta-ketoacyl synthase	orf17	Beta-ketoacyl synthase
D	MMPL domain-containing transport protein	orf21	MMPL domain-containing transport protein
E	polyprenol-monophosphomannose synthase ppm1	orf22	polyprenol-monophosphomannose synthase ppm1
F	AMP-dependent synthetase and ligase	orf23	AMP-dependent synthetase and ligase
G	putative acyl carrier protein	orf24	phosphopantetheine-binding domain-containing protein
H	putative acyl carrier protein	orf25	acyl carrier protein
		orf27	Peptidase_S9
		orf29	NAD+ synthase
		orf31	acetylglutamate kinase
		orf34	thioesterase superfamily protein

Figure 4.19 Arylpolyene (APE) BGCs in *E. salina* strains, *P. pacifica* DSM 14875 and the homologous known BGCs of *Vibrio fischeri* ES114 and *Escherichia coli* CFT073. Dark red indicates the biosynthetic genes and pink indicates the transporter genes. Green presents the regulatory genes. White and grey indicate other genes.

4.2.4.6 Siderophores

Siderophore BGCs (NRPS-independent) were only shared between the *E. salina* strains and *P. pacifica* DSM 14875. Each strain harbors two distinct siderophore BGCs. One of

them has only one conserved gene from IucA/IucC family of siderophore biosynthesis and the other has two IucA/IucC genes and a lysine/ornithine N-monooxygenase.

4.2.4.7 Ectoine and hydroxyectoine

A complete ectoine/hydroxyectoine BGC was detected only in *E. salina* SWB005 and SWB007. In *H. ochraceum* DSM 14365 only an ectoine synthase gene was detected, while all the other necessary genes were absent. In addition, the ectoine BGC in *E. salina* SWB007 contains a glycine/sarcosine N-methyltransferase (GSMT) and a sarcosine/dimethylglycine N-methyltransferase (SDMT), which are responsible for betaine biosynthesis (Figure 4.20)(Amiri Moghaddam et al. 2016).

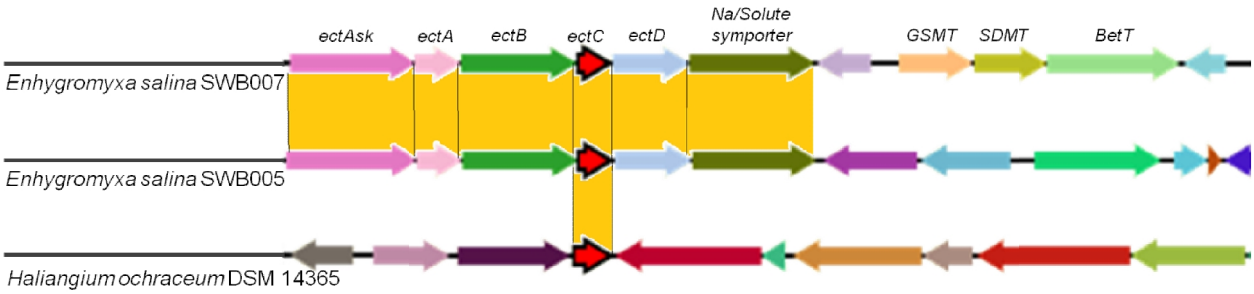


Figure 4.20 Ectoine/hydroxyectoine BGC in *E. salina* SWB005, *E. salina* SWB007 and *H. ochraceum*. Genes encode for: ectAsk: aspartokinase, ectA: diaminobutyric acid (DABA) acetyltransferase, ectB: DABA aminotransferase, ectC: ectoine synthase, ectD: ectoine hydroxylas, *Na/solute symporter*, *GSMT*: glycine/sarcosine N-methyltransferase, *SDMT*: sarcosine/dimethylglycine N-methyltransferase, *BetT*: high affinity choline uptake transporter *BetT*. Same colors indicate similar annotation.

4.2.4.8 Indole

All *E. salina* strains harbor a conserved indole prenyltransferase. However, the adjacent genes are either rearranged or not conserved (Figure 4.21).

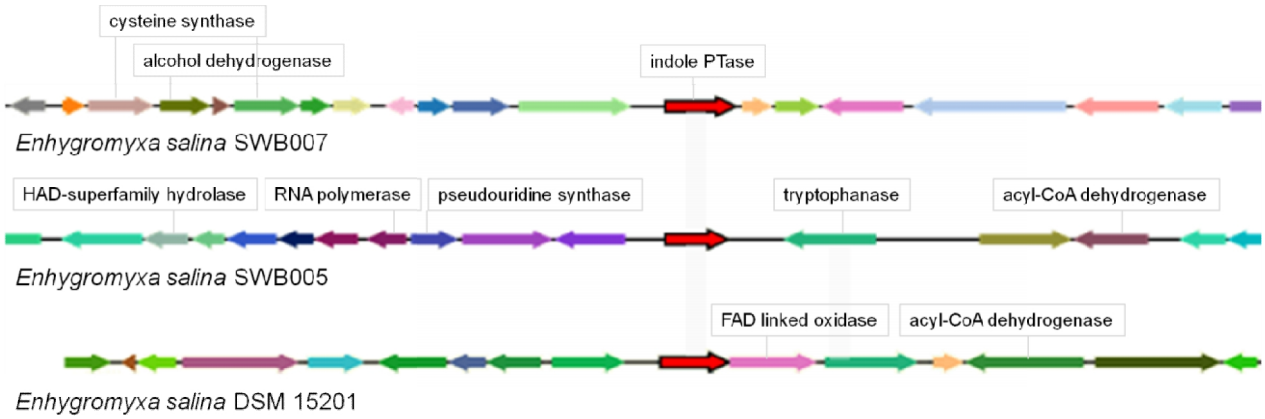


Figure 4.21 Genomic locus adjacent to the indole prenyltransferase in *E. salina* strains. The same genes have similar color code.

4.2.4.9 Ribosomally synthesized and post-translationally modified peptides (RiPPs)

BGCs coding for RiPPs were only found as unique BGCs in the genomes of *E. salina* SWB007 and *H. ochraceum* DSM 14365. A lanthipeptide and a thiopeptide BGC were detected in the genome of *E. salina* SWB007, and in *H. ochraceum* DSM 14365 a lanthipeptide and a lassopeptide BGC were detected.

4.2.4.10 Putative gene clusters

Many of the putative BGCs (29%) were shared as similar BGCs between *E. salina* strains and *P. pacifica* DSM 14875. They are mostly related to the biosynthesis of primary metabolites, such as a BGC putatively linked to the production of 3-crotonyl-CoA and 3-hydroxybutyryl-CoA. In addition, a conserved PHB synthase identified in *E. salina* strains and *P. pacifica* DSM 14875 are probably involved in the synthesis of polyhydroxybutyrate (PHB) from 3-hydroxybutyryl-CoA (Figure 4.22).

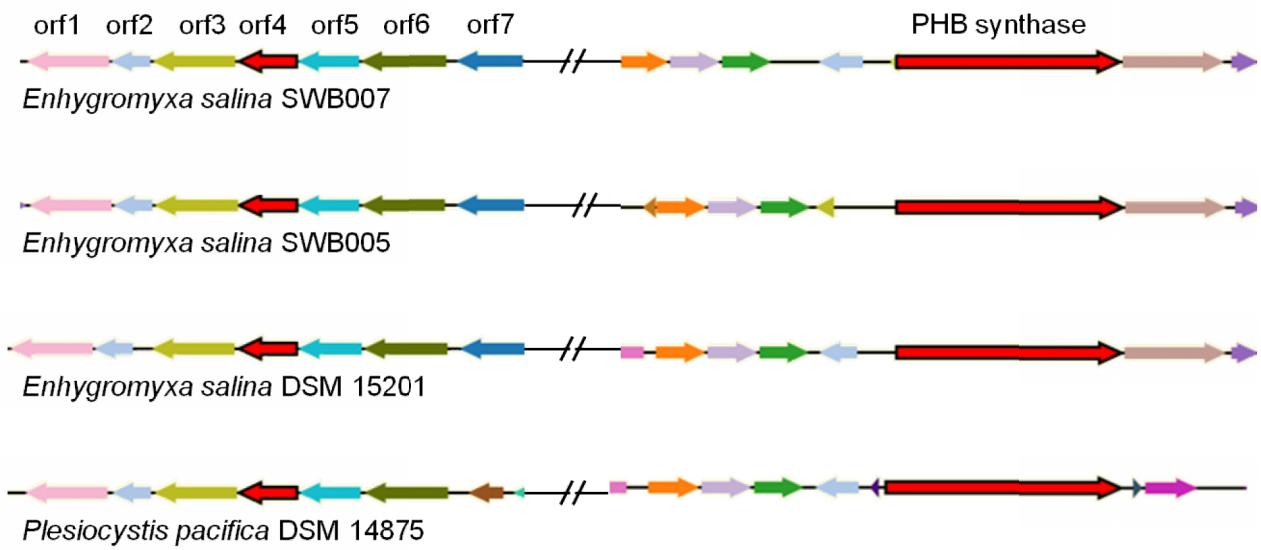


Figure 4.22 Conserved BGC for polyhydroxybutyrate biosynthesis and production of 3-crotonyl-CoA from 3-hydroxybutyryl-CoA. orf1: butyryl-CoA dehydrogenase (EC 1.3.99.2), orf2: hypothetical protein, orf3: acyl-CoA dehydrogenase, short-chain specific (EC 1.3.99.2), *orf4*: 3-hydroxybutyryl-CoA dehydratase (EC 4.2.1.55), *orf5*: 3-hydroxybutyryl-CoA dehydrogenase (EC 1.1.1.157), *orf6*: acetyl-CoA acetyltransferase (EC 2.3.1.9), *orf7*: alpha/beta hydrolase fold-1 precursor and PHB synthase: polyhydroxybutyrate synthase.

4.2.5 Metabolomic analysis of four marine myxobacterial strains

Next, we aimed to analyze and compare the metabolomes of the marine myxobacterial strains, in order to see if the bioinformatics results are translatable into actual metabolites. For this type of analysis, the more closely related strains *E. salina* SWB005, SWB007, DSM15201 and *P. pacifica* DSM 14875 were selected. The strains were cultivated in the liquid medium containing adsorber resin and subsequently extracted and fractionated. The crude extracts and all fractions were analyzed by HPLC coupled with high-resolution mass spectrometry and automated fragmentation (HPLC-HRMS/MS). The resulting MS^2 data were used to generate a molecular network consisting of 1251 nodes after removal of media blanks (Figure 4.23A). In accordance with the high abundance of strain-specific BGCs, also in the metabolomics analysis many strain specific nodes, corresponding to ions, were identified. The result was summarized in a Venn diagram (Figure 4.23B). Of the four strains, *E. salina* SWB005 and DSM15201 display the highest metabolic diversity with 584 and 556 nodes,

respectively. Interestingly, all four strains show a relatively high percentage of strain-specific nodes, *i.e.* nodes that were only found in one strain. The most unique metabolome shows *E. salina* SWB007, where almost half of all nodes (173 of 343) were found to be strain-specific. Surprisingly, only 6-11% of the nodes in each strain were shared in the network by all four strains. Taken together, this analysis points to a large degree of unique metabolism in all four investigated strains under laboratory conditions.

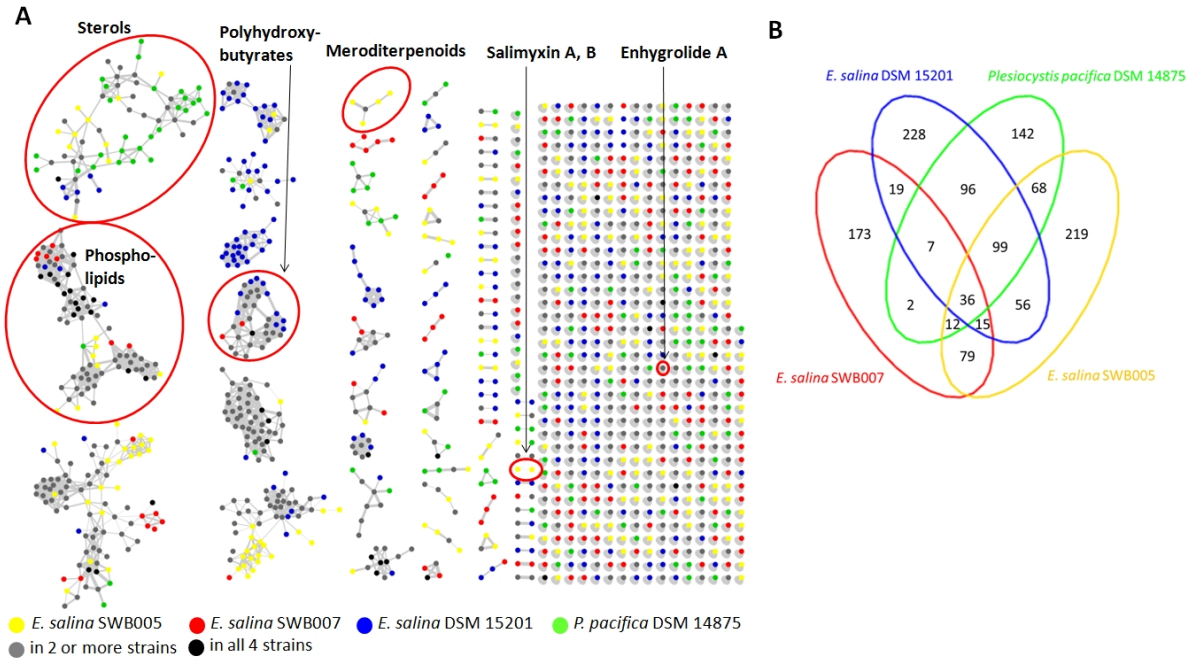


Figure 4.23 A: Molecular network of *E.salina* SWB005, SWB007, DSM15201 and *P. pacifica* DSM 14875 extracts. Network is color-coded according to detection from single or multiple strains. Identified specialized metabolites are marked. B: Venn diagram displaying node counts according to distribution in strains.

Only a few nodes in the network could be dereplicated as specialized metabolites using the GNPS and our metabolite libraries (Table 4.4). Salimyxin A and salimyxin B were previously isolated from *E. salina* (Felder et al. 2013b). Both compounds were detected as strain specific metabolites of *E. salina* SWB005 in this analysis. Retention time and exact mass of all compounds correspond to an authentic standard. Enhygrolide A (Felder et al. 2013b), was found in extracts from *E. salina* SWB005 and SWB007.

Other compound clusters in the network could be dereplicated with the help of the GNPS library search tool. These include a number of ions annotated as triterpenes/sterols and a large group of phospholipid-related molecules. With the DEREPLICATOR+ tool available on the GNPS platform (Mohimani et al. 2017), one metabolite cluster produced by *E. salina* SWB005 and *P. pacifica* could be annotated with high confidence as meroditerpenoids related to tetraprenyltoluquinols isolated from marine algae (Amico et al. 1985).

Table 4.4 Overview of the dereplicated compounds in this study

Compound name	Molecular family	Analogues	Molecular formula	m/z measured	m/z theor	Adduct	GNPS score	GNPS Library ID	strains detected
Enhygrolide A	Enhygrolides	1	$C_{12}H_{22}O_3$	335.17	335.17	$[M+H]^+$	0.76	CCMSLIB00000839211	*E. salina* SWB005,SWB007
Salimyxin A	Salimyxins	2	$C_{21}H_{30}O_4$	347.22	347.22	$[M+H]^+$		CCMSLIB00000579942	*E.salina* SWB005
Salimyxin B	Salimyxins	2	$C_{21}H_{32}O_4$	349.42	349.24	$[M+H]^+$		CCMSLIB00000839209	*E.salina* SWB005
1-Hexadecanoyl-2-octadecadienoyl-sn-glycero-3-phosphocholine	Phospholipids	80	$C_{42}H_{80}NO_8P$	758.55	758.57	$[M+H]^+$	0.84	CCMSLIB00003135073	*E. salina* DSM15201
Ergosterole Endoperoxide	Sterols	68	$C_{28}H_{44}O_3$	411.33	411.32	$[M-H_2O+H]^+$	0.79	CCMSLIB00000848042	*E. salina* SWB005, DSM15201, *P. pacifica*
Polyhydroxybutyrate (Fragments)	Polyhydroxyalkanoates, biopolymer	34	$[COCH_2CH(CH_3)O]n$				dereplicated by MS pattern, delta mass and fragmentaion analysis		*E. salina* SWB005, SWB007, DSM15201, *P. pacifica*
(6R,7S,10E,14E)-16-(2-hydroxy-5-methoxy-3-methylphenyl)-2,6,10,14-tetramethylhexadeca-2,10,14-triene-6,7-diol (IUPAC)	Linear meroterpenoids	5	$C_{28}H_{44}O_4$	445.33	445.33	$[M+H]^+$	dereplicated by DEREPLICATOR+ analysis		*E. salina* SWB005, *P. pacifica*

A large metabolite cluster with a mass range between 883.3554-1332.5815 *m/z* displayed several characteristic mass shifts of 86.04 Da, which correspond to the loss or gain of hydroxybutyric acid (Figure 4.24A). In addition, in the MS^2 spectra of these compounds, several hydroxybutyric acid mass shifts were observed (Figure 4.24B). Thus, we conclude that this metabolite cluster consists of different molecular weight fragments of the polymer polyhydroxybutyric acid (PHB), which is produced by all the strains. These biopolymers gained interest due to their biodegradability, biocompatibility, the possibility of biosynthesis from renewable resources, and similar physical and chemical characteristics to the ones of petrochemical polymers (Takahashi et al. 2017).

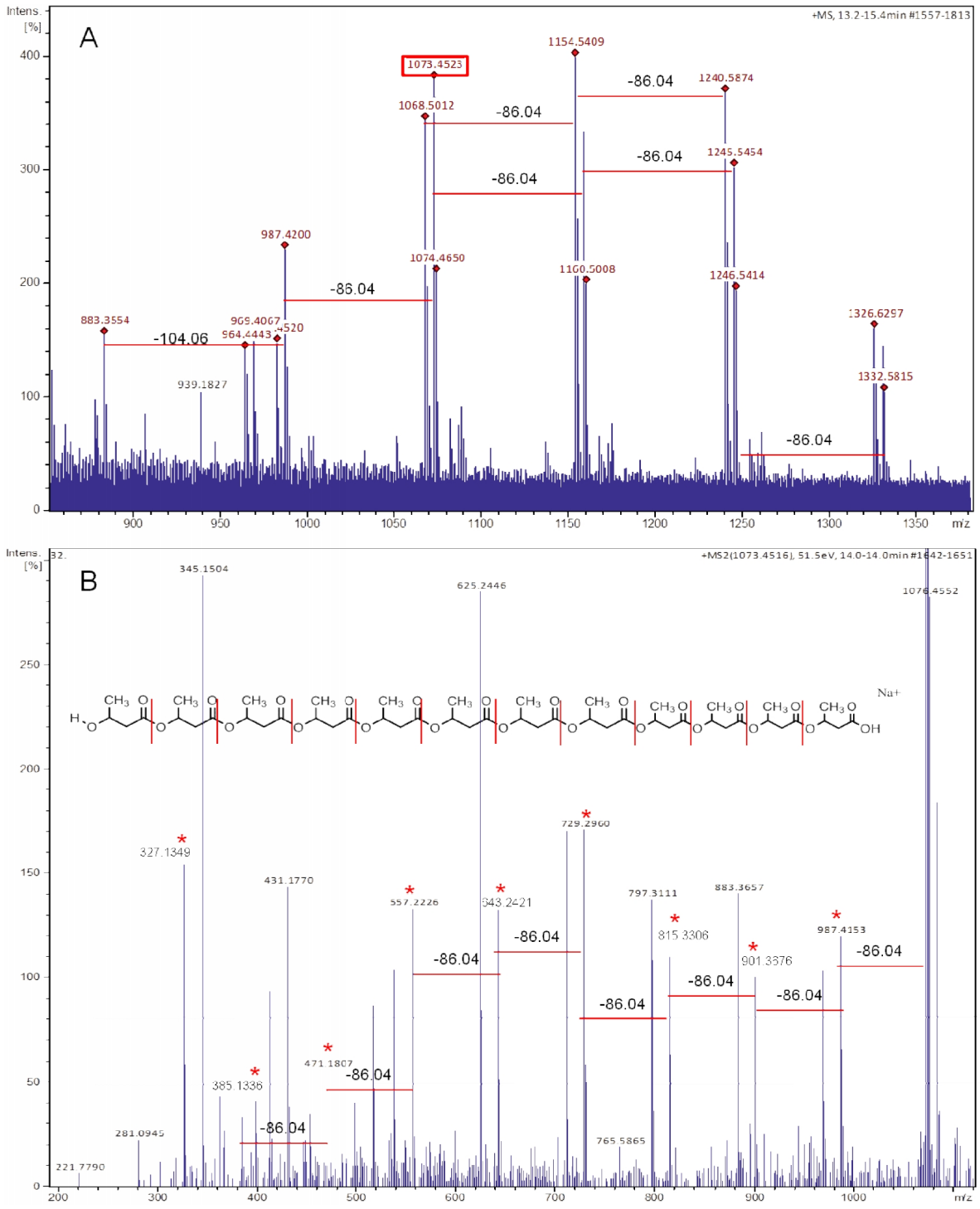


Figure 4.24 A: Mass spectra of different length of polyhydroxybutyric acid (PHB) found in *E. salina* SWB007 with characteristic mass shifts of 86.04 Da. B: MS2 spectra of PHB (N=12) with m/z = 1073.4516 and characteristic mass shifts of 86.04 Da.

In order to verify the metabolomic results, the completely uninvestigated *E. salina* strain SWB006 was included to the investigation. Hence, this strain was fermented, extracted

and its metabolomics profile analyzed. Also in this strain, enhygrolide A was identified (Figure 4.25). In order to verify the metabolomics results, large scale cultivation of this strain was required to enable isolation of the compound. By this approach, enhygrolide A was isolated and its structure confirmed by NMR spectroscopy (Table 9.1, Figure 9.1, and Figure 9.2).

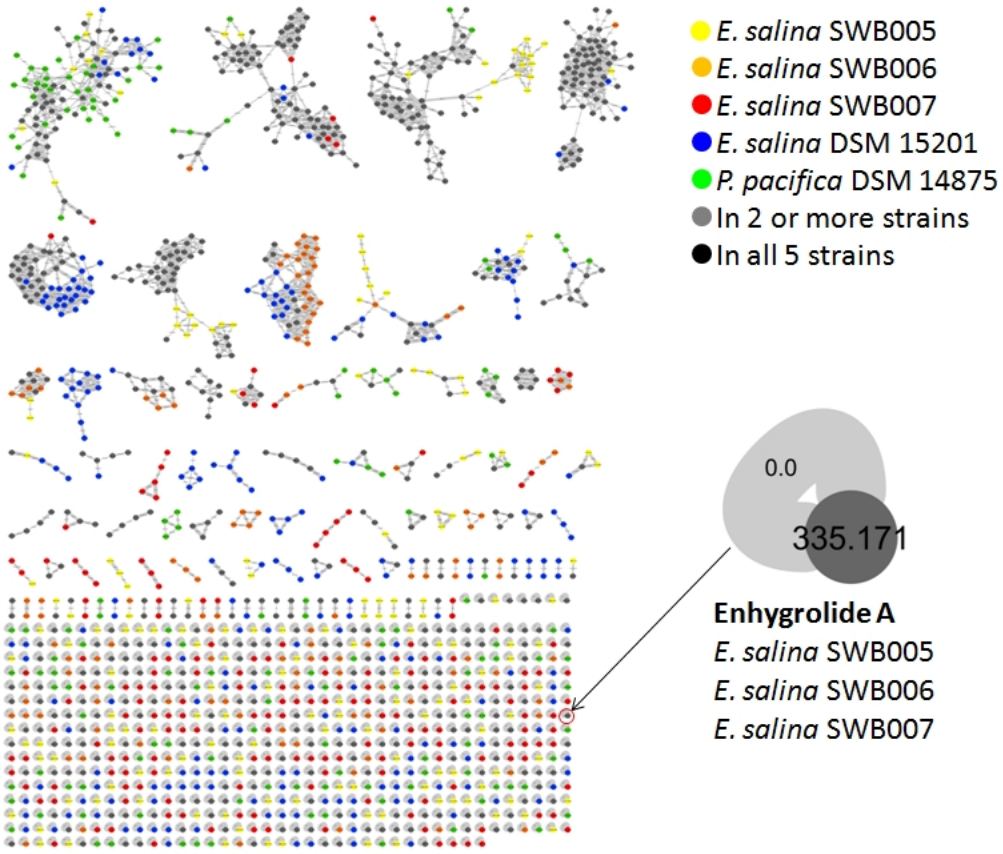


Figure 4.25 Molecular network of *E.salina* SWB005, SWB006, SWB007, DSM15217 and *P. pacifica* DSM 14875 extracts. Network is color-coded according to detection from single or multiple strains. Enhygrolide A is marked (identified mass was grouped for *E. salina* SWB005, SWB006 and SWB007).

4.3 **Discussion I**

Obligate marine myxobacteria have been discovered only recently compared to their terrestrial counterparts. Since then, a small number of marine myxobacterial strains and specialized metabolites were isolated (Albataineh and Stevens 2018). However, by metagenomics approaches, 16S rRNA gene sequences of marine myxobacteria were identified from sediments of different locations, depths, and climatic regions, indicating that they are widely distributed around the globe. This suggests that the vast majority of marine myxobacteria has yet to be discovered. Furthermore, they are separated from terrestrial myxobacteria at high levels of classification (Jiang et al. 2010; Brinkhoff et al. 2012). This indicates a high chance for the discovery of novel chemical scaffolds, since recently a correlation between taxonomic distance and the production of distinct secondary metabolite families was proven (Hoffmann et al. 2018). Therefore, marine myxobacteria should be a bioresource for novel specialized metabolites because their terrestrial counterparts are one of the prime sources of these bioactive compounds (Brinkhoff et al. 2012; Albataineh and Stevens 2018).

Similar to other marine Deltaproteobacteria, marine myxobacteria can be isolated from samples taken from benthic ecosystems such as sediments, sea weeds, sea grasses and aggregates close to the sediment surface (Zinger et al. 2011; Stevens et al. 2005; Iizuka et al. 1998). However, to date the cultivation of marine Myxobacteria lags far behind to terrestrial ones. One main obstacle to their isolation is the slow growth with the consequence that marine myxobacteria are easily overgrown by other faster growing bacteria. Another problem is, that usually more than one cell is needed for these social bacteria to start growing on agar plates and they usually prefer media poor in nutrients (Schäberle et al. 2010).

Here, we could show by comparative genomic analysis that the marine-derived species harbor an enormous potential for the discovery of novel natural products. The five available genomes of marine myxobacteria revealed that a relatively large portion of the genome (~10%) is dedicated to various classes of BGCs, corresponding to the production of specialized metabolites (Dávila-Céspedes et al. 2016).

The five marine myxobacteria from the family Nannocystaceae, for which genome information is available, are related to each other as evidenced by a conserved core

genome. However, *in silico* parameters, *i.e.* ANI, isDDH, and difference of the GC content, clearly indicate that all *E. salina* strains investigated in this work should be classified as different species. In fact, it seems that significant parts of the genomes are either from different ancestral origin or have diverged rapidly. The same situation was observed in terrestrial myxobacteria, which show a large variation in their genomes and a small core genome (Zaburannyi et al. 2016; Huntley et al. 2011).

For the unique BGCs of the marine strains, the corresponding metabolites are so far unknown. However, the observation that BGCs related to PK and terpene biosynthesis represent the most abundant BGC types, is in line with the fact that most of the compounds isolated so far from myxobacteria, terrestrial or marine ones, are terpenoids, PKs, NRPs and PK/NRP hybrids (Herrmann et al. 2017; Yamada et al. 2015; Desmond and Gribaldo 2009; Dickschat et al. 2007; Iniesta et al. 2007; Dávila-Céspedes et al. 2016; Wei et al. 2016).

Myxobacteria, along with actinobacteria and cyanobacteria harbor the majority of the annotated terpene synthases among all bacteria (Yamada et al. 2015). Many terpenes are volatile compounds and might play a communication role during the multicellular life stages in myxobacteria (Dickschat et al. 2007). Interestingly, conserved terpene BGCs of the marine strains can be attributed to different classes of terpenoids, e.g. carotenoids, sterols, and geosmin. The latter compound was thought to be indicative for terrestrial strains and was unexpected to be present in the genomes of the marine strains. Several sterols like lanosterol, cycloartenol and zymosterol were already reported from *E. salina* DSM15201 and *P. pacifica* DSM 14875 (Wei et al. 2016), and also the metabolomic analysis indicates a variety of sterols synthesized by these strains. The presence of the squalene BGC in all investigated strains emphasizes the importance of this compound as an intermediate in the biosynthesis of sterols, hopanoids, and related pentacyclic triterpenes with numerous essential functions, including the stabilization of lipid membranes and formation of membrane rafts (Pan et al. 2015). It can be speculated that this represents important features for the adaptation to the marine environment, like the presence of BGCs for compatible solutes, e.g. ectoine and betaine (Amiri Moghaddam et al. 2016). Further, the carotenoid BGC in marine myxobacteria is similar to the well-known carotenoid gene cluster in *Myxococcus*

xanthus, producing several different carotenoids, mainly phytoene, esterified carotenoids and all-trans-phytoene with different colors such as yellow, orange and red (Iniesta et al. 2007). Such a finding could be expected, since the phenotypic appearance of the strains on solid as well as in liquid medium is yellowish to orange. The presence of several strain specific terpene BGCs contributes to the remarkable complexity and diversity of terpene metabolism in these bacteria.

Our analysis shows that PK BGCs are abundant and conserved in the analyzed genomes. Several resistance and essential core genes detected within the cluster boundaries indicate that the corresponding metabolites will have antibacterial properties. However, only few metabolites with a polyketide backbone, e.g. haliamide and haliangicin, have been isolated so far from marine myxobacteria. In addition, salimabromide might be, partly of polyketide origin. The structures of these metabolites suggest the incorporation of malonate, methylmalonate and ethylmalonate units (Sun et al. 2016b; Sun et al. 2016a; Felder et al. 2013a). Accordingly, the biosynthetic pathways for all these predicted polyketide extender units were identified. Beside the pathways for mCoA biosynthesis, also the genes coding for the biosynthesis of mmCoA and pCoA are conserved in all analyzed strains, whereby emCoA, the rare extender unit putatively used in salimabromide biosynthesis, can be generated from butyryl-CoA via the side activity of the PCC or from crotonyl-CoA via carboxylase activity of CCR. Both coding genes were identified in all analyzed genomes .

Unlike the polyketide BGCs, which are often shared between the strains, NRPS and PKS/NRPS hybrid BGCs are rare and mostly strain specific. *P. pacifica* DSM 14875 carries no NRPS BGC. Examples of PKS/NRPS hybrid BGCs from marine myxobacteria are very limited, *e.g.* haliamide from *H. ochraceum* DSM14365 and phenylnannolone A from the halotolerant myxobacterium *Nannocystis pusilla* B150 were described in detail (Sun et al. 2016b; Bouhired et al. 2014). Here, we identified a PKS/NRPS hybrid gene cluster encoding for leupyrrin in the genome of *E. salina* SWB007. Leupyrrin was isolated and its gene cluster was reported before from the terrestrial *Sorangium cellulosum* strain So ce690 (Kopp et al. 2011). Comparison with this BGC reveals that all encoded proteins are homologues, showing over 50% amino acid identity and complete coverage. Only some rearrangements are observed overall.

Several bacteriocin BGCs were identified in the strains, shared as well as unique ones. It can be suggested that the marine strains use them to compete with other bacteria, since it was reported that *Myxococcus virescens* uses bacteriocins against *M. xanthus* as a competitive mechanism of territory establishment (Smith and Dworkin 1994). Further, it was speculated that specific bacteriocins contribute to the enrichment of species within myxobacterial fruiting bodies (Muñoz-Dorado et al. 2016). Fruiting body formation was also observed in these marine strains.

Additional genomic features, which might contribute to the adaptation to the marine environment, could be the capability for the biosynthesis of arylpolyenes and siderophores. The corresponding BGCs are widely distributed throughout Gram-negative bacteria (Cimermancic et al. 2014). Arylpolyenes are structurally and functionally similar to the well-known carotenoid pigments with respect to their polyene systems and protect bacteria against oxidative stress (Schöner et al. 2016). Siderophores, as iron scavengers, contribute to iron acquisition under low-iron conditions. Here, NRPS-independent siderophore BGCs were only identified in *E. salina* and *P. pacifica* strains, while *H. ochraceum* lacks these BGCs, like the terrestrial myxobacterium *M. xanthus*. It is reported that the presence of arylpolyene BGCs is changing within bacterial genera due to frequent BGC loss from the descendants of a cluster-harboring ancestor, and due to frequent horizontal gene transfer (Cimermancic et al. 2014). In the future, when more marine myxobacterial genomes will become available, it will be possible to judge which events took place. Within the here investigated strains the presence of the ectoine BGC was also specific for *E. salina* SWB005 and SWB007, while the other strains do not harbor this specific BGC. This might be due to different strategies to cope with salt stress. Our work (see 5) revealed *E. salina* SWB007 biosynthesizing ectoine, hydroxyectoine and betaine at high salt concentration, while *P. pacifica* does not produce any specialized organic solutes and relies on amino acids accumulation as osmoprotective agents (Amiri Moghaddam et al. 2016).

The metabolomic analysis revealed a high diversity of chemical features between the investigated bacteria. Despite the differences, one chemical feature is shared in all analyzed strains, *i.e.* polyhydroxybutyric acid (PHB), which was identified by a large

cluster of characteristic MS spectra (± 86.04 Da) belonging to PHBs different in length. The biopolymer PHB plays an important role in long-term survival of bacteria under nutrient-scarce conditions by acting as carbon and energy reserve (Tan et al. 2014). Additionally, polyhydroxyalkanoates enhance the stress tolerance of bacteria against transient environmental assaults such as ultraviolet (UV) irradiation, heat and osmotic shock (Tan et al. 2014). For the fruiting body forming myxobacteria PHB might act as energy supply at nutrient limited conditions, and as protective agent for myxospores.

From the few compounds previously isolated from the marine strains, enhygrolide A was detected from *E. salina* SWB005 and is now also proven to be produced by other *E. salina* strains, i.e. SWB006 and SWB007. Instead, the salimyxins A and B were only detected as strain specific features of *E. salina* SWB005. These compounds are degraded sterols and could hypothetically be modified/degraded from lanosterol or other sterols in this strain (Felder et al. 2013b). However, such modifications of sterols in myxobacteria are still elusive (Wei et al. 2016).

Overall, the percentage of chemical features (6-11%) shared between all analyzed strains is consistent with the small core genome of these bacteria (13-16% of each genome). In contrast, 30-50% of the chemical features are unique in single strains which is consistent with 21-40% of the singleton genes and 43-85% of strain specific BGCs. A similar trend was also revealed in a study of 13 *Pseudoalteromonas luteoviolacea* isolates. Only 2% of the metabolomics features and 7% of biosynthetic genes were shared between all strains, while 30% of all chemical features and 24% of the genes were unique to single strains (Maansson et al. 2016). Similarly, significant differences have been found in the specialized metabolomes of *M. xanthus* isolates from different locations (Krug et al. 2008) and also in the marine actinomycete *Salinispora*, where 75 strains were analyzed and compared (Ziemert et al. 2014). In conclusion, each of the investigated marine myxobacterial strains harbors a high unique genetic and metabolic diversity, rendering this group of microorganisms a promising source for novel specialized metabolites and predicting further diversity for future isolates.

However, the number of isolated compounds to date from these strains is much lower than this predicted potential. This can be mostly contributed to the fact that marine

myxobacteria are hard to isolate and cultivate due to their slow growth and difficult handling. Thus, improved cultivation techniques for these bacteria must be developed in the future (Timmermans et al. 2017) and optimal conditions for specialized metabolite production evaluated. Heterologous expression approaches of orphan gene clusters should be considered as an alternative strategy to tap the specific metabolome of these organisms. Molecular biological tools for such approaches are available and are undergoing a steady process of improvement (Fisch and Schäberle 2016).

The genomic and metabolomic analysis reveals the strain specific potential for specialized metabolite production, and which compounds are indeed accessible under given *in vitro* conditions. These are important data in the early stage of natural product discovery to select and prioritize strains and cultivation conditions.

5 Results II: Different strategies of osmoadaptation in the closely related marine Myxobacteria *Enhygromyxa salina* SWB007 and *Plesiocystis pacifica* DSM 14875

5.1 Background

Micro-organisms living in a highly osmotic environment have to deal with the problem that water follows the osmotic gradient. Cells unable to cope with osmotic stress will become dehydrated. This will eventually disrupt cellular metabolism, and so is used in food conservation by pickling. One strategy to thrive in such environments involves the production of so-called osmolytes to maintain osmotic equilibrium across the cytoplasmic membrane. Osmolytes are organic compounds of low molecular mass that have no influence on cellular metabolism and are non-toxic. Therefore, these highly water-soluble molecules are also called compatible solutes (Brown 1976; Held et al. 2010). Bacteria accumulate many different organic osmolytes in response to hypertonicity, including some amino acids, e.g. proline and glutamate, and some specialized compatible solutes, e.g. betaine and ectoine (da Costa et al. 1998; Burg and Ferraris 2008). Ectoine (1,4,5,6-tetrahydro- 2-methyl-4-pyrimidinecarboxylic acid) was discovered first in *Ectothiorhodospira halochloris* (Galinski et al. 1985) and later on in many bacterial species (Severin et al. 1992; Roberts 2005). The derivative hydroxyectoine was first discovered in *Streptomyces parvalus* (Inbar and Lapidot 1988), before its presence was proven in a wide range of bacteria, and recently also in Archaea (Widderich et al. 2016). These molecules either are synthesized de novo or are taken up from the medium. The biosynthesis is well studied and the underlying genes are known (Sadeghi et al. 2014). In addition to their role in osmoregulation, compatible solutes are known to stabilize cell components under abiotic stress conditions, e.g. temperature, desiccation or oxidative stress. Their properties render them interesting for application in biotechnological research, e.g. for the stabilization of cryocultures/proteins, and as additives for PCR enhancement. Owing to the same effect, they are also applied in formulations of fragile drugs and medical products. Furthermore, compatible solutes display some biological effects. Thus, ectoine was

reported to act as an anti-inflammatory upon particle induced lung inflammation (Sydlik et al. 2009), and inhibited the aggregation of β-amyloid peptides that are involved in senile diseases (Kanapathipillai et al. 2005). Ectoine is already on the market, owing to its moisturizing and protecting properties, in different skin care and cosmetic products, e.g. sun blockers (Kunte et al. 2014).

Despite the relatively moderate osmotic stress of marine habitats, which usually provide a salt concentration of about 3 % NaCl, the chance of finding producers of organic osmolytes is high. Marine *Vibrio* species, for example, are known for their ability to produce ectoine. In recent years the first halophilic Myxobacteria, which are not able to grow in the absence of NaCl, have been isolated. Their terrestrial counterparts instead grow without salt and can usually not grow at NaCl concentrations higher than 1.0 %. However, both have the ability to lyse a variety of bacteria and fungi to obtain nutrients. The few genera of halophilic marine Myxobacteria isolated to date, i.e. *Haliangium*, *Plesiocystis* and *Enhygromyxa* (Fudou et al. 2002a; Iizuka et al. 2003a; Schäberle et al. 2010), are of special interest, owing to their ability to produce unprecedented natural products like the antibiotically active salimabromide (Felder et al. 2013a). Despite this interest in these organisms, their salt tolerance mechanism remains to be investigated.

This study of the closely related marine Myxobacteria *Enhygromyxa salina* SWB007 and *Plesiocystis pacifica* DSM 14875 aimed at investigating the osmo-adaptation mechanisms using analytical experiments and comparative genomics.

5.2 Generation of a suitable medium to study the compatible solutes

Marine Myxobacteria are usually cultivated using a medium containing autoclaved baker´s yeast. Such media are not suitable for the analysis of osmoregulated compatible solute synthesis because compatible solutes originating from yeast cells can be accumulated by the Myxobacteria and falsify the results. Therefore, we established a novel medium, named ASW-Coli, which is based on *E. coli* cells as sole food source.

E. coli can synthesize the compatible solute trehalose *de novo* and can convert choline into glycine betaine. If provided with the medium, *E. coli* is able to transport and amass

a variety of different compatible solutes such as ectoine, proline and glycine betaine with the help of the transporters ProP and ProU (Milner et al. 1988; May et al. 1989). In order to feed Myxobacteria without contaminating compatible solutes, we used the trehalose-free *E. coli* strain BKA-13 that is based on *E. coli* MHK-13 (Haardt et al. 1995). *E. coli* MHK13 is devoid of the compatible solute transporters ProP and ProU and the proline transporter PutP. Like its parental strain MC4100 (Peters et al. 2003; Casadaban 1976), MHK13 is unable to transport and convert choline into glycine betaine (*betTIBA*).

5.3 Growth and salt tolerance of Myxobacteria

The salt tolerance ranges of the two myxobacterial strains *E. salina* SWB007 and *P. pacifica* DSM 14875 were determined in ASW-Coli medium. To enable a fast and reliable method for growth determination, an indirect measurement was performed; since the Myxobacteria rapidly form fruiting bodies instead of disperse growth. Thus, the decrease of OD_{600} was measured. The latter is directly related to lysis and consumption of the prey cells. In that way a decrease in OD_{600} indicates growth of *E. salina* SWB007, and *P. pacifica* DSM 14875 , respectively. *E. salina* SWB007 was able to grow in media containing 0.5% to 3% NaCl. The fastest lysis rate of the prey cells was in the salinity range of 0.5% to 2% NaCl (Figure 5.1). After 7 days all *E. coli* cells were consumed and orange fruiting bodies appeared. Cultures supplemented with 2.5% and 3% NaCl took two days longer to achieve similar results. With 0% NaCl the culture initially seemed to grow but after four days the OD_{600} increased again. In cultures containing more than 3% NaCl no growth was observed.

P. pacifica DSM 14875 showed no growth with 0.5% NaCl containing medium. The lower limit was 1% NaCl and growth was observed up to a concentration of 4% NaCl in ASW-Coli medium. However, the growth was decelerated, and therefore, a concentration of 3.5% NaCl was used for subsequent analyses.

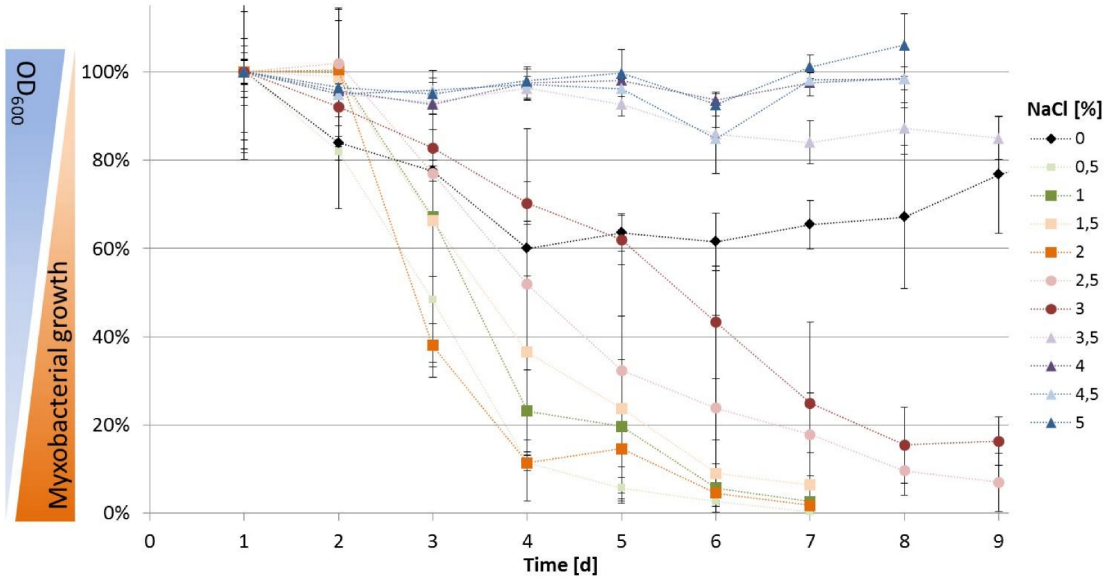


Figure 5.1 Growth of *E. salina* SWB007 in ASW-Coli-Medium with different salt concentrations. The growth was measured by the decrease in OD_{600} due to lysis of *E. coli* cells. For better visualisation of the datasets, only the mean values are given and the data points are connected with a dotted line. The original data as well as the standard deviations can be found in the supporting information (Table 9.2). Each point in this graph represents the mean value of three independent cultures. The OD_{600} at day one was set to 100 %. Fastest lysis of prey cells is obtained between 0.5-2% NaCl (indicated by circles), slower lysis between 2-3% NaCl (circles), and no lysis was observed at 3.5% or higher concentration of NaCl (triangles).

5.4 Identification of the compatible solutes of *E. salina* SWB007 and *P. pacifica* DSM 14875

Both strains, *E. salina* SWB007, and *P. pacifica* DSM 14875 , could grow in a relatively high NaCl concentration range, *i.e.* 0.5% - 3% NaCl for *E. salina* SWB007 and 1% - 4% NaCl for *P. pacifica* DSM 14875. Such a range is usually an indication that the organisms use organic osmolytes as a strategy to deal with the salt stress of the environment. To prove this hypothesis, we identified the compatible solutes produced by these organisms under different salt concentrations. Therefore, two 1 L cultures of each strain were grown in ASW-Coli medium: (i) 0.5% and 3% NaCl for *E. salina*

SWB007, and (ii) 1% and 3.5% NaCl for *P. pacifica* DSM 14875. These cultures were incubated until total clearing of the medium, before the cells were harvested. In total, 92.9 mg dry cell mass (0.5% NaCl) and 134.3 mg (3% NaCl) were obtained from *E. salina* SWB007, and 180 mg (1% NaCl) and 251 mg (3.5% NaCl) from *P. pacifica* DSM 14875 cultures, respectively. Subsequently, 30 mg of the dried cell masses were extracted and the aqueous phase was analysed by HPLC. Clear differences indicating accumulation of organic osmolytes were observed in *E. salina* SWB007 grown in 0.5% and 3% NaCl, respectively (Figure 5.2). The new peaks at high salt concentration were putatively identified as betaine and hydroxyectoine, due to comparison with standard substances. To confirm these results, HPLC-MS experiments followed. These confirmed the presence of betaine and hydroxyectoine, as well as minor amounts of ectoine (Figure 5.3).The most abundant compatible solute was hydroxyectoine. The peak at 20 min retention time which also increased with salinity however, could not be identified. The peak which eluted at 12 min was not dependent on the salt concentration, since it was present in both, the 0.5% NaCl culture and the 3% NaCl culture in the same order of magnitude.

In contrast, the HPLC analysis of *P. pacifica* DSM 14875 cultures grown in medium supplemented with either 1%, or 3.5% NaCl revealed no differences concerning the specialized osmolytes (Figure 9.3 and Figure 9.4) when analyzed under these conditions. Interestingly, this closely related organism did not produce any of the typical organic osmolytes, *e.g.* betaine, ectoine, or hydroxyectoine, even at a high salt concentration.

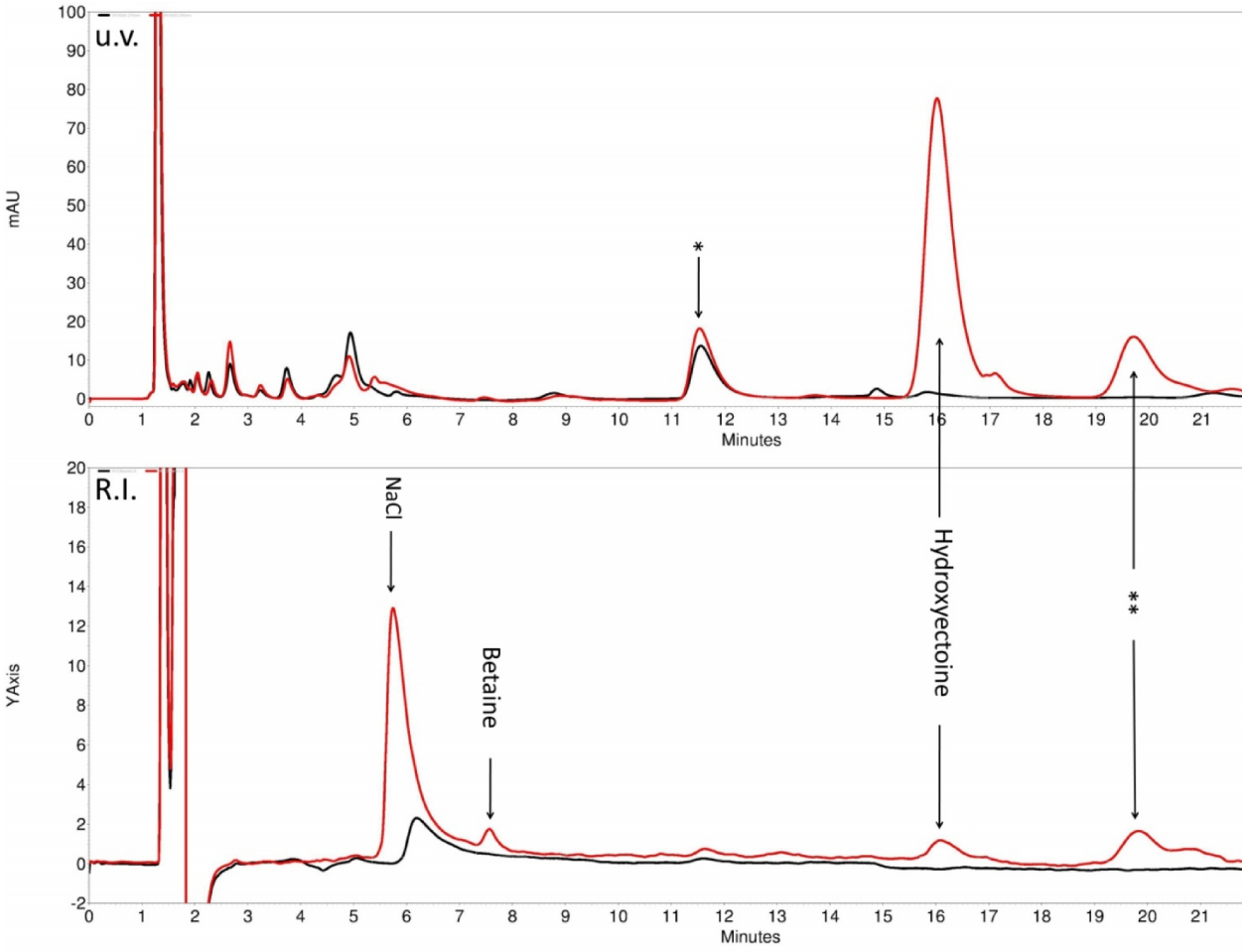


Figure 5.2 Comparison of the compatible solute content of *E. salina* SWB007 in medium containing either 0.5% NaCl, or 3% NaCl. Black line: 0.5% NaCl, red line: 3.0% NaCl. Betaine and hydroxyectoine were present in cells grown in 3% NaCl while in 0.5% NaCl these substances were not detectable. The peak at 20 min (marked by **) could not clearly be assigned. The peak at 12 min (marked by *) was present in both cultures in a similar amount, and therefore this peak is not dependent on the salt concentration. The peak at 6 min, which is only detectable in the refractive index, marks NaCl.

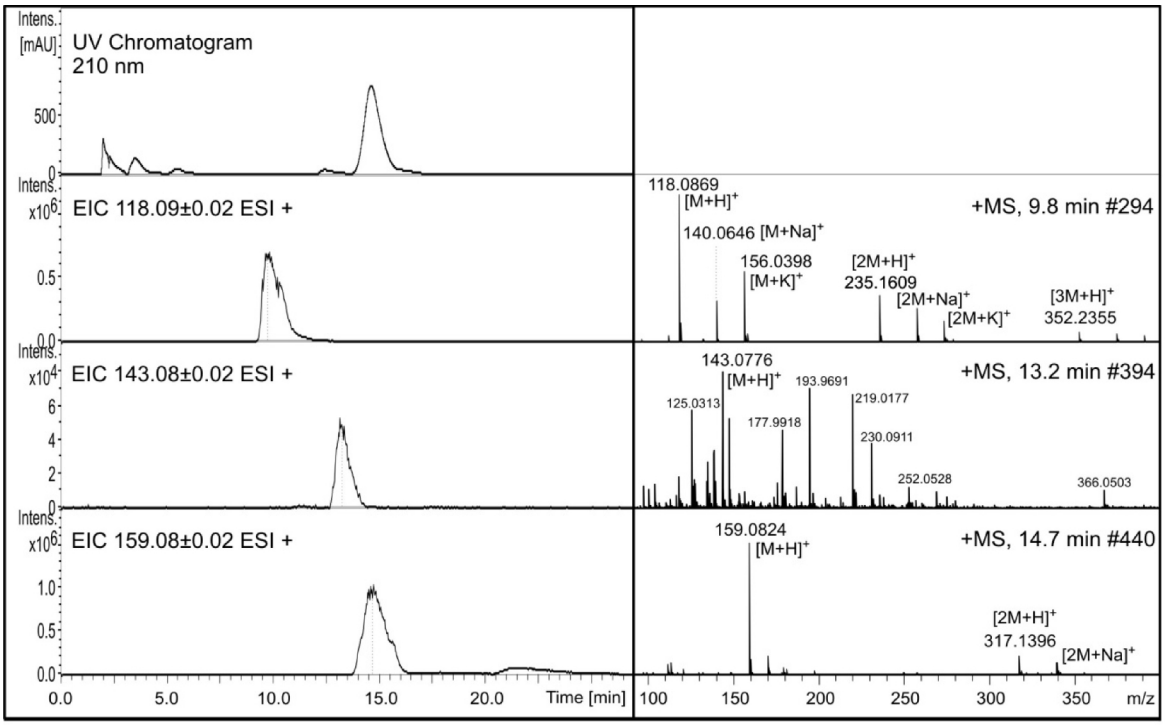


Figure 5.3 HPLC-MS-analysis of compatible solutes produced by *E. salina* SWB007. On the left side the UV chromatogram as well as the extracted ion counts for betaine *m/z* 118.09 $[M+H]^+$, ectoine *m/z* 143.08 $[M+H]^+$, and hydroxyectoine *m/z* 159.08 $[M+H]^+$ are given. On the right side the mass traces of the respective compound are shown. The different adducts are indicated.

5.5 **Free Amino acid analysis**

In addition to specialized compounds, common substances like amino acids may also serve as compatible solutes. Thus, the content of free amino acids in the cells was determined using FMOC ADAM HPLC analysis, revealing significant changes (Table 5.1).

Table 5.1 Content of compatible solutes and free amino acids in *E. salina* SWB007 and *P. pacifica* DSM 14875

Osmolytes/ Amino acid	*E. salina* 0.5% NaCl[§]	*E. salina* 3% NaCl[§]	*P. pacifica* 1% NaCl[§]	*P. pacifica* 3.5% NaCl[§]
Hydroxyectoine [µmol/g]	n.d.	97.56	n.d.	n.d.
Betaine [µmol/g]	n.d.	48.38	n.d.	n.d.
Glutamine [µmol/g]	6.91	49.92	n.d.	8.24
Glutamate [µmol/g]	76.75	272.63	33.24	295.97
Glycine [µmol/g]	n.d.	n.d.	n.d.	124.89
Alanine [µmol/g]	n.d.	40.94	n.d.	3.11
Proline [µmol/g]	n.d.	11.99	n.d.	13.06
Hydroxyectoine [mg/g]	n.d.	15.42	n.d.	n.d.
Betaine [mg/g]	n.d.	5.31	n.d.	n.d.
Glutamine [mg/g]	1.01	7.30	n.d.	1.20
Glutamate [mg/g]	11.29	40.11	4.89	43.55
Glycine [mg/g]	n.d.	n.d.	n.d.	9.38
Alanine [mg/g]	n.d.	3.65	n.d.	0.28
Proline [mg/g]	n.d.	1.38	n.d.	1.50

[§]Chromatograms are given in appendix (Figure 9.3 and Figure 9.8); n.d. = not detectable.

E. salina SWB007 grown under high salt conditions accumulated the amino acids alanine, glutamate, glutamine, and proline (Table 5.1, Figure 9.7, and Figure 9.8). Of these, the level of glutamine increased seven fold and the level of glutamate four folds at elevated salinity. Further, alanine and proline were accumulated, whereas both amino acids were below the detection limit under low salt conditions. Free amino acids analysis of *P. Pacifica* DSM 14875 revealed accumulation of alanine, glutamate, glutamine, glycine and proline at high salinity (Table 5.1, Figure 9.5, and Figure 9.6). The level of glutamate showed a nine fold increase compared to the level at low salinity (296 µmol/g vs 33 µmol/g). Alanine, glutamine, glycine and proline were also significantly accumulated at high salt conditions, while they were below the detection limit at low salt conditions. Glycine was identified as the second major amino acid, contributing to 28% of the *P. pacifica* DSM 14875 organic osmolyte pool. Comparing the

two strains one common feature could be detected, i.e. glutamate was the major player, making up 52% and 66% of the total organic osmolyte pool in *E. salina* SWB007 and *P. pacifica* DSM 14875 , respectively. In *E. salina* SWB007 alanine and glutamine followed with amounts between 40-50 µmol/g, adding up to 19% of the organic osmolyte pool. In *P. pacifica* DSM 14875 instead, these two amino acids were only abundant in minor amounts, adding up to 2.6% of the organic osmolyte pool. In contrast, glycine was identified as the second major amino acid used by *P. pacifica* DSM 14875 , resulting in a share of 28% within the organic osmolyte pool. Proline was accumulated by both organisms only in minor amounts, contributing only less than 3% to the organic osmolytes.

5.6 Genome sequencing, annotation, and bioinformatics analysis for compatible solutes biosynthesis

To elucidate the reason for the observed differences between these closely related myxobacterial strains, *in silico* analysis of the genomes was performed. The available draft genome of *P. pacifica* DSM 14875 was screened in respect to gene clusters associated with betaine, ectoine, and hydroxyl ectoine biosynthesis. However, no such gene cluster was identified, supporting the previous results. Instead, many solute transporters like Na^+/proline(solutes) symporter (GenBank: EDM75864.1, EDM79015.1, EDM74523.1, EDM80875.1, EDM77400.1), proton/sodium-glutamate symporter (GenBank: EDM74633.1), sodium: alanine symporter (GenBank: EDM78123.1), Choline-glycine betaine transporter (GenBank: EDM75025.1) and many ABC transporters were detected in the genome.

In *E. salina* SWB007 biosynthetic gene clusters for the organic osmolytes had been expected, due to the confirmation of these compounds in the extracts. The draft genome of strain SWB007 was screened for the presence of genes corresponding to organic osmolyte synthesis. Hence, a gene locus associated with ectoine biosynthesis was identified (accession number: KU237243). In addition to *ectA*, *ectB*, *ectC* and *ectD* (Figure 5.4 and Table 5.2), further open reading frames putatively coding for enzymes involved in organic osmolyte biosynthesis and for osmolyte transporters are present in

this gene cluster, *i.e.* a sensor histidine kinase (*ectR*) and an aspartokinase (*ectAsk*). These genes represent the complete biosynthetic gene cluster for the synthesis of the osmolyte hydroxyectoine.

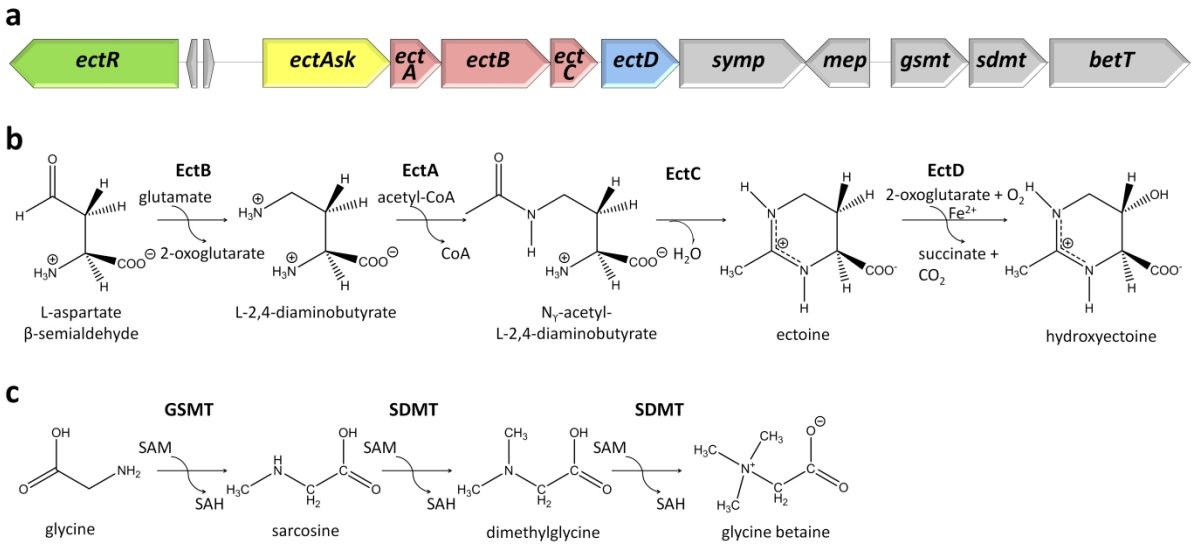


Figure 5.4 Hypothesis for organic osmolyte biosynthesis in *E. salina* SWB007. a) Gene locus coding for genes involved in organic osmolyte biosynthesis. Genes from left to right: *ectR*, sensor histidine kinase; *orf2*, hypothetical protein; *orf3*, hypothetical protein; *ectAsk*, aspartokinase; *ectA*, diaminobutyric acid (DABA) acetyltransferase; *ectB*, DABA aminotransferase; *ectC*, ectoine synthase; *ectD*, ectoine hydroxylase; *symp*, Na/Solute symporter; *mep*, membrane protein (MarC family); *gsmt*, glycine/sarcosine N-methyltransferase; *sdmt*, sarcosine/dimethylglycine N-methyltransferase; *betT*, high affinity choline uptake transporter BetT. It can be seen that the genes for ectoine/hydroxyectoine (*ectAskABCD*) and for betaine (*gsmt and sdmt*) biosynthesis are clustered in *E. salina* SWB007. b) Hydroxyectoine biosynthetic pathway (modified from (Bursy et al. 2008). c) Glycine betaine biosynthetic pathway through methylation steps of glycine (modified from (Khan et al. 2009). SAM: *S*-adenosyl-L-methionine, SAH: *S*-adenosyl-L-homocysteine.

Table 5.2 Genes encoded in the gene locus putatively corresponding to organic osmolytes biosynthesis in *E. salina* SWB007

ORF	Gene	Length(bp)	Highest homology	Identity (%)
1	*ectR*	1899	Histidine kinase (*Marichromatium purpuratum* 984)	43.3
2	*orf2*	126	Uncharacterized protein (*Sorangium cellulosum*)	61.3
3	*orf3*	153	Outer membrane protein and related peptidoglycan-associated (Lipo)protein (*Comamonadaceae bacterium* A1)	52.6
4	*ectAsk*	1458	Aspartate kinase (*Rhodothermus marinus* SG0.5JP)	46.6
5	*ectA*	537	Diaminobutyrate acetyltransferase (*Alkalilimnicola ehrlichii*)	53.7
6	*ectB*	1335	Diaminobutyrate aminotransferase apoenzyme (*Streptomyces* sp. SM8)	63.4
7	*ectC*	381	L-ectoine synthase (*Pontibacillus yanchengensis*)	59.2
8	*ectD*	897	Ectoine hydroxylase (*Castellaniella defragrans* 65)	57.8
9	*solute transporter*	1446	SSS sodium solute transporter superfamily (*Rubinisphaera brasiliensis*)	58.5
10	*mep*	633	UPF0056 membrane protein (*Lyngbya* sp. strain PCC 8106)	53.8
11	*GSMT*	861	SAM-dependent methyltransferase (*Thioploca ingrica*)	71.4
12	*SDMT*	864	Dimethylglycine methyltransferase (*Marichromatium purpuratum* 984)	54.3
13	*BetT*	1530	Choline transporter (*Sphingomonas* sp. Ant20)	56.6

Comparison of this gene cluster with reported ectoine/hydroxyectoine gene loci revealed that *E. salina* SWB007 possesses a particular gene order (Figure 5.5). A putative regulatory gene, i.e. *ectR*, was found in about one quarter of the putative ectoine/hydroxyectoine producers (Widderich et al. 2016). The functional association of an *ectR* gene with ectoine biosynthesis was first demonstrated in the halotolerant alkaliphilic *Methylomicrobium alcaliphilum*. This cluster, including *ectR*, shows the highest overall identity to *E. salina* SWB007 (Figure 5.5). However, the grade of homology varies between the different genes. *ectA*, *ectB* and *ectC* show 35-56%

identity on protein level, while *ectR* has only about 10% identity. This might be an indication that the regulatory mechanism differs from *M. alcaliphilum*, in which it acts as a transcriptional repressor (Overbeek et al. 2014). Of special interest are the two methyltransferases (MTs), *i.e.*, glycine/sarcosine N-MT (GS-MT) and sarcosine/dimethylglycine N-MT (SD-MT), and a *betT* homologue, which follow directly downstream of the ectoine-related genes, oriented in the same direction. These MTs are required for betaine synthesis from glycine, while *betT* codes for the transporter. Thus, the genes corresponding to the biosynthesis of betaine and hydroxyectoine are clustered together.

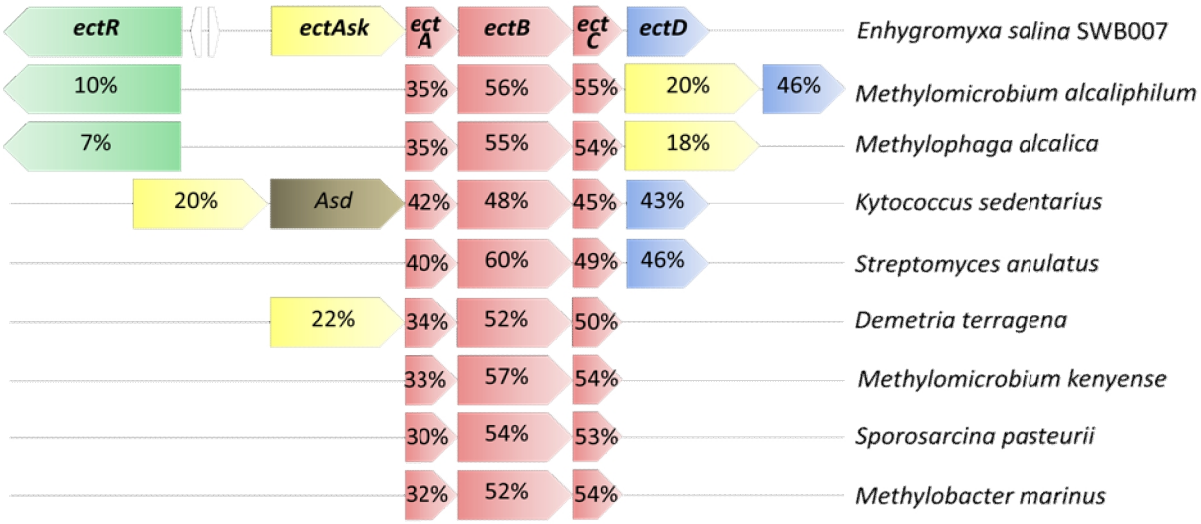


Figure 5.5 Genetic organization of ectoine/hydroxyectoine biosynthesis gene clusters. The *E. salina* SWB007 gene locus is shown in comparison to closely related gene loci, whereby related genes are shown in the same color. The sequence homology on protein level is given in per cent.

5.7 Discussion II

Marine Myxobacteria , *e.g. E. salina* strains, are coming more and more into focus of research due to their biosynthetic potential (Felder et al. 2013a; Felder et al. 2013b). To get these slow growing organisms adapted to growth under laboratory conditions, insights into their ecology, *e.g.* analysis of their salt tolerance mechanism, will be beneficial. To overcome the hurdle that *E. salina* strains were routinely grown in media containing a yeast cell suspension which provides an unfavourable background, the medium ASW-Coli, based on *E. coli* BKA13 instead of baker's yeast as nutrient source, was developed. This enabled to follow the growth of marine Myxobacteria by measuring the optical density (OD) at 600 nm. Hence, the OD decrease due to the consumption of the *E. coli* cells by the predatory bacteria was measured. The values for the salt tolerance of these strains (0.5-3% NaCl for *E. salina* SWB007; 1-4% NaCl for *P. pacifica* DSM 14875) determined with this method are in accordance with previous results (Iizuka et al. 2003a; Schäberle et al. 2010). Such a relatively high degree of flexibility to environmental salt concentrations indicates an adaptation strategy involving organic osmolyte accumulation, rather than the salt-in-strategy (Sleator and Hill 2002).

At high salt concentrations a total of 521 µmol/g solutes was detected in *E. salina* SWB007. The bigger part of this consisted of glutamate (52%), followed by hydroxyectoine and betaine, which accumulated to 19% and 10%, respectively. Hence, glutamate, hydroxyectoine, and betaine seem to be the main players in osmoadaptation of *E. salina*. A comparison with marine *Vibrio fischeri* species revealed that the observed total solute pool falls within expectations (451 µmol/g for *V. fischeri* DSM 7151). Here also, the major compound was glutamate (approx. 50%), followed by ectoine (approx. 24%) (Schmitz and Galinski 1996). Accumulation of glutamate as well as of hydroxyectoine is fully congruous with metabolism and other cellular functions (Reuter et al. 2010). However, it seems like this mechanism cannot be regarded as a general rule for Myxobacteria . For the terrestrial myxobacterium *Myxococcus xanthus*, the model organism for Myxobacteria, it was shown that it produces the compatible solutes betaine and trehalose under salt stress conditions (Kimura et al. 2010; McBride and Zusman 1989). Our results indicate that in addition to betaine the marine strain

acquired, or retained, the ability to use hydroxyectoine as main compatible solute. Hydroxyectoine, a derivative of ectoine, is widespread among halophilic bacteria and compared to ectoine it provides better protection against various stress conditions, especially desiccation and heat (Lippert and Galinski 1992; Louis et al. 1994; Tanne et al. 2014). In the intertidal zones of the littoral environment, desiccation is at least a temporary stress factor. This might be an explanation why the organism favours hydroxyectoine as compatible solute. The preferential accumulation of hydroxyectoine has also been observed in other salt-stressed bacteria, such as *S. coelicolor*A3(2) and *Virgibacillus salexigens* (Sadeghi et al. 2014). Recently the ability to synthesize hydroxyectoine was also demonstrated in the acidophilic *Acidiphilium cryptum* (Moritz et al. 2015) which displays only limited salt tolerance. The authors concluded that hydroxyectoine, besides osmoadaptation, may serve still unknown other functions.

In *P. pacifica* DSM 14875 the total amount of detected solutes was 445 µmol/g. The major solute at high salt concentrations was also glutamate, contributing to two-thirds (66%) of the composition, followed by the amino acid glycine, covering 28%. Thus, *P. pacifica* DSM 14875 , the closest relative to the *Enhygromyxa* clade, did not produce any of the well-established compatible solutes under high salt conditions. This marine myxobacterium instead accumulates the amino acids glutamate, glycine and proline. Glycine and proline can be enzymatically synthesized or might be taken up from the environment. The accumulation of glutamate as primary response is the same as in *E. salina* SWB007. However, glycine has so far not been reported as a major bacterial osmolyte. The few reports available refer to its osmotic use in marine mussels and ciliates, in combination with alanine and taurine in the former and alanine and proline in the latter (Ellis et al. 1985; Kaneshiro et al. 1969). It is worthy of note that glycine has been applied empirically for protein stabilization during freeze-thawing of phosphate buffer systems and as a bulking agent during lyophilization of monoclonal antibodies (Pikal-Cleland et al. 2002; Meyer et al. 2009). Therefore, the fact that glycine is one of the dominant amino acids in *P. pacifica* deserves further investigations.

In accordance with the above observations a gene locus associated with solute biosynthesis could be identified in *E. salina* SWB007 (Figure 5.4). The biosynthesis of

ectoine and hydroxyectoine has been firmly established for many microorganisms (Sadeghi et al. 2014). Accordingly, a biosynthetic hypothesis was deduced for hydroxyectoine and betaine in *E.salina* (Figure 5.4b and Figure 5.4c).

It is striking that only hydroxyectoine is accumulated and not the direct precursor ectoine. Thus, the EctD-catalyzed step has to be very efficient. In *Streptomyces rimosus* it was shown that the effect of salinity on *ectD* transcription level is 3-fold more prominent than on ectC (Sadeghi et al. 2014). Further, it was suggested that the localization of *ectD* at the 3'-end of the *ectABCD* mRNA contributes positively to its stability. A putative regulator, encoded by *ectR*, is clustered with the *ectABCD* genes. EctR is expected to be responsible for the regulation of hydroxyectoine synthesis, dependent on natural stimuli. It was shown before, that in Gram-negative and Gram-positive bacteria the genes are induced by osmotic or temperature stress (Widderich et al. 2016). The presence of genes for several different osmolyte transporters in the genome can be regarded as a complementary mechanism for osmoadaptation in *E. salina* SWB007. A Na^+/proline symporter is encoded directly downstream of *ectD*, and such transporters can mediate the accumulation of organic osmolytes, *e.g.* betaine, ectoine, proline at high osmotic pressure (Wood 2006). Further, two *N*-methyltransferase genes are clustered with the hydroxyectoine gene locus of *E. salina* SWB007 (Figure 5.4a). It can be assumed that these genes are corresponding to betaine biosynthesis (Waditee et al. 2003). Downstream of these two methyltransferase genes a gene putatively coding for the high affinity betaine/carnitine/choline transporter BetT can be found.

In contrast, none of the established compatible solutes were detected in *P. pacifica* DSM 14875 . This is in accordance with the genome analysis, providing no gene cluster corresponding to their biosynthesis. However, the presence of many osmolyte transporters and symporters indicates that this bacterium counts on uptake to cope with salt stress. Such an uptake of osmolytes, *e.g.* amino acids, as major factor against the osmotic gradient represents an intelligent strategy for a predatory bacterium, which is lysing prey cells and thereby creates a source for these molecules itself. It is therefore

at present unclear whether the markedly increased glycine level (28% of total solutes) is a result of uptake or genuine osmoregulated biosynthesis.

6 Results III: Genome mining of *Labrenzia* spp. for bioactive metabolites

6.1 Draft Genome Sequence of the Marine Bacterium *Labrenzia* sp. 011, a Potential Protective Agent of Mollusks

6.1.1 Background

Bacteria of the genus *Labrenzia* colonize surfaces, such as oyster shells and may produce antibacterial compounds inhibiting the growth of other bacteria. *Labrenzia* sp. 011 showed activity against the oyster pathogen *Roseovarius crassostreae*. Therefore, its genome will enable the identification of the biosynthetic gene clusters corresponding to the protective compounds and provide potential biotechnological applications.

6.1.2 Genome assembly and phylogeny of *Labrenzia* sp. 011

Labrenzia sp. 011 was isolated from marine sediment of the coastal area of Kronsgaard, Germany. The phenotypic appearance of its colonies is creamy yellow on Difco™ marine agar 2216 (Table 6.1). The draft genome sequence of *Labrenzia* sp. 011 consists of 65 contigs (> 1000 bp) with 5,102,962 bp in length with a G+C content of 61.6%. 4,812 coding sequences (CDSs) were predicted (this number includes proteins annotated as hypothetical), whereof 2280 CDSs (48%) were categorized in 473 different subsystems with identified functional roles. The whole-genome shotgun project has been deposited at DDBJ/ENA/GenBank under the accession no. QCYM00000000. The strain was deposited at the German collection of microorganisms and cell cultures DSMZ with the accession number DSM 107099.

Table 6.1 General features of *Labrenzia* sp. 011, and the minimum information about a genome sequence (MIGS mandatory information).

Items	Description
Investigation_type	Bacteria
Strain	*Labrenzia* sp. 011
Gram stain	Negative
Cell shape	Rod
Pigmentation	Creamy yellow
Temperature optimum	30 °C
Lat_lon	54.731111 N 9.964167 E
Geo_loc_name	Kronsgaard, Germany
Collection_date	15-Aug-2012
Env_biome	Marine biome (ENVO:00000447)
Env_feature	Sea coast (ENVO:00000303)
Env_material	Marine sediment (ENVO_03000033)
Env_package	Surface sediment
Rel_to_oxygen	Aerobe
Num_replicons	1
Seq_meth	Illumina

A phylogenetic tree of all *Labrenzia* strains with available genomes based on the nucleotide sequence alignment of the core genomes revealed *Labrenzia* sp. OB1 and *L. marina* DSM 17023 that were isolated respectively from coastal seawater in La Jolla, CA, USA and South Korea, as the most closely related strains to *Labrenzia* sp. 011 (Figure 6.1).

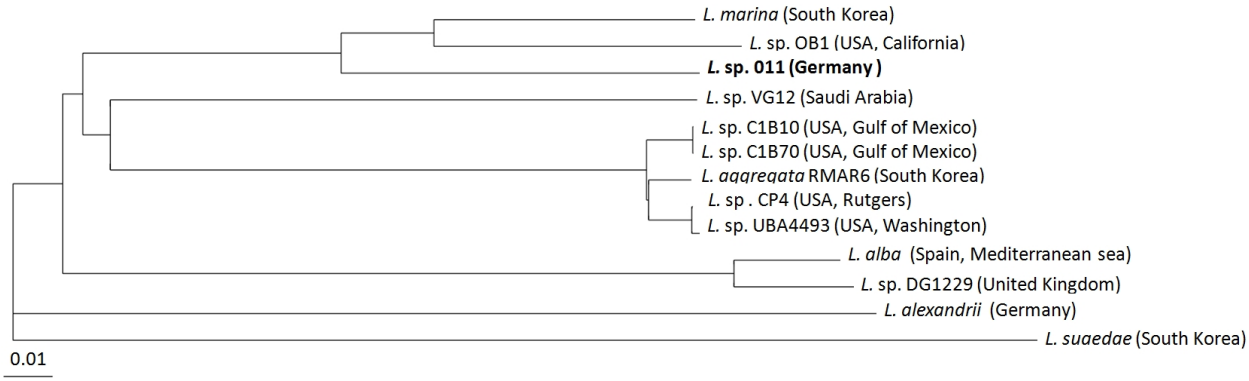


Figure 6.1 Phylogenetic tree of selected *Labrenzia* strains with available genomes. Tree was build out of a core of 2131 genes per genome. Geographic origins of the strains are given in parentheses.

6.1.3 Comparative genomics of genus *Labrenzia*

In order to obtain further insights into the degree of similarity between the analyzed genomes, the number of core genes and of singletons was determined. 2131 CDS contributed to the core genome of the *Labrenzia* strains, equivalent to ~44% of the *Labrenzia* sp. 011 genome (Figure 6.2A). It was calculated that the core genome will be around 2113 CDS, based on a decay function (2929.005*exp(-x/3.229) + 2112.783, see Figure 6.2B). The pan genome increases by every additional *Labrenzia* strain, indicating an open pan genome of *Labrenzia* (Heaps' law function: 5736.13*x^0.462, see Figure 6.2C).

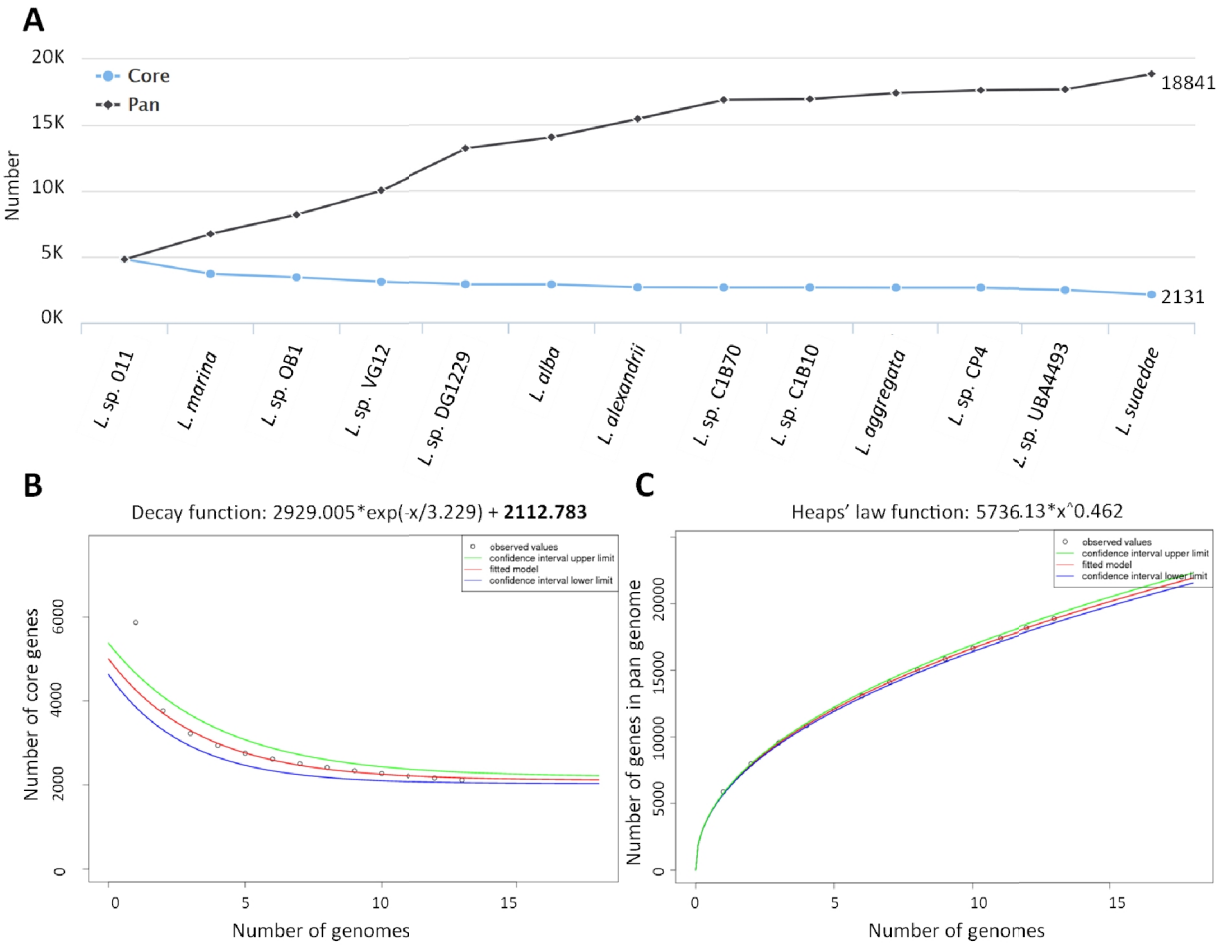


Figure 6.2 Core vs. pan genome plot of the genomes. B: Core genome development plot. C: Pan genome development plot

The Average nucleotide identity (ANI) values between *Labrenzia* sp. 011 and all analyzed *Labrenzia* strains were between 73.55% to 84.85% in pair-wise comparisons of the sequences (Figure 6.3). This puts the strain only into distant relation to other strains, since values smaller than 80-85% ANI must be regarded as distantly related (Goris et al. 2007). The *In silico* DNA-DNA hybridization (isDDH) values between *Labrenzia* sp. 011 and other *Labrenzia* strains were between 22.7% to 33.1%, whereby the highest values were obtained for *Labrenzia* sp. OB1 and *L. marina* DSM 17023, verifying the phylogenetic relationship between these strains and 011. Further, differences in G+C content between *Labrenzia* sp. 011 and other *Labrenzia* strains were between 1.32-5.38%, which supports the species delineation (Table 6.2). Therefore, the *in silico* parameters (ANI ≥96%, isDDH ≥70% and difference in G+C content of ≤1%)

(Meier-Kolthoff et al. 2014; Colston et al. 2014; Goris et al. 2007) define *Labrenzia* sp. 011 as a distinct new species of the genus *Labrenzia* (Figure 6.3 and Table 6.2). Instead, CP4, UBA4493, C1B70, C1B10 seem to be closely related strains to *L. aggregata* RMAR6 with ANI values between 97-100% (Figure 6.3).

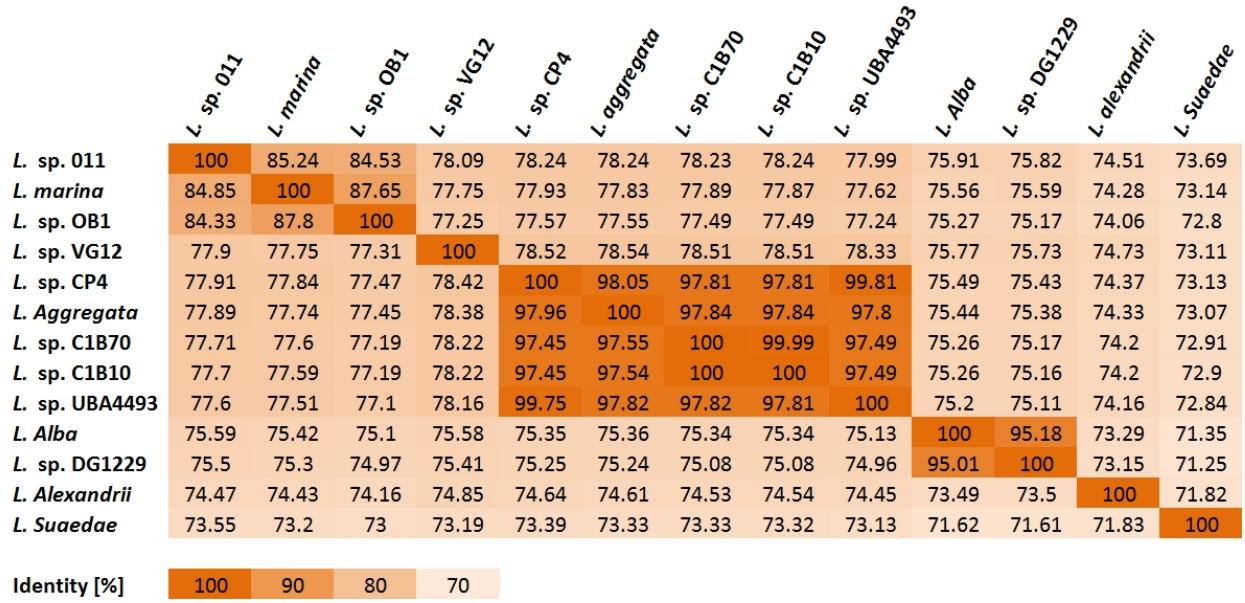


Figure 6.3 Average nucleotide identity (ANI) heat map of the selected *Labrenzia* strains

Table 6.2 *In silico* DNA-DNA hybridization (isDDH) and G+C difference of *Labrenzia* sp. 011 vs. other *Labrenzia* strains

***Labrenzia* sp. 011 vs.**	**isDDH%**	**G+C difference %**
Labrenzia sp. OB1	33.1	2.20
Labrenzia marina	30.2	1.41
Labrenzia sp. C1B70	26.7	2.71
Labrenzia sp. C1B10	26.7	2.71
Labrenzia sp. CP4	26.6	2.52
Labrenzia sp. VG12	26.5	1.62
Labrenzia aggregata	26.5	2.57
Labrenzia sp.	26.3	2.56
Labrenzia sp. DG1229	24.8	5.38
Labrenzia alba	24.4	5.25
Labrenzia alexandrii	23.5	5.26
Labrenzia suaedae	22.7	1.32

6.1.4 Biosynthetic potential of the newly sequenced *Labrenzia* sp. 011 strain

The genome of *Labrenzia* sp. 011 carries genes related to nitrogen metabolism and denitrification (56 CDSs), polyhydroxybutyrate metabolism (32 CDSs), and many genes that are related to stress response, e.g. heat and cold shock (169 CDSs) (Figure 6.4). *Labrenzia* sp. 011 belongs to the family of *Rhodobacteraceae*, which is a sister family of the *Rhizobiales.* The latter fix nitrogen in plant roots (Williams et al. 2007). This data may explain the denitrification potential of the oyster microbiome, which is dominated by *Rhodobacteraceae* (Arfken et al. 2017). Bacteria of this family rapidly colonize surfaces, such as oysters shells and may produce antibacterial components, which in turn shape the microbiome by inhibiting the growth of other bacteria (Arfken et al. 2017; Dang et al. 2008).

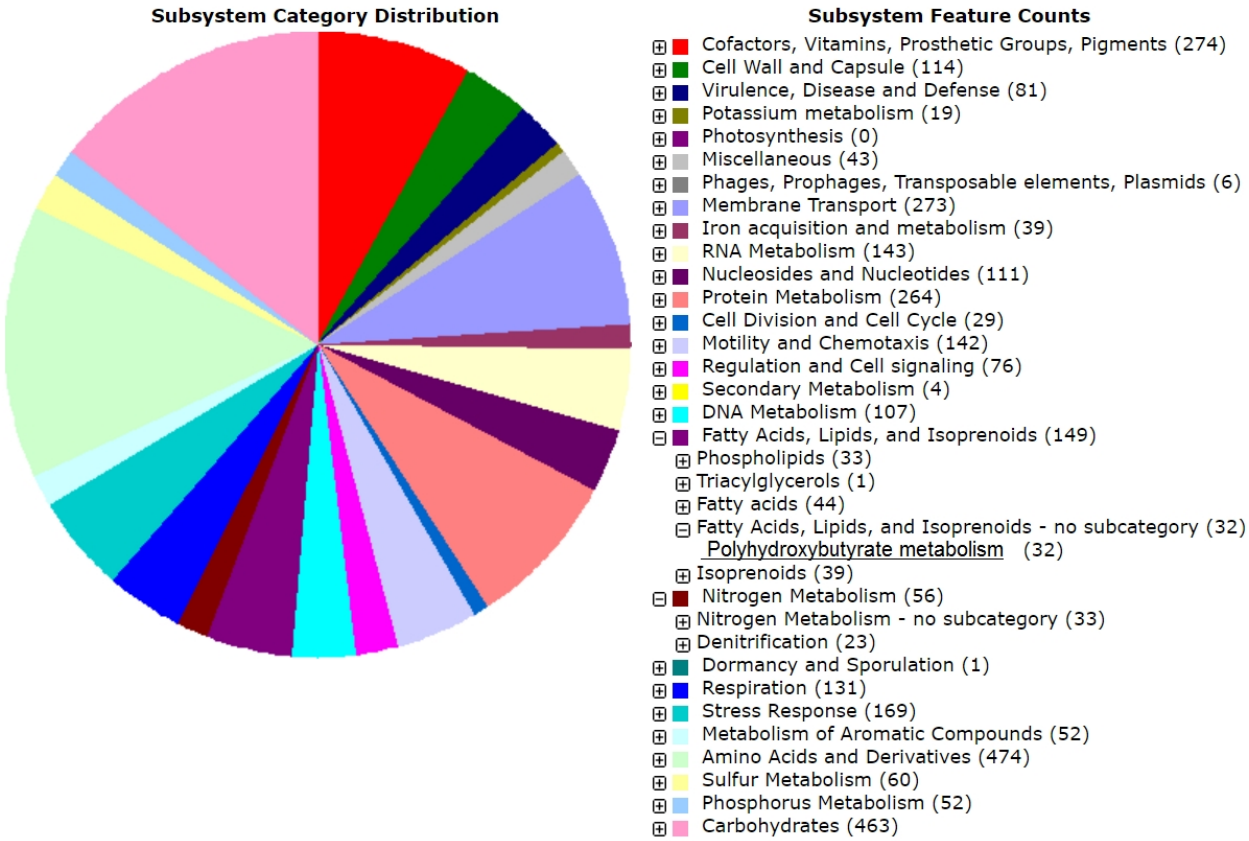


Figure 6.4 Subsystem category distribution and feature counts in the genome of *Labrenzia* sp. 011

In total, 11 biosynthetic gene clusters (BGCs) were identified (5.3% of the genome) including one type-I PKS, one terpene, one bacteriocin, four fatty acids, and four saccharide BGCs (Figure 6.5). Additionally, 23 putative gene clusters were identified by using the cluster finder algorithm (3.8% of the genome), thereunder three BGCs for cyclopropane fatty acid synthases (Figure 6.5).

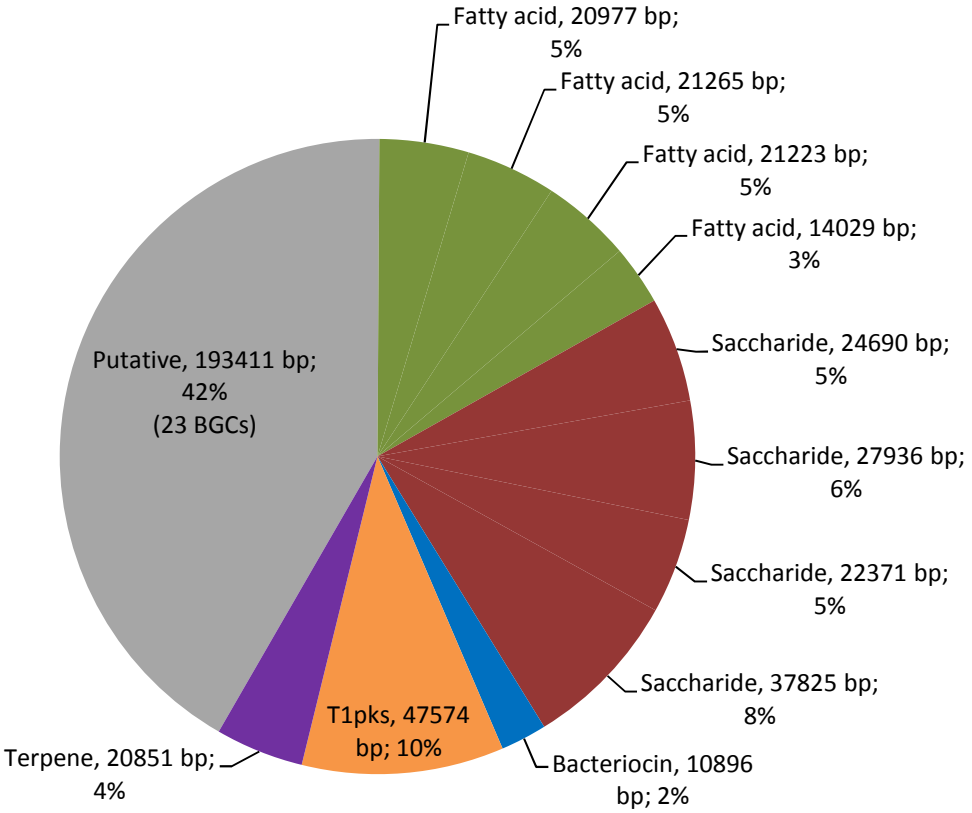


Figure 6.5 Distribution of the BGCs in the genome of *Labrenzia* sp. 011. In total 46,3048 bp (equal to 9.1% of the genome) were identified. Identified regions and percentages to the total are given.

6.2 Biosythesis of cyclopropane-containing fatty acids produced by *Labrenzia* sp. 011 inhibit a causative agent of the roseovarius oyster disease

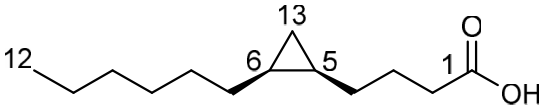
6.2.1 Background

Interactions in marine ecosystems are often mediated by chemicals (Boyd et al. 1999; Ganapiriya et al. 2012; Clare 1996). Indeed, it has been established that epibiotic or symbiotic bacteria and metabolites produced by them, play a major role in keeping their host healthy. A specific example of such a relation is the protective mechanism displayed by *Alteromonas* sp. I-2, which colonizes the embryos of the Caribbean shrimp *Palaemon macrodactylus* (Gil-Turnes et al. 1989). This bacterium produces the antifungal agent 2,3-indolinedione, which prevents infection with the fungus *Lagenidium callinectes*, a frequent crustacean pathogen (Gil-Turnes et al. 1989). In the same manner, embryos of the American lobster *Homarus americanus* were reported to be chemically protected against the same pathogen by the antifungal agent 4-hydroxyphenethyl alcohol, produced by the Gram-negative bacterium SG-76 (Gil-Turnes and Fenical 1992).

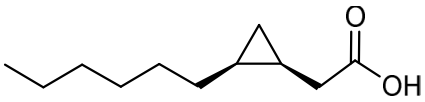
Within this context, species of the genus *Labrenzia*, previously classified *Stappia* (Biebl et al. 2007), have been suggested to be protective agents of mollusks, e.g. *Crassostrea virginica* against the bacterial pathogen *Pseudoroseovarius crassostreae* (also *Aliiroseovarius crassostreae*), the causative agent of roseovarius oyster disease (Boettcher et al. 2000; Maloy et al. 2007; Pujalte et al. 2005). This pathogen has a negative impact on marine ecology and on aquaculture activities related to oysters (Méndez-Vilas 2010). A mechanism mediated by chemical substances is thought to be involved in the putative protective role of *Labrenzia* spp.

However, hardly any metabolites from *Labrenzia* spp. are known. To date, there is only one report of a pederine-analog isolated from one member of the genus *Labrenzia*, i.e. *Labrenzia* sp. PHM005, which showed cytotoxic activity against different cancer cell lines *in vitro* (Schleissner et al. 2017). The latter report reflects the potential of these organisms to produce bioactive substances. With this background, we decided to find substances to support the putative protective role of bacteria of the genus *Labrenzia* in the marine environment. Our findings suggest that the halophilic bacterium *Labrenzia*

sp. strain 011, isolated from marine sediments is able to produce antimicrobial metabolites (**1**, **2**) (Figure 6.6) that could provide protection to their hosts against fungal and bacterial pathogens, especially *Pseudoroseovarius crassostreae*. Based on genome analyses, all to date sequenced *Labrenzia* spp. have the ability for the biosynthesis of cyclopropyl-containing fatty acids, which could all be active against ROD.



cis - 4 - (2 - hexylcyclopropyl) - butanoic acid **(1)**



cis - 2 - (2 - hexylcyclopropyl) - acetic acid **(2)**

Figure 6.6 Structures of the cyclopropane fatty acids **1** and **2** (relative configuration) produced by *Labrenzia* sp. strain 011.

6.2.2 Bioinformatic analysis of the genome of *Labrenzia* sp. 011 for the identification of the putative biosynthetic genes of 1 and 2

In order to elucidate the putative biosynthesis of compounds **1** and **2**, the genome of *Labrenzia* sp. 011 has recently been sequenced (accession number: QCYM00000000) (Amiri Moghaddam et al. in press). The resulting genome was compared with further genome sequences available of the genus *Labrenzia*. It was found that all *Labrenzia* strains share between 71.1-99.9% identities at the amino acid level (Figure 6.3), and orthologous genes in these strains might perform the same functional roles. In a next step, the genomes of *Labrenzia* strains were screened *in silico* to identify the presence of genes or BGCs coding for cyclopropane fatty acid synthases (CFAS), which are reported to catalyze the cyclopropanation of unsaturated lipids in bacteria (Courtois et al. 2004; Guangqi et al. 2010).

Specifically, regions with conserved CFAS genes in *Labrenzia* strains were aligned. All investigated *Labrenzia* strains harbor between 2 - 5 different types of CFAS genes in their genomes, with some of them being highly conserved (Figure 6.7). *Labrenzia* sp. 011 possesses four different types of CFAS genes, assigned as CFAS1 - 3 and CAFS7. The first three, i.e. CFAS1 - 3 are conserved among most *Labrenzia* strains analyzed, while CFAS7 was only in *Labrenzia* sp. 011 and *L. marina* strains.

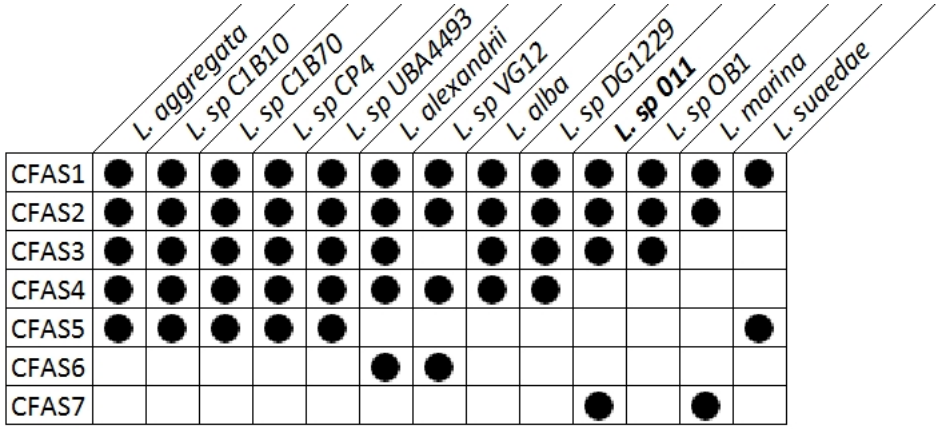


Figure 6.7 Presence of different types of CFAS genes in *Labrenzia* strains

In a next step, phylogenetic analyses of the CFAS genes have been performed (Figure 6.8). These included all CFAS genes detected in *Labrenzia* strains, and all described CFAS genes from UniProtKB protein database, as well as CFAS genes from the marine cyanobacterium *Moorea producens* (formerly classified as *lyngbya majuscula*) (Engene et al. 2012), which was earlier reported to produce similar medium chain length fatty acids with a cyclopropane ring as in **1** and **2** (Macmillan and Molinski 2005).

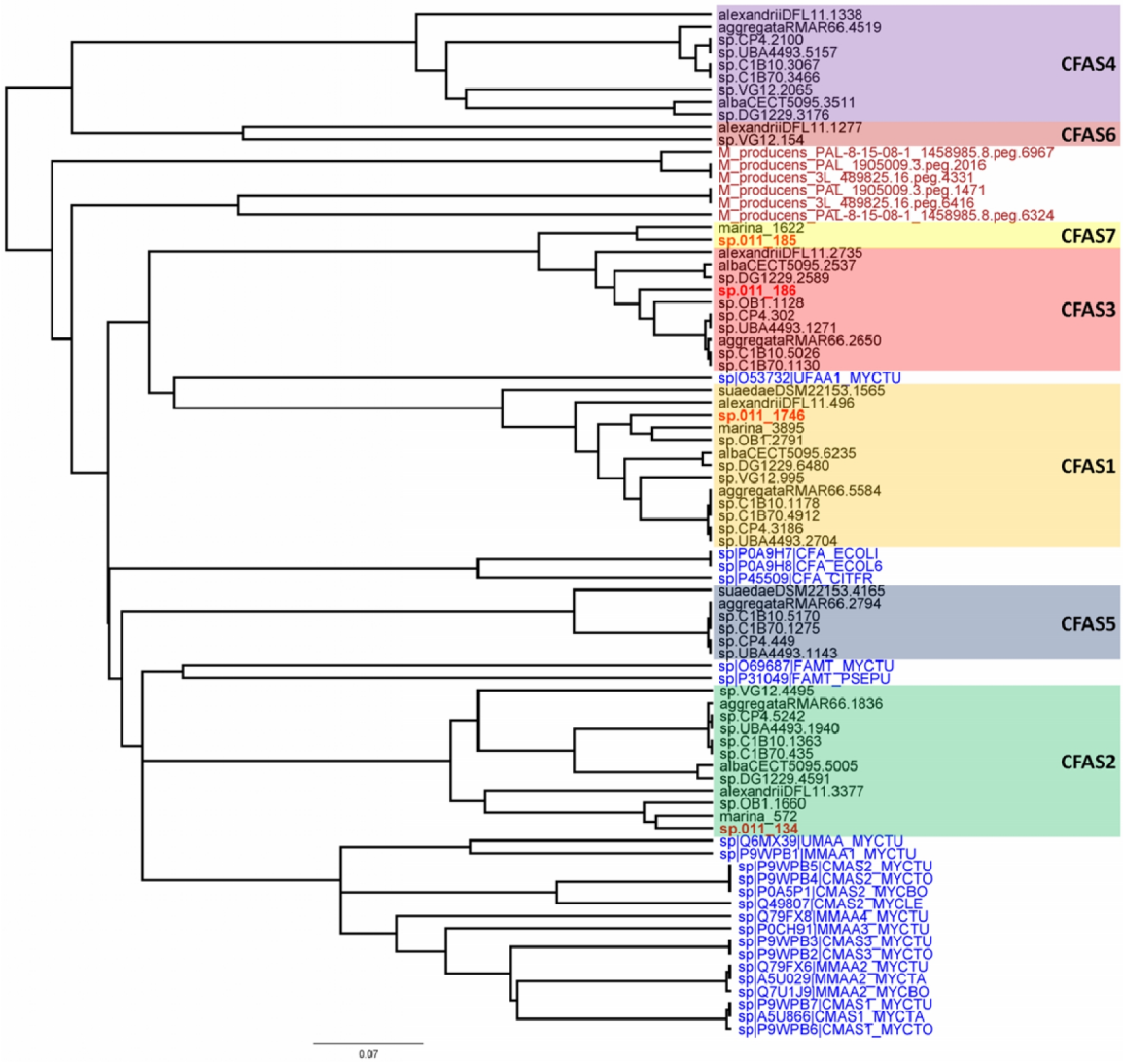


Figure 6.8 Phylogeny tree of different clades of CFAS genes detected in *Labrenzia* strains. Text in red indicates the CFAS genes from *Labrenzia* sp. 011 and black indicates the annotated CFAS genes from other *Labrenzia* strains. Text in blue indicates the CFAS gene obtained from UniProtKB protein database with respective codes and dark red indicates CFAS genes from *Moorea producens* strains.

CFAS1 is a conserved gene in all the investigated *Labrenzia* strains (Figure 6.9A and Figure 6.10) (75-91% protein identity) with sequences for a short-chain dehydrogenase/reductase (SDR) and a FAD-dependent oxidoreductase in its vicinity (Figure 6.9A). The closest described CFAS to the CFAS1 clade of *Labrenzia* spp. was *ufaA1* from *Mycobacterium tuberculosis* (strain ATCC 25618) (Figure 6.8). The enzyme encoded by the latter catalyzes the transfer of a methyl group from S-adenosyl-L-

methionine (SAM) to the double bond of oleic acid to produce tuberculostearic acid (syn. 10-methylstearic-acid).

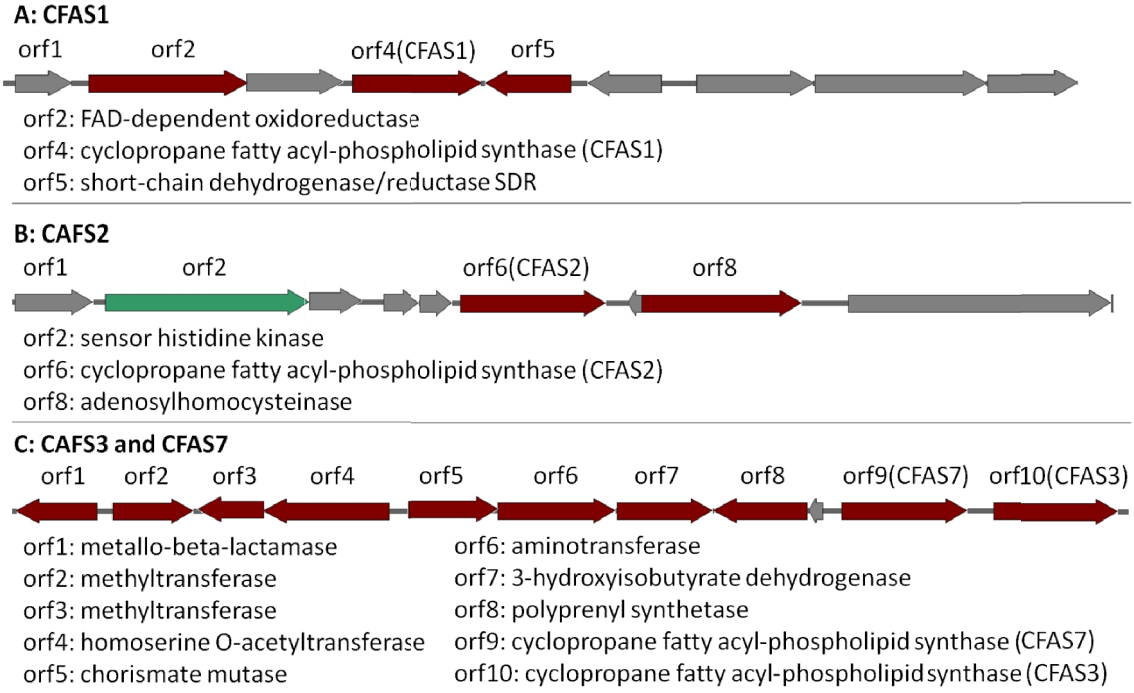


Figure 6.9 CAFS1, CFAS2 and CFAS3/7 gene clusters in *Labrenzia* sp. 011 assigned as putative BGCs by AntiSMASH. Dark red indicates the biosynthetic genes. Green indicates regulatory genes and grey indicates other genes. Locus accession numbers: QCYM01000001 (CFAS1), QCYM01000010 (CFAS2), and QCYM01000011 (CFAS3/7).

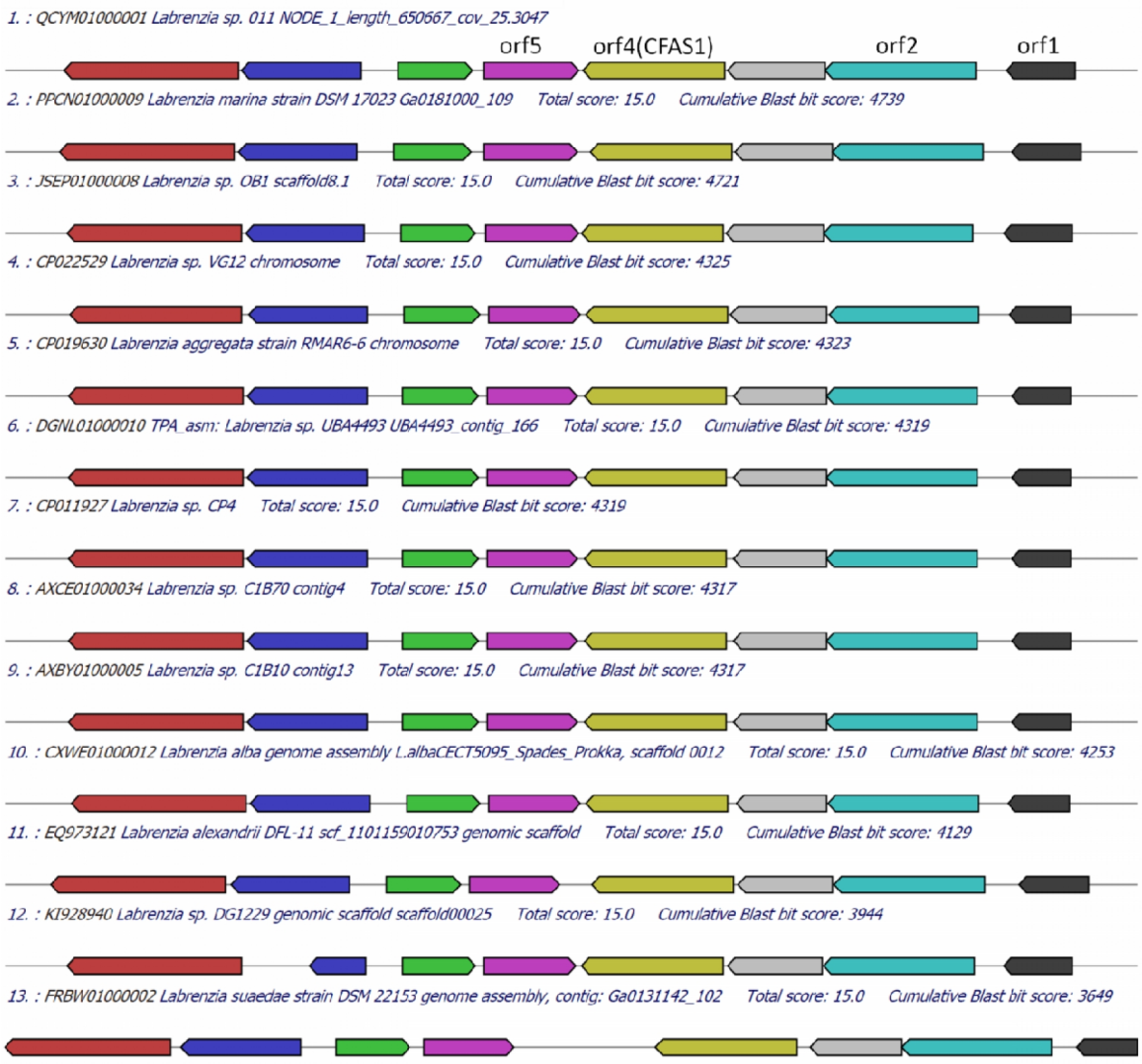


Figure 6.10 CAFS1 gene cluster alignment in *Labrenzia* strains, orf2: FAD-dependent oxidoreductase, orf4: Cyclopropane fatty acyl-phospholipid synthase (CFAS1), and orf5: short-chain dehydrogenase/reductase SDR. Same colors represent the same annotation.

CFAS2 is also conserved among all *Labrenzia* strains (Figure 6.9B and Figure 6.11), except *L. suaedae* (72-90% protein identity) with sequences encoding an adenosylhomocysteinase in its vicinity, which converts S-adenosylhomocysteine to homocysteine and adenosine in methionine biosynthesis. The closest described CFAS to the CFAS2 clade in *Labrenzia* spp. was the methyltransferases of *M. tuberculosis*

(Figure 6.8) which catalyzes the cyclopropanation of the double bonds in the mycolates by the addition of a methyl group derived from SAM (Barkan et al. 2010).

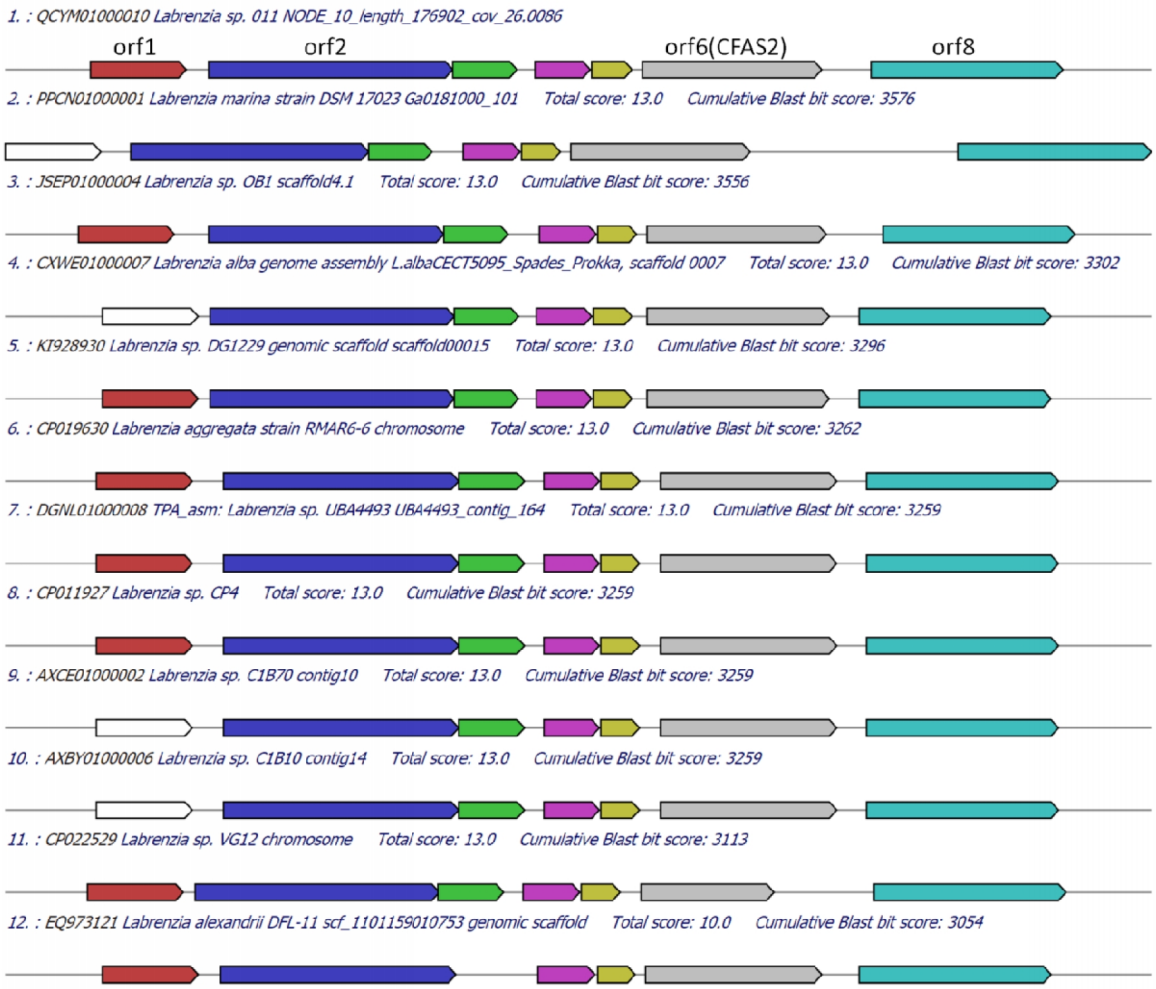


Figure 6.11 CAFS2 gene cluster alignment in *Labrenzia* strains, orf2: Sensor histidine kinase, orf4: Cyclopropane fatty acyl-phospholipid synthase (CFAS2), and orf8: Adenosylhomocysteinase. Same colors represent the same annotation.

Also, the CFAS3 gene is conserved in most *Labrenzia* strains (84-92% protein identity), and clustered with a unique CFAS7, which is detected only in *Labrenzia* sp. 011 and *L. marina* (Figure 6.8C and Figure 6.12). This BGC also includes sequences for a conserved homoserine O-acetyltransferase and a methyltransferase (*MetW*). These might participate in the metabolism of methionine and subsequently provide SAM by

using a well-conserved S-adenosylmethionine synthetase present in all the *Labrenzia* strains.

CFAS7 in *Labrenzia* sp. 011 is a distinct CFAS, albeit related to the CFAS3 clade (Figure 6.8). To date, no other CFASs related to the CAFS3 clade have been reported. *M. producens* produces similar cyclopropane-containing fatty acids i.e. lyngbyoic acid with 12 carbon chain as **1** and majusculoic acid with 14 carbon chain (Kwan et al. 2011; Macmillan and Molinski 2005). Indeed genome of *M. producens* strains harbor two distinct clades of CFAS genes (Figure 6.8). Therefore, the CFAS3 and CFAS7 gene clusters might be responsible for biosynthesis of **1** and **2** in *Labrenzia* sp. 011.

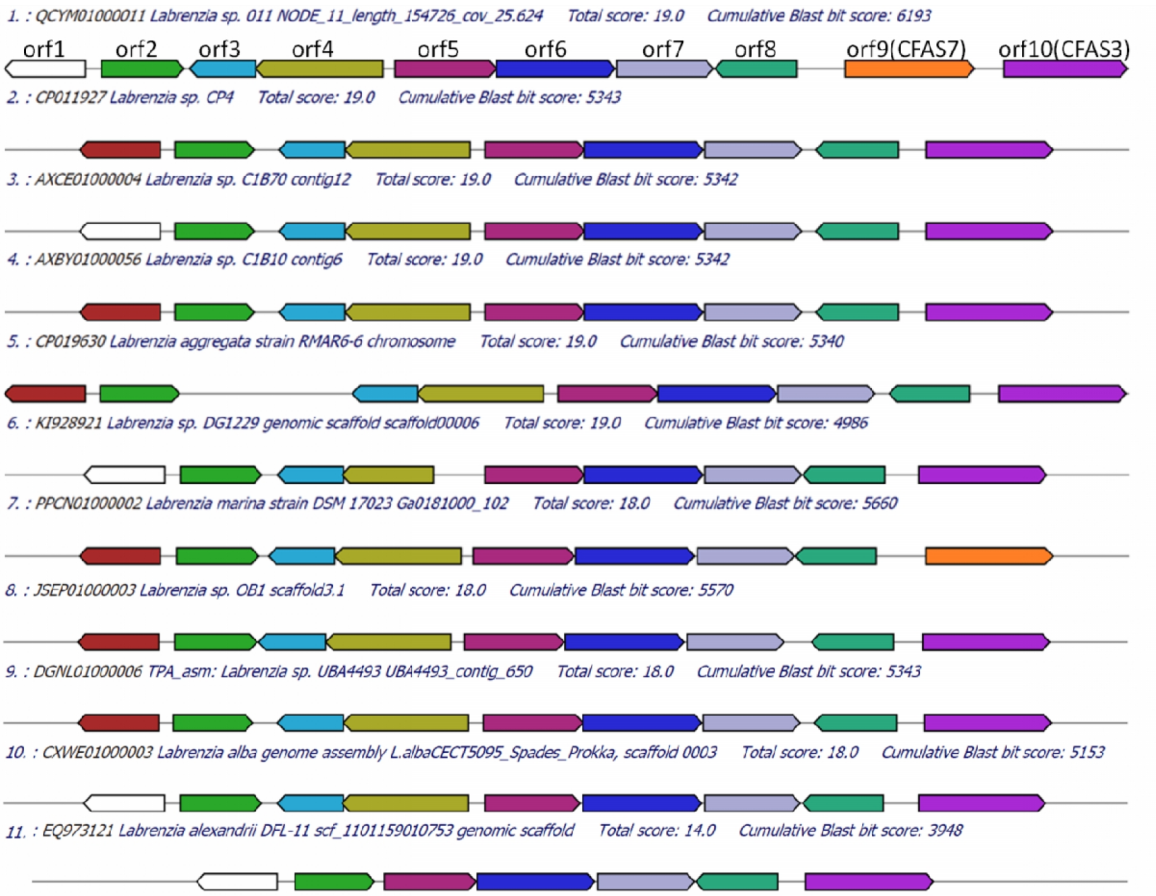


Figure 6.12 CAFS3 gene cluster alignment in *Labrenzia* strains, orf1: metallo-beta-lactamase, orf2: methyltransferase, orf3: methyltransferase, orf4: homoserine_O-acetyltransferase, orf5: Chorismate mutase, orf6: aminotransferase, orf7: 3-hydroxyisobutyrate dehydrogenase, orf8: Polyprenyl synthetase, orf9: Cyclopropane fatty acyl-phospholipid synthase (CFAS7), and orf10: Cyclopropane fatty acyl-phospholipid synthase (CFAS3). Same colors represent the same annotation.

6.3 **Discussion III**

The biosynthesis of cyclopropane fatty acids uses unsaturated fatty acids as a substrate (Guangqi et al. 2010). Their modification, catalyzed by the CFAS enzymes, occurs in many bacteria and is recognized to play an important role in the adaptation of bacteria to drastic environmental perturbation such as acid or freeze-drying stress (To et al. 2011; Velly et al. 2015). The cyclopropanation reaction of unsaturated lipids is well described for long chain fatty acids of *M. tuberculosis* and *E. coli*, and mainly associated with bacterial membranes (Guangqi et al. 2010; Xu et al. 2018). In general, bacterial CFASs catalyze the transfer of a methyl group from SAM to an inactivated double bond of a lipid chain, followed by deprotonation of the newly attached methyl group and ring closure to form a cyclopropane ring (Guangqi et al. 2010). CFAS genes have highly conserved amino acid sequences and several clades of functional CFAS genes have been identified by phylogenetic analysis (Grogan and Cronan 1997; Banerjee et al. 2011). Species of genus *Labrenzia* possess several different CFAS genes in their genomes, which indicates their ability to biosynthesize different cyclopropane fatty acids (Biebl et al. 2007).

In *Labrenzia* sp. 011 four different CFAS clades were identified and a phylogenetic analysis revealed their relationship with known functional CFAS genes. This analysis grouped the CFAS1 gene with UfaA1 from *M. tuberculosis* H_{37}Rv, which synthesizes 19:0Me10 from 18:1ω9 and NADPH (Machida et al. 2017). Heterologous expression of UfaA1 revealed this enzyme to be co-located with an FAD-dependent oxidoreductase, which contributes to the reduction reaction that is required for the conversion of 10-methylene octadecanoic acid to 19:0Me10 (Machida et al. 2017). Interestingly, in the CFAS1 clade, which is the closest clade to *ufaA1*, the orf2 in the vicinity of CFAS1 is also a FAD-dependent oxidoreductase. Therefore, members of this CFAS1 clade probably perform the same function in *Labrenzia* strains. The fatty acids 11-methyl 18:ω6t and 11-methyl 20:ω6t were indeed isolated from different *Labrenzia* strains (Biebl et al. 2007).

On the other hand, the CFAS2 clade was grouped with SAM-dependent methyltransferases of *Mycobacterium* strains, which are involved in the mycolic acid biosynthesis pathway (Banerjee et al. 2011). Mycolic acids are long-chain fatty acids

contain several cyclopropane rings and found in the hydrophobic cell walls of Mycobacteria. Functional cyclopropane groups are introduced to the mycolate chain by numerous cyclopropane synthases (Banerjee et al. 2011). The fatty acids cyclo 21:0 and cyclo 19:0 were formerly isolated from *L. alexandrii* and *L. marina* (Biebl et al. 2007). Therefore, production of such long chain cyclopropane fatty acids is probably related to the CFAS2 clade in *Labrenzia* strains.

To the best of our knowledge, there is no report of CFASs, which catalyze the cyclopropanation of medium-chain fatty acids such as **1** and **2** in bacteria. Therefore, the CFAS3 and CFAS7 in *Labrenzia* sp. 011 are probably responsible for the biosynthesis of these molecules. Additionally, the presence of aminotransferase and chorismate mutase genes in the respective CFAS gene cluster (Figure 4C) is probably related to hitherto unknown compounds with a structure supposedly similar to that of grenadamide (Figure 1) from *M. producens* (Kwan et al. 2011; Sitachitta and Gerwick 1998), *i.e*. the cyclopropane containing fatty acids **1** and **2** may also occur as amides. However, experimental studies such as knock-outs of the potential biosynthetic genes have to be performed in order to confirm the biosynthetic origin of compounds **1** and **2**, and the presence of further derivatives of the latter.

7 Summary

Marine bacteria have come into the focus for natural product discovery as a consequence of the emergence of antibiotic resistance. This was additionally boosted by the limitations encountered in drug developments from traditional producers of drug leads from terrestrial environments. The comparative genomic/metabolomic analysis is considered as a new and powerful tool to disclose the potential of microbes for the biosynthesis of novel specialized metabolites.

In the group of marine myxobacteria, only a limited number of isolated species and sequenced genomes is available. However, the few compounds isolated thereof to date show interesting bioactivities and even novel chemical scaffolds, indicating a huge potential for natural product discovery. In this study, all marine myxobacteria with accessible genome data (n=5), including *Haliangium ochraceum* DSM 14365, *Plesiocystis pacifica* DSM 14875*, Enhygromyxa salina* DSM 15201, and the two newly sequenced species *Enhygromyxa salina* SWB005 and SWB007, were analyzed.

All of these genomes are large (~10 Mb), with a relatively small core genome and many unique coding sequences in each strain. Genome analysis revealed a high variety of biosynthetic gene clusters (BGCs) between the strains (Figure 1A, in summary), and several resistance models and essential core genes indicated the potential to biosynthesize antimicrobial molecules. Polyketides (PKs) and terpenes represented the majority of predicted specialized metabolite BGCs and contributed to the highest share between the strains. BGCs coding for non-ribosomal peptides (NRPs), PK/NRP hybrids and ribosomally synthesized and post-translationally modified peptides were mostly strain specific.

These results were in line with the metabolomic analysis, which revealed a high diversity of the chemical features between the strains (Figure 1B, in summary). Only 6-11% of the metabolome was shared between all the investigated strains, which correlates to the small core genome of these bacteria (13-16% of each genome). In addition, the compound enhygrolide A, known from *E. salina* SWB005, was detected for the first time and structurally elucidated from *E. salina* SWB006. The here acquired data corroborate that these microorganisms represent a most promising source for the detection of novel specialized metabolites.

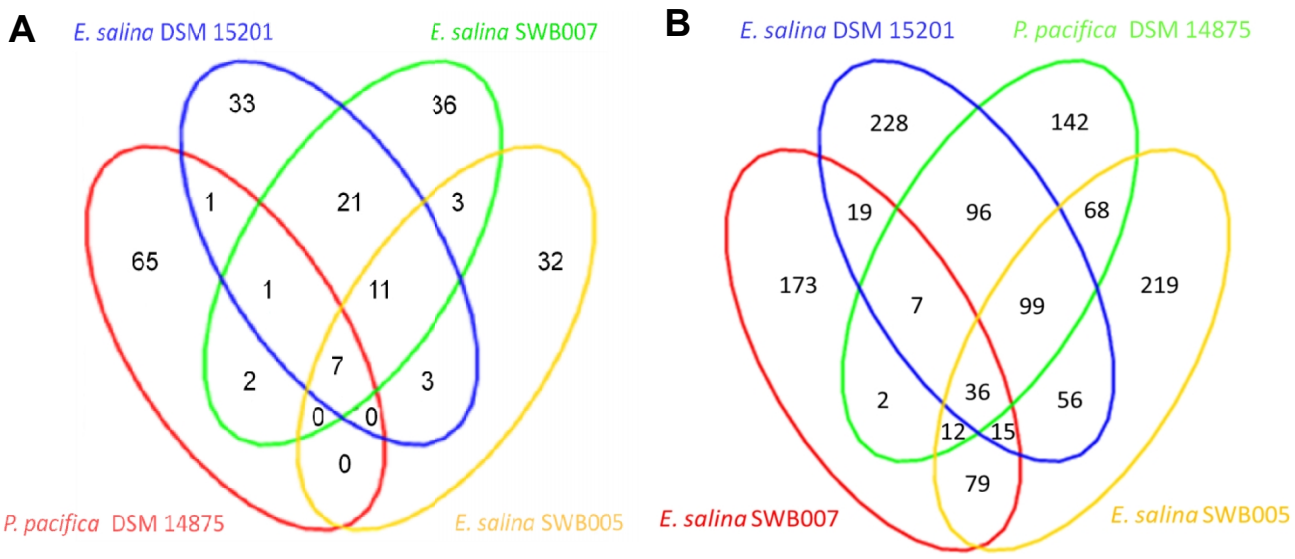


Figure1. A: Venn diagram displaying biosynthetic gene clusters (BGCs) counts according to distribution in closely related strains of marine myxobacteria. (acquired from the similarity network of the predicted BGCs in the analyzed genomes). B: Venn diagram displaying chemical features counts according to detection from single or multiple strains of marine myxobacteria. (acquired from the molecular network of analyzed strains).

To date, only a few myxobacteria are known as marine, owing to the salt dependency of their growth. Thus, the salt tolerance mechanism of these bacteria was investigated. To this end, a growth medium was designed in which the mutated *Escherichia coli* strain BKA13 served as a sole food source for the predatory, heterotrophic myxobacteria. This enabled measurement of the osmolytes without any background and revealed that the closely related strains *E. salina* SWB007 and *P. pacifica* DSM 14875 had developed different strategies to handle salt stress. *P. pacifica* DSM 14875, which was grown

between 1 and 4% NaCl, relies solely on the accumulation of amino acids, while *E. salina* SWB007, which was grown between 0.5 and 3% NaCl, employs, besides betaine, hydroxyectoine as the major compatible solute. In accordance with this analysis, only in the latter strain was a locus identified that codes for genes corresponding to the biosynthesis of betaine, ectoine, and hydroxyectoine.

The second group of marine bacteria investigated here belongs to the family Rhodobacteraceae and is frequently encountered in the oceans. The marine bacterium *Labrenzia* sp. 011 was isolated from the coastal sediment of Kronsgaard, Germany. This strain produces two cyclopropane-containing medium-chain fatty acids, namely *cis*-4-(2-hexylcyclopropyl)-butanoic acid and *cis*-2-(2-hexylcyclopropyl)-acetic acid (Figure 2A, in summary), which showed activity against a range of microorganisms. It is of special interest, that these compounds strongly inhibit *Pseudoroseovarius crassostreae* DSM 16950 (genus *Roseovarius*), the causative agent of Roseovarius oyster disease (ROD). The latter is a bacterial-induced infection and causes major problems in oyster aquaculture. Bacteria of the genus *Labrenzia* have been proposed as protective agents against ROD.

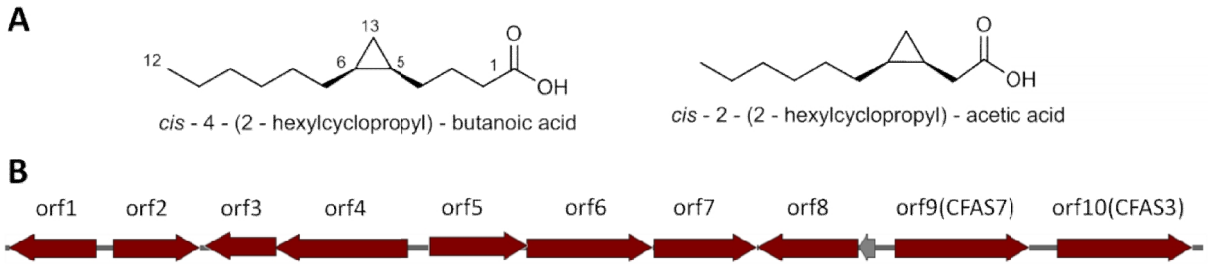


Figure 2. A: Cyclopropane fatty acids from *Labrenzia* sp. 011. B: Proposed biosynthetic gene cluster of the isolated cyclopropane fatty acids.

The genome analysis of bacteria of the genus *Labrenzia* was expected to provide information to understand the mollusk-protective role of *Labrenzia* spp.. Therefore, the genome of *Labrenzia* sp. 011 was sequenced and assembled into 65 contigs. It has a size of 5.1 Mbp and a G+C content of 61.6%. Comparative genome analysis defined *Labrenzia* sp. 011 as a distinct new species within the genus *Labrenzia*, whereby 44% of the genome was contributed to the *Labrenzia* core genome.

The genomic analysis revealed several conserved cyclopropane fatty acid synthases (CFAS) genes in bacteria of the genus *Labrenzia*, putatively responsible for methylation and cyclopropanation of long-chain fatty acids. In addition, a gene cluster encoding for two distinct CFAS genes is proposed as the biosynthetic origin of the isolated cyclopropane fatty acids from *Labreniza* sp. 011 (Figure 2B, in summary).

In conclusion, each of the investigated marine bacteria, myxobacteria and *Labrenzia,* harbors a high unique genetic and metabolic diversity, rendering these groups of microorganisms promising sources for novel specialized metabolites.

8 References

Agarwal, Vinayak; Blanton, Jessica M.; Podell, Sheila; Taton, Arnaud; Schorn, Michelle A.; Busch, Julia et al. (2017): Metagenomic discovery of polybrominated diphenyl ether biosynthesis by marine sponges. In *Nature chemical biology* 13 (5), pp. 537–543. DOI: 10.1038/nchembio.2330.

Alanjary, Mohammad; Kronmiller, Brent; Adamek, Martina; Blin, Kai; Weber, Tilmann; Huson, Daniel et al. (2017): The Antibiotic Resistant Target Seeker (ARTS), an exploration engine for antibiotic cluster prioritization and novel drug target discovery. In *Nucleic acids research. DOI:* 10.1093/nar/gkx360.

Albataineh, Hanan; Stevens, D. Cole (2018): Marine Myxobacteria. A Few Good Halophiles. In *Marine drugs* 16 (6). DOI: 10.3390/md16060209.

Amico, Vincenzo; Cunsolo, Francesca; Piattelli, Mario; Ruberto, Giuseppe (1985): Acyclic tetraprenyltoluquinols from Cystoseira sauvageuana and their possible role as biogenetic precursors of the cyclic cystoseira metabolites. In *Phytochemistry* 24 (11), pp. 2663–2668. DOI: 10.1016/S0031-9422(00)80689-5.

Amiri Moghaddam, Jamshid; Boehringer, Nils; Burdziak, Amal; Kunte, Hans-Jörg; Galinski, Erwin A.; Schäberle, Till F. (2016): Different strategies of osmoadaptation in the closely related marine myxobacteria Enhygromyxa salina SWB007 and Plesiocystis pacifica SIR-1. In *Microbiology (Reading, England). DOI:* 10.1099/mic.0.000250.

Amiri Moghaddam, Jamshid; Dávila-Céspedes, Antonio; Alanjary, Mohammad; Blom, Jochen; König, Gabriele M.; Schäberle, Till F. (under review): Draft Genome Sequence of the Marine Bacterium Labrenzia sp. 011, a Potential Protective Agent of Mollusks. Data in Brief.

Amiri Moghaddam, Jamshid; Poehlein, Anja; Fisch, Katja; Alanjary, Mohammad; Daniel, Rolf; König, Gabriele M.; Schäberle, Till F. (2018): Draft Genome Sequences of the

Obligatory Marine Myxobacterial Strains Enhygromyxa salina SWB005 and SWB007. In *Genome Announc.* 6 (17). DOI: 10.1128/genomeA.00324-18.

Arfken, Ann; Song, Bongkeun; Bowman, Jeff S.; Piehler, Michael (2017): Denitrification potential of the eastern oyster microbiome using a 16S rRNA gene based metabolic inference approach. In *PloS one* 12 (9), e0185071. DOI: 10.1371/journal.pone.0185071.

Arnison, Paul G.; Bibb, Mervyn J.; Bierbaum, Gabriele; Bowers, Albert A.; Bugni, Tim S.; Bulaj, Grzegorz et al. (2013): Ribosomally synthesized and post-translationally modified peptide natural products. Overview and recommendations for a universal nomenclature. In *Natural product reports* 30 (1), pp. 108–160. DOI: 10.1039/c2np20085f.

Aziz, Ramy K.; Bartels, Daniela; Best, Aaron A.; DeJongh, Matthew; Disz, Terrence; Edwards, Robert A. et al. (2008): The RAST Server. Rapid annotations using subsystems technology. In *BMC genomics* 9, p. 75. DOI: 10.1186/1471-2164-9-75.

Banerjee, Ruma; Vats, Pankaj; Dahale, Sonal; Kasibhatla, Sunitha Manjari; Joshi, Rajendra (2011): Comparative genomics of cell envelope components in mycobacteria. In *PloS one* 6 (5), e19280. DOI: 10.1371/journal.pone.0019280.

Bankevich, Anton; Nurk, Sergey; Antipov, Dmitry; Gurevich, Alexey A.; Dvorkin, Mikhail; Kulikov, Alexander S. et al. (2012): SPAdes. A new genome assembly algorithm and its applications to single-cell sequencing. In *Journal of computational biology : a journal of computational molecular cell biology* 19 (5), pp. 455–477. DOI: 10.1089/cmb.2012.0021.

Barkan, Daniel; Rao, Vivek; Sukenick, George D.; Glickman, Michael S. (2010): Redundant function of cmaA2 and mmaA2 in Mycobacterium tuberculosis cis cyclopropanation of oxygenated mycolates. In *Journal of bacteriology* 192 (14), pp. 3661–3668. DOI: 10.1128/JB.00312-10.

Biebl, Hanno; Pukall, Rüdiger; Lünsdorf, Heinrich; Schulz, Stefan; Allgaier, Martin; Tindall, Brian J.; Wagner-Döbler, Irene (2007): Description of Labrenzia alexandrii gen. nov., sp. nov., a novel alphaproteobacterium containing bacteriochlorophyll a,

and a proposal for reclassification of Stappia aggregata as Labrenzia aggregata comb. nov., of Stappia marina as Labrenzia marina comb. nov. and of Stappia alba as Labrenzia alba comb. nov., and emended descriptions of the genera Pannonibacter, Stappia and Roseibium, and of the species Roseibium denhamense and Roseibium hamelinense. In *International journal of systematic and evolutionary microbiology* 57 (Pt 5), pp. 1095–1107. DOI: 10.1099/ijs.0.64821-0.

Blin, Kai; Medema, Marnix H.; Kottmann, Renzo; Lee, Sang Yup; Weber, Tilmann (2017a): The antiSMASH database, a comprehensive database of microbial secondary metabolite biosynthetic gene clusters. In *Nucleic acids research* 45 (D1), D555-D559. DOI: 10.1093/nar/gkw960.

Blin, Kai; Wolf, Thomas; Chevrette, Marc G.; Lu, Xiaowen; Schwalen, Christopher J.; Kautsar, Satria A. et al. (2017b): antiSMASH 4.0-improvements in chemistry prediction and gene cluster boundary identification. In *Nucleic acids research. DOI:* 10.1093/nar/gkx319.

Blom, Jochen; Kreis, Julian; Spänig, Sebastian; Juhre, Tobias; Bertelli, Claire; Ernst, Corinna; Goesmann, Alexander (2016): EDGAR 2.0. An enhanced software platform for comparative gene content analyses. In *Nucleic acids research* 44 (W1), W22-8. DOI: 10.1093/nar/gkw255.

Blunt, John W.; Copp, Brent R.; Munro, Murray H. G.; Northcote, Peter T.; Prinsep, Michèle R. (2004): Marine natural products. In *Natural product reports* 21 (1), pp. 1–49. DOI: 10.1039/B305250H.

Boettcher, K. J.; Barber, B. J.; Singer, J. T. (2000): Additional Evidence that Juvenile Oyster Disease Is Caused by a Member of the Roseobacter Group and Colonization of Nonaffected Animals by Stappia stellulata-Like Strains. In *Applied and environmental microbiology* 66 (9), pp. 3924–3930. DOI: 10.1128/AEM.66.9.3924-3930.2000.

Bolger, Anthony M.; Lohse, Marc; Usadel, Bjoern (2014): Trimmomatic. A flexible trimmer for Illumina sequence data. In *Bioinformatics (Oxford, England)* 30 (15), pp. 2114–2120. DOI: 10.1093/bioinformatics/btu170.

Bouhired, Sarah M.; Crüsemann, Max; Almeida, Celso; Weber, Tilmann; Piel, Jörn; Schäberle, Till F.; König, Gabriele M. (2014): Biosynthesis of phenylnannolone A, a multidrug resistance reversal agent from the halotolerant myxobacterium Nannocystis pusilla B150. In *Chembiochem : a European journal of chemical biology* 15 (5), pp. 757–765. DOI: 10.1002/cbic.201300676.

Boyd, Kenneth G.; Adams, David R.; Burgess, J. Grant (1999): Antibacterial and repellent activities of marine bacteria associated with algal surfaces. In *Biofouling* 14 (3), pp. 227–236. DOI: 10.1080/08927019909378414.

Brinkhoff, Thorsten; Fischer, Doreen; Vollmers, John; Voget, Sonja; Beardsley, Christine; Thole, Sebastian et al. (2012): Biogeography and phylogenetic diversity of a cluster of exclusively marine myxobacteria. In *The ISME journal* 6 (6), pp. 1260–1272. DOI: 10.1038/ismej.2011.190.

Brown, A. D. (1976): Microbial water stress. In *Bacteriological reviews* 40 (4), pp. 803–846.

Buntin, Kathrin; Irschik, Herbert; Weissman, Kira J.; Luxenburger, Eva; Blöcker, Helmut; Müller, Rolf (2010): Biosynthesis of thuggacins in myxobacteria. Comparative cluster analysis reveals basis for natural product structural diversity. In *Chemistry & biology* 17 (4), pp. 342–356. DOI: 10.1016/j.chembiol.2010.02.013.

Burg, Maurice B.; Ferraris, Joan D. (2008): Intracellular organic osmolytes. Function and regulation. In *The Journal of biological chemistry* 283 (12), pp. 7309–7313. DOI: 10.1074/jbc.R700042200.

Bursy, Jan; Kuhlmann, Anne U.; Pittelkow, Marco; Hartmann, Holger; Jebbar, Mohamed; Pierik, Antonio J.; Bremer, Erhard (2008): Synthesis and uptake of the compatible solutes ectoine and 5-hydroxyectoine by Streptomyces coelicolor A3(2) in response to salt and heat stresses. In *Applied and environmental microbiology* 74 (23), pp. 7286–7296. DOI: 10.1128/AEM.00768-08.

Casadaban, M. J. (1976): Transposition and fusion of the lac genes to selected promoters in Escherichia coli using bacteriophage lambda and Mu. In *Journal of molecular biology* 104 (3), pp. 541–555.

Ceniceros, Ana; Dijkhuizen, Lubbert; Petrusma, Mirjan; Medema, Marnix H. (2017): Genome-based exploration of the specialized metabolic capacities of the genus Rhodococcus. In *BMC genomics* 18 (1), p. 593. DOI: 10.1186/s12864-017-3966-1.

Chan, Yolande A.; Podevels, Angela M.; Kevany, Brian M.; Thomas, Michael G. (2009): Biosynthesis of polyketide synthase extender units. In *Nat. Prod. Rep.* 26 (1), pp. 90–114. DOI: 10.1039/B801658P.

Chen, I-Min A.; Markowitz, Victor M.; Chu, Ken; Palaniappan, Krishna; Szeto, Ernest; Pillay, Manoj et al. (2017): IMG/M. Integrated genome and metagenome comparative data analysis system. In *Nucleic acids research* 45 (D1), D507-D516. DOI: 10.1093/nar/gkw929.

Cimermancic, Peter; Medema, Marnix H.; Claesen, Jan; Kurita, Kenji; Wieland Brown, Laura C.; Mavrommatis, Konstantinos et al. (2014): Insights into secondary metabolism from a global analysis of prokaryotic biosynthetic gene clusters. In *Cell* 158 (2), pp. 412–421. DOI: 10.1016/j.cell.2014.06.034.

Clare, Anthony (1996): Marine natural product antifoulants. Status and potential. In *Biofouling* 9 (3), pp. 211–229. DOI: 10.1080/08927019609378304.

Colston, Sophie M.; Fullmer, Matthew S.; Beka, Lidia; Lamy, Brigitte; Gogarten, J. Peter; Graf, Joerg (2014): Bioinformatic genome comparisons for taxonomic and phylogenetic assignments using Aeromonas as a test case. In *mBio* 5 (6), e02136. DOI: 10.1128/mBio.02136-14.

Cotter, Paul D.; Ross, R. Paul; Hill, Colin (2013): Bacteriocins - a viable alternative to antibiotics? In *Nature reviews. Microbiology* 11 (2), pp. 95–105. DOI: 10.1038/nrmicro2937.

Courtois, Fabienne; Guérard, Christine; Thomas, Xavier; Ploux, Olivier (2004): Escherichia coli cyclopropane fatty acid synthase. In *European journal of biochemistry* 271 (23-24), pp. 4769–4778. DOI: 10.1111/j.1432-1033.2004.04441.x.

Crüsemann, Max; O'Neill, Ellis C.; Larson, Charles B.; Melnik, Alexey V.; Floros, Dimitrios J.; da Silva, Ricardo R. et al. (2017): Prioritizing Natural Product Diversity in

a Collection of 146 Bacterial Strains Based on Growth and Extraction Protocols. In *Journal of natural products* 80 (3), pp. 588–597. DOI: 10.1021/acs.jnatprod.6b00722.

Curson, Andrew R. J.; Liu, Ji; Bermejo Martínez, Ana; Green, Robert T.; Chan, Yohan; Carrión, Ornella et al. (2017): Dimethylsulfoniopropionate biosynthesis in marine bacteria and identification of the key gene in this process. In *Nature microbiology* 2, p. 17009. DOI: 10.1038/nmicrobiol.2017.9.

Curtis, Thomas P.; Sloan, William T.; Scannell, Jack W. (2002): Estimating prokaryotic diversity and its limits. In *Proceedings of the National Academy of Sciences of the United States of America* 99 (16), pp. 10494–10499. DOI: 10.1073/pnas.142680199.

da Costa, M. S.; Santos, H.; Galinski, E. A. (1998): An overview of the role and diversity of compatible solutes in Bacteria and Archaea. In *Advances in biochemical engineering/biotechnology* 61, pp. 117–153.

Dang, Hongyue; Li, Tiegang; Chen, Mingna; Huang, Guiqiao (2008): Cross-ocean distribution of Rhodobacterales bacteria as primary surface colonizers in temperate coastal marine waters. In *Applied and environmental microbiology* 74 (1), pp. 52–60. DOI: 10.1128/AEM.01400-07.

Dávila-Céspedes, Antonio; Hufendiek, Peter; Crüsemann, Max; Schäberle, Till F.; König, Gabriele M. (2016): Marine-derived myxobacteria of the suborder Nannocystineae. An underexplored source of structurally intriguing and biologically active metabolites. In *Beilstein journal of organic chemistry* 12, pp. 969–984. DOI: 10.3762/bjoc.12.96.

Dawid, W. (2000): Biology and global distribution of myxobacteria in soils. In *FEMS microbiology reviews* 24 (4), pp. 403–427.

Debbab, Abdessamad; Aly, Amal H.; Lin, Wen H.; Proksch, Peter (2010): Bioactive compounds from marine bacteria and fungi. In *Microbial biotechnology* 3 (5), pp. 544–563. DOI: 10.1111/j.1751-7915.2010.00179.x.

Desjardine, Kelsey; Pereira, Alban; Wright, Helen; Matainaho, Teatulohi; Kelly, Michael; Andersen, Raymond J. (2007): Tauramamide, a lipopeptide antibiotic produced in culture by Brevibacillus laterosporus isolated from a marine habitat. Structure

elucidation and synthesis. In *Journal of natural products* 70 (12), pp. 1850–1853. DOI: 10.1021/np070209r.

Desmond, Elie; Gribaldo, Simonetta (2009): Phylogenomics of sterol synthesis. Insights into the origin, evolution, and diversity of a key eukaryotic feature. In *Genome biology and evolution* 1, pp. 364–381. DOI: 10.1093/gbe/evp036.

Dias, Daniel A.; Urban, Sylvia; Roessner, Ute (2012): A historical overview of natural products in drug discovery. In *Metabolites* 2 (2), pp. 303–336. DOI: 10.3390/metabo2020303.

Dickschat, Jeroen S.; Nawrath, Thorben; Thiel, Verena; Kunze, Brigitte; Müller, Rolf; Schulz, Stefan (2007): Biosynthesis of the off-flavor 2-methylisoborneol by the myxobacterium Nannocystis exedens. In *Angewandte Chemie (International ed. in English)* 46 (43), pp. 8287–8290. DOI: 10.1002/anie.200702496.

Duboc, Philippe; Mollet, Beat (2001): Applications of exopolysaccharides in the dairy industry. In *International Dairy Journal* 11 (9), pp. 759–768. DOI: 10.1016/S0958-6946(01)00119-4.

Edgar, Robert C. (2004): MUSCLE. Multiple sequence alignment with high accuracy and high throughput. In *Nucleic acids research* 32 (5), pp. 1792–1797. DOI: 10.1093/nar/gkh340.

Ellis, L. L.; Burcham, J. M.; Paynter, K. T.; Bishop, S. H. (1985): Amino acid metabolism in euryhaline bivalves. Regulation of glycine accumulation in ribbed mussel gills. In *The Journal of experimental zoology* 233 (3), pp. 347–358. DOI: 10.1002/jez.1402330303.

Engene, Niclas; Rottacker, Erin C.; Kaštovský, Jan; Byrum, Tara; Choi, Hyukjae; Ellisman, Mark H. et al. (2012): Moorea producens gen. nov., sp. nov. and Moorea bouillonii comb. nov., tropical marine cyanobacteria rich in bioactive secondary metabolites. In *International journal of systematic and evolutionary microbiology* 62 (Pt 5), pp. 1171–1178. DOI: 10.1099/ijs.0.033761-0.

Felder, Stephan; Dreisigacker, Sandra; Kehraus, Stefan; Neu, Edith; Bierbaum, Gabriele; Wright, Patrick R. et al. (2013a): Salimabromide. Unexpected chemistry

from the obligate marine myxobacterium Enhygromxya salina. In *Chemistry (Weinheim an der Bergstrasse, Germany)* 19 (28), pp. 9319–9324. DOI: 10.1002/chem.201301379.

Felder, Stephan; Kehraus, Stefan; Neu, Edith; Bierbaum, Gabriele; Schäberle, Till F.; König, Gabriele M. (2013b): Salimyxins and enhygrolides. Antibiotic, sponge-related metabolites from the obligate marine myxobacterium Enhygromyxa salina. In *Chembiochem : a European journal of chemical biology* 14 (11), pp. 1363–1371. DOI: 10.1002/cbic.201300268.

Feling, Robert H.; Buchanan, Greg O.; Mincer, Tracy J.; Kauffman, Christopher A.; Jensen, Paul R.; Fenical, William (2003): Salinosporamide A. A highly cytotoxic proteasome inhibitor from a novel microbial source, a marine bacterium of the new genus salinospora. In *Angewandte Chemie (International ed. in English)* 42 (3), pp. 355–357. DOI: 10.1002/anie.200390115.

Felsenstein, Joe (1989): PHYLIP-phylogeny inference package (version 3.2). In *Cladistics* 5 (2), pp. 163–166. DOI: 10.1111/j.1096-0031.1989.tb00562.x.

Fernández-Gómez, Beatriz; Fernàndez-Guerra, Antonio; Casamayor, Emilio O.; González, José M.; Pedrós-Alió, Carlos; Acinas, Silvia G. (2012): Patterns and architecture of genomic islands in marine bacteria. In *BMC genomics* 13, p. 347. DOI: 10.1186/1471-2164-13-347.

Fisch, Katja M.; Schäberle, Till F. (2016): Toolbox for Antibiotics Discovery from Microorganisms. In *Archiv der Pharmazie* 349 (9), pp. 683–691. DOI: 10.1002/ardp.201600064.

Fu, Yunyun; Keats, Kimberley F.; Rivkin, Richard B.; Lang, Andrew S. (2013): Water mass and depth determine the distribution and diversity of Rhodobacterales in an Arctic marine system. In *FEMS microbiology ecology* 84 (3), pp. 564–576. DOI: 10.1111/1574-6941.12085.

Fudou, Ryosuke; Jojima, Yasuko; Iizuka, Takashi; Yamanaka, Shigeru (2002a): Haliangium ochraceum gen. nov., sp. nov. and Haliangium tepidum sp. nov. Novel moderately halophilic myxobacteria isolated from coastal saline environments. In

The Journal of general and applied microbiology 48 (2), pp. 109–115. DOI: 10.2323/jgam.48.109.

Fudou, Ryosuke; Jojima, Yasuko; Iizuka, Takashi; Yamanaka, Shigeru (2002b): Haliangium ochraceum gen. nov., sp. nov. and Haliangium tepidum sp. nov. Novel moderately halophilic myxobacteria isolated from coastal saline environments. In *The Journal of general and applied microbiology* 48 (2), pp. 109–116.

Galinski, Erwin A.; Herzog, Ruth M. (1990): The role of trehalose as a substitute for nitrogen-containing compatible solutes (Ectothiorhodospira halochloris). In *Arch. Microbiol.* 153 (6), pp. 607–613. DOI: 10.1007/BF00245273.

Galinski, Erwin A.; PFEIFFER, Heinz-Peter; TRUPER, Hans G. (1985): 1,4,5,6-Tetrahydro-2-methyl-4-pyrimidinecarboxylic acid. A novel cyclic amino acid from halophilic phototrophic bacteria of the genus Ectothiorhodospira. In *Eur J Biochem* 149 (1), pp. 135–139. DOI: 10.1111/j.1432-1033.1985.tb08903.x.

Ganapiriya, Viswambaran; Maharajan, Athisuyambulingam; Kumarasamy, Pothiappan (2012): Antifouling effect of bioactive compounds from marine sponge Acanthella elongata and different species of bacterial film on larval attachment of Balanus amphitrite (cirripedia, crustacea). In *Braz. arch. biol. technol.* 55 (3), pp. 395–402. DOI: 10.1590/S1516-89132012000300010.

García-Alcalde, Fernando; Okonechnikov, Konstantin; Carbonell, José; Cruz, Luis M.; Götz, Stefan; Tarazona, Sonia et al. (2012): Qualimap. Evaluating next-generation sequencing alignment data. In *Bioinformatics (Oxford, England)* 28 (20), pp. 2678–2679. DOI: 10.1093/bioinformatics/bts503.

Gehring, A. M.; Mori, I.; Walsh, C. T. (1998): Reconstitution and characterization of the Escherichia coli enterobactin synthetase from EntB, EntE, and EntF. In *Biochemistry* 37 (8), pp. 2648–2659. DOI: 10.1021/bi9726584.

Ghani, Norshazliza Ab; Norizan, Siti Nur Maisarah; Chan, Xin Yue; Yin, Wai-Fong; Chan, Kok-Gan (2014): Labrenzia sp. BM1. A quorum quenching bacterium that degrades N-acyl homoserine lactones via lactonase activity. In *Sensors (Basel, Switzerland)* 14 (7), pp. 11760–11769. DOI: 10.3390/s140711760.

Giglio, Steven; Jiang, Jiaoyang; Saint, Christopher P.; Cane, David E.; Monis, Paul T. (2008): Isolation and characterization of the gene associated with geosmin production in cyanobacteria. In *Environmental science & technology* 42 (21), pp. 8027–8032.

Gil-Turnes, M. S.; Fenical, W. (1992): Embryos of Homarus americanus are Protected by Epibiotic Bacteria. In *The Biological bulletin* 182 (1), pp. 105–108. DOI: 10.2307/1542184.

Gil-Turnes, M. S.; Hay, M. E.; Fenical, W. (1989): Symbiotic marine bacteria chemically defend crustacean embryos from a pathogenic fungus. In *Science (New York, N.Y.)* 246 (4926), pp. 116–118.

Giordano, Daniela; Coppola, Daniela; Russo, Roberta; Denaro, Renata; Giuliano, Laura; Lauro, Federico M. et al. (2015): Marine Microbial Secondary Metabolites. Pathways, Evolution and Physiological Roles. In *Advances in microbial physiology* 66, pp. 357–428. DOI: 10.1016/bs.ampbs.2015.04.001.

Goris, Johan; Konstantinidis, Konstantinos T.; Klappenbach, Joel A.; Coenye, Tom; Vandamme, Peter; Tiedje, James M. (2007): DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. In *International journal of systematic and evolutionary microbiology* 57 (Pt 1), pp. 81–91. DOI: 10.1099/ijs.0.64483-0.

Grammann, K.; Volke, A.; Kunte, H. J. (2002): New Type of Osmoregulated Solute Transporter Identified in Halophilic Members of the Bacteria Domain. TRAP Transporter TeaABC Mediates Uptake of Ectoine and Hydroxyectoine in Halomonas elongata DSM 2581T. In *Journal of bacteriology* 184 (11), pp. 3078–3085. DOI: 10.1128/JB.184.11.3078-3085.2002.

Grant, Jason R.; Arantes, Adriano S.; Stothard, Paul (2012): Comparing thousands of circular genomes using the CGView Comparison Tool. In *BMC genomics* 13, p. 202. DOI: 10.1186/1471-2164-13-202.

Grogan, D. W.; Cronan, J. E. (1997): Cyclopropane ring formation in membrane lipids of bacteria. In *Microbiology and Molecular Biology Reviews* 61 (4), pp. 429–441.

Grünewald, Jan; Marahiel, Mohamed A. (2013): Nonribosomal Peptide Synthesis. In Abba J. Kastin (Ed.): Handbook of biologically active peptides. Second edition. Amsterdam: AP, pp. 138–149.

Guangqi, E.; Lesage, Denis; Ploux, Olivier (2010): Insight into the reaction mechanism of the Escherichia coli cyclopropane fatty acid synthase. Isotope exchange and kinetic isotope effects. In *Biochimie* 92 (10), pp. 1454–1457. DOI: 10.1016/j.biochi.2010.05.019.

Guo, Shoudong; Mao, Wenjun; Han, Yin; Zhang, Xiaohua; Yang, Chunli; Chen, Yin et al. (2010): Structural characteristics and antioxidant activities of the extracellular polysaccharides produced by marine bacterium Edwardsiella tarda. In *Bioresource technology* 101 (12), pp. 4729–4732. DOI: 10.1016/j.biortech.2010.01.125.

Gurevich, Alexey; Saveliev, Vladislav; Vyahhi, Nikolay; Tesler, Glenn (2013): QUAST. Quality assessment tool for genome assemblies. In *Bioinformatics (Oxford, England)* 29 (8), pp. 1072–1075. DOI: 10.1093/bioinformatics/btt086.

Guthals, Adrian; Watrous, Jeramie D.; Dorrestein, Pieter C.; Bandeira, Nuno (2012): The spectral networks paradigm in high throughput mass spectrometry. In *Molecular bioSystems* 8 (10), pp. 2535–2544. DOI: 10.1039/c2mb25085c.

Haardt, Martin; Kempf, Bettina; Faatz, Elke; Bremer, Erhard (1995): The osmoprotectant proline betaine is a major substrate for the binding-protein-dependent transport system ProU of Escherichia coli K-12. In *Molec. Gen. Genet.* 246 (6), pp. 783–796. DOI: 10.1007/BF00290728.

Hadjithomas, Michalis; Chen, I-Min A.; Chu, Ken; Huang, Jinghua; Ratner, Anna; Palaniappan, Krishna et al. (2017): IMG-ABC. New features for bacterial secondary metabolism analysis and targeted biosynthetic gene cluster discovery in thousands of microbial genomes. In *Nucleic acids research* 45 (D1), D560-D565. DOI: 10.1093/nar/gkw1103.

Harms, Henrik; Poehlein, Anja; Thürmer, Andrea; König, Gabriele M.; Schäberle, Till F. (2017): Draft Genome Sequence of Zobellia sp. Strain OII3, Isolated from the

Coastal Zone of the Baltic Sea. In *Genome Announc.* 5 (36), e00737-17. DOI: 10.1128/genomeA.00737-17.

Haroun-Bouhedja, F.; Ellouali, M.; Sinquin, C.; Boisson-Vidal, C. (2000): Relationship between sulfate groups and biological activities of fucans. In *Thrombosis research* 100 (5), pp. 453–459.

Held, Christoph; Neuhaus, Thorsten; Sadowski, Gabriele (2010): Compatible solutes. Thermodynamic properties and biological impact of ectoines and prolines. In *Biophysical chemistry* 152 (1-3), pp. 28–39. DOI: 10.1016/j.bpc.2010.07.003.

Herrmann, J.; Fayad, A. Abou; Müller, R. (2017): Natural products from myxobacteria. Novel metabolites and bioactivities. In *Natural product reports* 34 (2), pp. 135–160. DOI: 10.1039/c6np00106h.

Hertweck, Christian (2009): The biosynthetic logic of polyketide diversity. In *Angewandte Chemie (International ed. in English)* 48 (26), pp. 4688–4716. DOI: 10.1002/anie.200806121.

Hoffmann, Thomas; Krug, Daniel; Bozkurt, Nisa; Duddela, Srikanth; Jansen, Rolf; Garcia, Ronald et al. (2018): Correlating chemical diversity with taxonomic distance for discovery of natural products in myxobacteria. In *Nature Communications* 9 (1), p. 803. DOI: 10.1038/s41467-018-03184-1.

Huang, K. H.; Chen, B. Y.; Shen, F. T.; Young, C. C. (2012): Optimization of exopolysaccharide production and diesel oil emulsifying properties in root nodulating bacteria. In *World journal of microbiology & biotechnology* 28 (4), pp. 1367–1373. DOI: 10.1007/s11274-011-0936-7.

Huntley, Stuart; Hamann, Nils; Wegener-Feldbrügge, Sigrun; Treuner-Lange, Anke; Kube, Michael; Reinhardt, Richard et al. (2011): Comparative genomic analysis of fruiting body formation in Myxococcales. In *Molecular biology and evolution* 28 (2), pp. 1083–1097. DOI: 10.1093/molbev/msq292.

Iizuka, T.; Jojima, Y.; Fudou, R.; Yamanaka, S. (1998): Isolation of myxobacteria from the marine environment. In *FEMS microbiology letters* 169 (2), pp. 317–322.

Iizuka, Takashi; Jojima, Yasuko; Fudou, Ryosuke; Hiraishi, Akira; Ahn, Jong-Woong; Yamanaka, Shigeru (2003a): Plesiocystis pacifica gen. nov., sp. nov., a marine myxobacterium that contains dihydrogenated menaquinone, isolated from the Pacific coasts of Japan. In *International journal of systematic and evolutionary microbiology* 53 (Pt 1), pp. 189–195. DOI: 10.1099/ijs.0.02418-0.

Iizuka, Takashi; Jojima, Yasuko; Fudou, Ryosuke; Tokura, Mitsunori; Hiraishi, Akira; Yamanaka, Shigeru (2003b): Enhygromyxa salina gen. nov., sp. nov., a slightly halophilic myxobacterium isolated from the coastal areas of Japan. In *Systematic and applied microbiology* 26 (2), pp. 189–196. DOI: 10.1078/072320203322346038.

Inbar, L.; Lapidot, A. (1988): The structure and biosynthesis of new tetrahydropyrimidine derivatives in actinomycin D producer Streptomyces parvulus. Use of 13C- and 15N-labeled L-glutamate and 13C and 15N NMR spectroscopy. In *The Journal of biological chemistry* 263 (31), pp. 16014–16022.

Iniesta, Antonio A.; Cervantes, María; Murillo, Francisco J. (2007): Cooperation of two carotene desaturases in the production of lycopene in Myxococcus xanthus. In *The FEBS journal* 274 (16), pp. 4306–4314. DOI: 10.1111/j.1742-4658.2007.05960.x.

Ivanova, Natalia; Daum, Chris; Lang, Elke; Abt, Birte; Kopitz, Markus; Saunders, Elizabeth et al. (2010): Complete genome sequence of Haliangium ochraceum type strain (SMP-2). In *Standards in genomic sciences* 2 (1), pp. 96–106. DOI: 10.4056/sigs.69.1277.

Jensen, P. R.; Fenical, W. (1994): Strategies for the discovery of secondary metabolites from marine bacteria. Ecological perspectives. In *Annual review of microbiology* 48, pp. 559–584. DOI: 10.1146/annurev.mi.48.100194.003015.

Jiang, De-Ming; Kato, Chiaki; Zhou, Xiu-Wen; Wu, Zhi-Hong; Sato, Takako; Li, Yue-Zhong (2010): Phylogeographic separation of marine and soil myxobacteria at high levels of classification. In *The ISME journal* 4 (12), pp. 1520–1530. DOI: 10.1038/ismej.2010.84.

Joint, Ian; Mühling, Martin; Querellou, Joël (2010): Culturing marine bacteria - an essential prerequisite for biodiscovery. In *Microbial biotechnology* 3 (5), pp. 564–575. DOI: 10.1111/j.1751-7915.2010.00188.x.

Kanapathipillai, Mathumai; Lentzen, Georg; Sierks, Michael; Park, Chan Beum (2005): Ectoine and hydroxyectoine inhibit aggregation and neurotoxicity of Alzheimer's beta-amyloid. In *FEBS letters* 579 (21), pp. 4775–4780. DOI: 10.1016/j.febslet.2005.07.057.

Kaneshiro, E. S.; Holz, G. G.; Dunham, P. B. (1969): Osmoregulation in a marine ciliate, Miamiensis avidus. II. Regulation of intracellular free amino acids. In *The Biological bulletin* 137 (1), pp. 161–169. DOI: 10.2307/1539939.

Khan, Mohammad Sayyar; Yu, Xiang; Kikuchi, Akira; Asahina, Masashi; Watanabe, Kazuo N. (2009): Genetic engineering of glycine betaine biosynthesis to enhance abiotic stress tolerance in plants. In *Plant Biotechnology* 26 (1), pp. 125–134. DOI: 10.5511/plantbiotechnology.26.125.

Kimura, Yoshio; Kawasaki, Shinji; Yoshimoto, Hinae; Takegawa, Kaoru (2010): Glycine betaine biosynthesized from glycine provides an osmolyte for cell growth and spore germination during osmotic stress in Myxococcus xanthus. In *Journal of bacteriology* 192 (5), pp. 1467–1470. DOI: 10.1128/JB.01118-09.

Kopp, Maren; Irschik, Herbert; Gemperlein, Katja; Buntin, Kathrin; Meiser, Peter; Weissman, Kira J. et al. (2011): Insights into the complex biosynthesis of the leupyrrins in Sorangium cellulosum So ce690. In *Molecular bioSystems* 7 (5), pp. 1549–1563. DOI: 10.1039/c0mb00240b.

Krug, Daniel; Zurek, Gabriela; Revermann, Ole; Vos, Michiel; Velicer, Gregory J.; Müller, Rolf (2008): Discovering the hidden secondary metabolome of Myxococcus xanthus. A study of intraspecific diversity. In *Applied and environmental microbiology* 74 (10), pp. 3058–3068. DOI: 10.1128/AEM.02863-07.

Krzywinski, Martin; Schein, Jacqueline; Birol, Inanç; Connors, Joseph; Gascoyne, Randy; Horsman, Doug et al. (2009): Circos. An information aesthetic for

comparative genomics. In *Genome research* 19 (9), pp. 1639–1645. DOI: 10.1101/gr.092759.109.

Kunte, Hans; Lentzen, Georg; Galinski, Erwin (2014): Industrial Production of the Cell Protectant Ectoine. Protection Mechanisms, Processes, and Products. In *CBIOT* 3 (1), pp. 10–25. DOI: 10.2174/22115501113026660037.

Kwan, Jason Christopher; Meickle, Theresa; Ladwa, Dheran; Teplitski, Max; Paul, Valerie; Luesch, Hendrik (2011): Lyngbyoic acid, a "tagged" fatty acid from a marine cyanobacterium, disrupts quorum sensing in Pseudomonas aeruginosa. In *Molecular bioSystems* 7 (4), pp. 1205–1216. DOI: 10.1039/c0mb00180e.

Larsen, P. I.; Sydnes, L. K.; Landfald, B.; Strm, A. R. (1987): Osmoregulation in Escherichia coli by accumulation of organic osmolytes. Betaines, glutamic acid, and trehalose. In *Arch. Microbiol.* 147 (1), pp. 1–7. DOI: 10.1007/BF00492896.

Letzel, Anne-Catrin; Pidot, Sacha J.; Hertweck, Christian (2014): Genome mining for ribosomally synthesized and post-translationally modified peptides (RiPPs) in anaerobic bacteria. In *BMC genomics* 15, p. 983. DOI: 10.1186/1471-2164-15-983.

Li, Yue Zhong; Hu, Wei; Zhang, Yu Qing; Qiu, Zhi jun; Zhang, Yong; Wu, Bin Hui (2002): A simple method to isolate salt-tolerant myxobacteria from marine samples. In *Journal of microbiological methods* 50 (2), pp. 205–209.

Lippert, Karin; Galinski, ErwinA. (1992): Enzyme stabilization be ectoine-type compatible solutes. Protection against heating, freezing and drying. In *Appl Microbiol Biotechnol* 37 (1). DOI: 10.1007/BF00174204.

Louis, P.; Trper, H. G.; Galinski, E. A. (1994): Survival of Escherichia coli during drying and storage in the presence of compatible solutes. In *Appl Microbiol Biotechnol* 41 (6), pp. 684–688. DOI: 10.1007/BF00167285.

Maansson, Maria; Vynne, Nikolaj G.; Klitgaard, Andreas; Nybo, Jane L.; Melchiorsen, Jette; Nguyen, Don D. et al. (2016): An Integrated Metabolomic and Genomic Mining Workflow To Uncover the Biosynthetic Potential of Bacteria. In *mSystems* 1 (3). DOI: 10.1128/mSystems.00028-15.

Machida, Shuntaro; Bakku, Ranjith K.; Suzuki, Iwane (2017): Expression of Genes for a Flavin Adenine Dinucleotide-Binding Oxidoreductase and a Methyltransferase from Mycobacterium chlorophenolicum Is Necessary for Biosynthesis of 10-Methyl Stearic Acid from Oleic Acid in Escherichia coli. In *Frontiers in microbiology* 8, p. 2061. DOI: 10.3389/fmicb.2017.02061.

Macmillan, John B.; Molinski, Tadeusz F. (2005): Majusculoic acid, a brominated cyclopropyl fatty acid from a marine cyanobacterial mat assemblage. In *J. Nat. Prod.* 68 (4), pp. 604–606. DOI: 10.1021/np049596k.

Maloy, Aaron P.; Ford, Susan E.; Karney, Richard C.; Boettcher, Katherine J. (2007): Roseovarius crassostreae, the etiological agent of Juvenile Oyster Disease (now to be known as Roseovarius Oyster Disease) in Crassostrea virginica. In *Aquaculture* 269 (1-4), pp. 71–83. DOI: 10.1016/j.aquaculture.2007.04.008.

Marris, Emma (2006): Marine natural products. Drugs from the deep. In *Nature* 443 (7114), pp. 904–905. DOI: 10.1038/443904a.

Marshall, C. Gary; Hillson, Nathan J.; Walsh, Christopher T. (2002): Catalytic Mapping of the Vibriobactin Biosynthetic Enzyme VibF †. In *Biochemistry* 41 (1), pp. 244–250. DOI: 10.1021/bi011852u.

May, G.; Faatz, E.; Lucht, J. M.; Haardt, M.; Bolliger, M.; Bremer, E. (1989): Characterization of the osmoregulated Escherichia coli proU promoter and identification of ProV as a membrane-associated protein. In *Mol Microbiol* 3 (11), pp. 1521–1531. DOI: 10.1111/j.1365-2958.1989.tb00138.x.

McBride, M. J.; Zusman, D. R. (1989): Trehalose accumulation in vegetative cells and spores of Myxococcus xanthus. In *Journal of bacteriology* 171 (11), pp. 6383–6386.

McIntosh, John A.; Donia, Mohamed S.; Schmidt, Eric W. (2009): Ribosomal peptide natural products. Bridging the ribosomal and nonribosomal worlds. In *Natural product reports* 26 (4), pp. 537–559.

Medema, Marnix H. (2018): Computational Genomics of Specialized Metabolism. From Natural Product Discovery to Microbiome Ecology. In *mSystems* 3 (2). DOI: 10.1128/mSystems.00182-17.

Medema, Marnix H.; Kottmann, Renzo; Yilmaz, Pelin; Cummings, Matthew; Biggins, John B.; Blin, Kai et al. (2015): Minimum Information about a Biosynthetic Gene cluster. In *Nature chemical biology* 11 (9), pp. 625–631. DOI: 10.1038/nchembio.1890.

Medema, Marnix H.; Takano, Eriko; Breitling, Rainer (2013): Detecting sequence homology at the gene cluster level with MultiGeneBlast. In *Molecular biology and evolution* 30 (5), pp. 1218–1223. DOI: 10.1093/molbev/mst025.

Meier-Kolthoff, Jan P.; Klenk, Hans-Peter; Göker, Markus (2014): Taxonomic use of DNA G+C content and DNA-DNA hybridization in the genomic age. In *International journal of systematic and evolutionary microbiology* 64 (Pt 2), pp. 352–356. DOI: 10.1099/ijs.0.056994-0.

Méndez-Vilas, A. (2010): Current research, technology and education topics in applied microbiology and microbial biotechnology. Bacteria in molluscs: good and bad guys. With assistance of J.L. Romalde, J.L. Barja. Badajoz, Spain: Formatex Research Center (Microbiology book series, no. 2).

Meyer, Jeffrey D.; Nayar, Rajiv; Manning, Mark Cornell (2009): Impact of bulking agents on the stability of a lyophilized monoclonal antibody. In *European journal of pharmaceutical sciences : official journal of the European Federation for Pharmaceutical Sciences* 38 (1), pp. 29–38. DOI: 10.1016/j.ejps.2009.05.008.

Milner, J. L.; Grothe, S.; Wood, J. M. (1988): Proline porter II is activated by a hyperosmotic shift in both whole cells and membrane vesicles of Escherichia coli K12. In *The Journal of biological chemistry* 263 (29), pp. 14900–14905.

Mohimani, Hosein; Gurevich, Alexey; Mikheenko, Alla; Garg, Neha; Nothias, Louis-Felix; Ninomiya, Akihiro et al. (2017): Dereplication of peptidic natural products through database search of mass spectra. In *Nature chemical biology* 13 (1), pp. 30–37. DOI: 10.1038/nchembio.2219.

Mohr, Kathrin I.; Moradi, Azam; Glaeser, Stefanie P.; Kämpfer, Peter; Gemperlein, Katja; Nübel, Ulrich et al. (2018): Nannocystis konarekensis sp. nov., a novel

myxobacterium from an Iranian desert. In *International journal of systematic and evolutionary microbiology* 68 (3), pp. 721–729. DOI: 10.1099/ijsem.0.002569.

Mootz, Henning D.; Schwarzer, Dirk; Marahiel, Mohamed A. (2002): Ways of Assembling Complex Natural Products on Modular Nonribosomal Peptide Synthetases A list of abbreviations can be found at the end of the text. In *ChemBioChem* 3 (6), p. 490. DOI: 10.1002/1439-7633(20020603)3:6<490::AID-CBIC490>3.0.CO;2-N.

Moritz, Katharina D.; Amendt, Birgit; Witt, Elisabeth M. H. J.; Galinski, Erwin A. (2015): The hydroxyectoine gene cluster of the non-halophilic acidophile Acidiphilium cryptum. In *Extremophiles : life under extreme conditions* 19 (1), pp. 87–99. DOI: 10.1007/s00792-014-0687-0.

Mühling, Martin; Joint, Ian; Willetts, Andrew J. (2013): The biodiscovery potential of marine bacteria. An investigation of phylogeny and function. In *Microbial biotechnology* 6 (4), pp. 361–370. DOI: 10.1111/1751-7915.12054.

Muñoz-Dorado, José; Marcos-Torres, Francisco J.; García-Bravo, Elena; Moraleda-Muñoz, Aurelio; Pérez, Juana (2016): Myxobacteria. Moving, Killing, Feeding, and Surviving Together. In *Frontiers in microbiology* 7, p. 781. DOI: 10.3389/fmicb.2016.00781.

Naughton, Lynn M.; Romano, Stefano; O'Gara, Fergal; Dobson, Alan D. W. (2017): Identification of Secondary Metabolite Gene Clusters in the Pseudovibrio Genus Reveals Encouraging Biosynthetic Potential toward the Production of Novel Bioactive Compounds. In *Frontiers in microbiology* 8, p. 1494. DOI: 10.3389/fmicb.2017.01494.

Nichols, D.; Cahoon, N.; Trakhtenberg, E. M.; Pham, L.; Mehta, A.; Belanger, A. et al. (2010): Use of ichip for high-throughput in situ cultivation of "uncultivable" microbial species. In *Applied and environmental microbiology* 76 (8), pp. 2445–2450. DOI: 10.1128/AEM.01754-09.

Nikapitiya, Chamilani (2012): Bioactive secondary metabolites from marine microbes for drug discovery. In *Advances in food and nutrition research* 65, pp. 363–387. DOI: 10.1016/B978-0-12-416003-3.00024-X.

Niquille, David L.; Hansen, Douglas A.; Mori, Takahiro; Fercher, David; Kries, Hajo; Hilvert, Donald (2018): Nonribosomal biosynthesis of backbone-modified peptides. In *Nature chemistry* 10 (3), pp. 282–287. DOI: 10.1038/nchem.2891.

Overbeek, Ross; Olson, Robert; Pusch, Gordon D.; Olsen, Gary J.; Davis, James J.; Disz, Terry et al. (2014): The SEED and the Rapid Annotation of microbial genomes using Subsystems Technology (RAST). In *Nucleic acids research* 42 (Database issue), D206-14. DOI: 10.1093/nar/gkt1226.

P, Priyanka; A B, Arun; P D, Rekha (2014): Sulfated exopolysaccharide produced by Labrenzia sp. PRIM-30, characterization and prospective applications. In *International journal of biological macromolecules* 69, pp. 290–295. DOI: 10.1016/j.ijbiomac.2014.05.054.

Pan, Jian-Jung; Solbiati, Jose O.; Ramamoorthy, Gurusankar; Hillerich, Brandan S.; Seidel, Ronald D.; Cronan, John E. et al. (2015): Biosynthesis of Squalene from Farnesyl Diphosphate in Bacteria. Three Steps Catalyzed by Three Enzymes. In *ACS central science* 1 (2), pp. 77–82. DOI: 10.1021/acscentsci.5b00115.

Parks, Donovan H.; Imelfort, Michael; Skennerton, Connor T.; Hugenholtz, Philip; Tyson, Gene W. (2015): CheckM. Assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. In *Genome research* 25 (7), pp. 1043–1055. DOI: 10.1101/gr.186072.114.

Penn, Kevin; Jenkins, Caroline; Nett, Markus; Udwary, Daniel W.; Gontang, Erin A.; McGlinchey, Ryan P. et al. (2009): Genomic islands link secondary metabolism to functional adaptation in marine Actinobacteria. In *The ISME journal* 3 (10), pp. 1193–1203. DOI: 10.1038/ismej.2009.58.

Pérez-Gil, Jordi; Rodríguez-Concepción, Manuel (2013): Metabolic plasticity for isoprenoid biosynthesis in bacteria. In *The Biochemical journal* 452 (1), pp. 19–25. DOI: 10.1042/BJ20121899.

Peters, J. E.; Thate, T. E.; Craig, N. L. (2003): Definition of the Escherichia coli MC4100 Genome by Use of a DNA Array. In *Journal of bacteriology* 185 (6), pp. 2017–2021. DOI: 10.1128/JB.185.6.2017-2021.2003.

Piel, J.; Hofer, I.; Hui, D. (2004): Evidence for a Symbiosis Island Involved in Horizontal Acquisition of Pederin Biosynthetic Capabilities by the Bacterial Symbiont of Paederus fuscipes Beetles. In *Journal of bacteriology* 186 (5), pp. 1280–1286. DOI: 10.1128/JB.186.5.1280-1286.2004.

Pikal-Cleland, Katherine A.; Cleland, Jeffrey L.; Anchordoquy, Thomas J.; Carpenter, John F. (2002): Effect of glycine on pH changes and protein stability during freeze-thawing in phosphate buffer systems. In *Journal of pharmaceutical sciences* 91 (9), pp. 1969–1979. DOI: 10.1002/jps.10184.

Plaza, A; Müller, R (Eds.) (2014): Natural Products: Discourse, Diversity, and Design,: eds Osbourn A., Goss R. J. & Carter G. T.), 103-124 (Wiley-Blackwell.

Pratt, Dexter; Chen, Jing; Welker, David; Rivas, Ricardo; Pillich, Rudolf; Rynkov, Vladimir et al. (2015): NDEx, the Network Data Exchange. In *Cell systems* 1 (4), pp. 302–305. DOI: 10.1016/j.cels.2015.10.001.

Pujalte, Maria J.; Carmen Macián, M.; Arahal, David R.; Garay, Esperanza (2005): Stappia alba sp. nov., isolated from Mediterranean oysters. In *Systematic and applied microbiology* 28 (8), pp. 672–678. DOI: 10.1016/j.syapm.2005.05.010.

Ray, Lauren; Moore, Bradley S. (2016): Recent advances in the biosynthesis of unusual polyketide synthase substrates. In *Natural product reports* 33 (2), pp. 150–161. DOI: 10.1039/c5np00112a.

Reuter, Klaus; Pittelkow, Marco; Bursy, Jan; Heine, Andreas; Craan, Tobias; Bremer, Erhard (2010): Synthesis of 5-hydroxyectoine from ectoine. Crystal structure of the non-heme iron(II) and 2-oxoglutarate-dependent dioxygenase EctD. In *PloS one* 5 (5), e10647. DOI: 10.1371/journal.pone.0010647.

Richter, Michael; Rosselló-Móra, Ramon; Oliver Glöckner, Frank; Peplies, Jörg (2016): JSpeciesWS. A web server for prokaryotic species circumscription based on

pairwise genome comparison. In *Bioinformatics (Oxford, England)* 32 (6), pp. 929–931. DOI: 10.1093/bioinformatics/btv681.

Rissman, Anna I.; Mau, Bob; Biehl, Bryan S.; Darling, Aaron E.; Glasner, Jeremy D.; Perna, Nicole T. (2009): Reordering contigs of draft genomes using the Mauve aligner. In *Bioinformatics (Oxford, England)* 25 (16), pp. 2071–2073. DOI: 10.1093/bioinformatics/btp356.

Roberts, Mary F. (2005): Organic compatible solutes of halotolerant and halophilic microorganisms. In *Saline systems* 1, p. 5. DOI: 10.1186/1746-1448-1-5.

Robinson, J. A. (1991): Polyketide synthase complexes. Their structure and function in antibiotic biosynthesis. In *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 332 (1263), pp. 107–114. DOI: 10.1098/rstb.1991.0038.

Ruiz-Ruiz, Carmen; Srivastava, Girish K.; Carranza, Diana; Mata, Juan A.; Llamas, Inmaculada; Santamaría, Manuel et al. (2011): An exopolysaccharide produced by the novel halophilic bacterium Halomonas stenophila strain B100 selectively induces apoptosis in human T leukaemia cells. In *Applied microbiology and biotechnology* 89 (2), pp. 345–355. DOI: 10.1007/s00253-010-2886-7.

Rutledge, Peter J.; Challis, Gregory L. (2015): Discovery of microbial natural products by activation of silent biosynthetic gene clusters. In *Nature reviews. Microbiology* 13 (8), pp. 509–523. DOI: 10.1038/nrmicro3496.

Sadeghi, Akram; Soltani, Bahram M.; Nekouei, Mojtaba Khayam; Jouzani, Gholamreza Salehi; Mirzaei, Hossein Hadavand; Sadeghizadeh, Majid (2014): Diversity of the ectoines biosynthesis genes in the salt tolerant Streptomyces and evidence for inductive effect of ectoines on their accumulation. In *Microbiological research* 169 (9-10), pp. 699–708. DOI: 10.1016/j.micres.2014.02.005.

Satpute, Surekha K.; Banat, Ibrahim M.; Dhakephalkar, Prashant K.; Banpurkar, Arun G.; Chopade, Balu A. (2010): Biosurfactants, bioemulsifiers and exopolysaccharides from marine microorganisms. In *Biotechnology advances* 28 (4), pp. 436–450. DOI: 10.1016/j.biotechadv.2010.02.006.

Schäberle, Till F.; Goralski, Emilie; Neu, Edith; Erol, Ozlem; Hölzl, Georg; Dörmann, Peter et al. (2010): Marine myxobacteria as a source of antibiotics--comparison of physiology, polyketide-type genes and antibiotic production of three new isolates of Enhygromyxa salina. In *Marine drugs* 8 (9), pp. 2466–2479. DOI: 10.3390/md8092466.

Schleissner, Carmen; Cañedo, Librada M.; Rodríguez, Pilar; Crespo, Cristina; Zúñiga, Paz; Peñalver, Ana et al. (2017): Bacterial Production of a Pederin Analogue by a Free-Living Marine Alphaproteobacterium. In *Journal of natural products* 80 (7), pp. 2170–2173. DOI: 10.1021/acs.jnatprod.7b00408.

Schmitz, Roland P.H.; Galinski, Erwin A. (1996): Compatible solutes in luminescent bacteria of the genera Vibrio, Photobacterium and Photorhabdus (Xenorhabdus). Occurrence of ectoine, betaine and glutamate. In *FEMS microbiology letters* 142 (2-3), pp. 195–201. DOI: 10.1111/j.1574-6968.1996.tb08430.x.

Schofield, Michael M.; Sherman, David H. (2013): Meta-omic characterization of prokaryotic gene clusters for natural product biosynthesis. In *Current opinion in biotechnology* 24 (6), pp. 1151–1158. DOI: 10.1016/j.copbio.2013.05.001.

Schöner, Tim A.; Gassel, Sören; Osawa, Ayako; Tobias, Nicholas J.; Okuno, Yukari; Sakakibara, Yui et al. (2016): Aryl Polyenes, a Highly Abundant Class of Bacterial Natural Products, Are Functionally Related to Antioxidative Carotenoids. In *Chembiochem : a European journal of chemical biology* 17 (3), pp. 247–253. DOI: 10.1002/cbic.201500474.

Seemann, Torsten (2014): Prokka. Rapid prokaryotic genome annotation. In *Bioinformatics (Oxford, England)* 30 (14), pp. 2068–2069. DOI: 10.1093/bioinformatics/btu153.

Severin, J.; Wohlfarth, A.; Galinski, E. A. (1992): The predominant role of recently discovered tetrahydropyrimidines for the osmoadaptation of halophilic eubacteria. In *Journal of General Microbiology* 138 (8), pp. 1629–1638. DOI: 10.1099/00221287-138-8-1629.

Sharma, Gaurav; Narwani, Tarun; Subramanian, Srikrishna (2016): Complete Genome Sequence and Comparative Genomics of a Novel Myxobacterium Myxococcus hansupus. In *PloS one* 11 (2), e0148593. DOI: 10.1371/journal.pone.0148593.

Shen, B.; Du, L.; Sanchez, C.; Edwards, D. J.; Chen, M.; Murrell, J. M. (2001): The biosynthetic gene cluster for the anticancer drug bleomycin from Streptomyces verticillus ATCC15003 as a model for hybrid peptide-polyketide natural product biosynthesis. In *Journal of Industrial Microbiology and Biotechnology* 27 (6), pp. 378–385. DOI: 10.1038/sj.jim.7000194.

Sievers, Fabian; Wilm, Andreas; Dineen, David; Gibson, Toby J.; Karplus, Kevin; Li, Weizhong et al. (2011): Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. In *Molecular systems biology* 7, p. 539. DOI: 10.1038/msb.2011.75.

Sitachitta, N.; Gerwick, W. H. (1998): Grenadadiene and grenadamide, cyclopropyl-containing fatty acid metabolites from the marine cyanobacterium Lyngbya majuscula. In *J. Nat. Prod.* 61 (5), pp. 681–684. DOI: 10.1021/np970576a.

Sleator, Roy D.; Hill, Colin (2002): Bacterial osmoadaptation. The role of osmolytes in bacterial stress and virulence. In *FEMS Microbiol Rev* 26 (1), pp. 49–71. DOI: 10.1111/j.1574-6976.2002.tb00598.x.

Slot, Jason C. (2017): Fungal Gene Cluster Diversity and Evolution. In *Advances in genetics* 100, pp. 141–178. DOI: 10.1016/bs.adgen.2017.09.005.

Smith, D. R.; Dworkin, M. (1994): Territorial interactions between two Myxococcus Species. In *Journal of bacteriology* 176 (4), pp. 1201–1205.

Staunton, J.; Weissman, K. J. (2001): Polyketide biosynthesis. A millennium review. In *Natural product reports* 18 (4), pp. 380–416.

Stevens, H.; Brinkhoff, T.; Simon, M. (2005): Composition of free-living, aggregate-associated and sediment surface-associated bacterial communities in the German Wadden Sea. In *Aquat. Microb. Ecol.* 38, pp. 15–30. DOI: 10.3354/ame038015.

Sun, Yuwei; Feng, Zhiyang; Tomura, Tomohiko; Suzuki, Akira; Miyano, Seishi; Tsuge, Takashi et al. (2016a): Heterologous Production of the Marine Myxobacterial

Antibiotic Haliangicin and Its Unnatural Analogues Generated by Engineering of the Biochemical Pathway. In *Scientific reports* 6, p. 22091. DOI: 10.1038/srep22091.

Sun, Yuwei; Tomura, Tomohiko; Sato, Junichi; Iizuka, Takashi; Fudou, Ryosuke; Ojika, Makoto (2016b): Isolation and Biosynthetic Analysis of Haliamide, a New PKS-NRPS Hybrid Metabolite from the Marine Myxobacterium Haliangium ochraceum. In *Molecules (Basel, Switzerland)* 21 (1), p. 59. DOI: 10.3390/molecules21010059.

Suo, Z.; Tseng, C. C.; Walsh, C. T. (2001): Purification, priming, and catalytic acylation of carrier protein domains in the polyketide synthase and nonribosomal peptidyl synthetase modules of the HMWP1 subunit of yersiniabactin synthetase. In *Proceedings of the National Academy of Sciences of the United States of America* 98 (1), pp. 99–104. DOI: 10.1073/pnas.021537498.

Sydlik, Ulrich; Gallitz, Inka; Albrecht, Catrin; Abel, Josef; Krutmann, Jean; Unfried, Klaus (2009): The compatible solute ectoine protects against nanoparticle-induced neutrophilic lung inflammation. In *American journal of respiratory and critical care medicine* 180 (1), pp. 29–35. DOI: 10.1164/rccm.200812-1911OC.

Takahashi, Rodrigo Yoji Uwamori; Castilho, Nathalia Aparecida Santos; Silva, Marcus Adonai Castro da; Miotto, Maria Cecilia; Lima, André Oliveira de Souza (2017): Prospecting for Marine Bacteria for Polyhydroxyalkanoate Production on Low-Cost Substrates. In *Bioengineering (Basel, Switzerland)* 4 (3). DOI: 10.3390/bioengineering4030060.

Tan, Giin-Yu; Chen, Chia-Lung; Li, Ling; Ge, Liya; Wang, Lin; Razaad, Indah et al. (2014): Start a Research on Biopolymer Polyhydroxyalkanoate (PHA). A Review. In *Polymers* 6 (3), pp. 706–754. DOI: 10.3390/polym6030706.

Tanne, Christoph; Golovina, Elena A.; Hoekstra, Folkert A.; Meffert, Andrea; Galinski, Erwin A. (2014): Glass-forming property of hydroxyectoine is the cause of its superior function as a desiccation protectant. In *Frontiers in microbiology* 5, p. 150. DOI: 10.3389/fmicb.2014.00150.

Thibodeau, A (2005): Protecting the Skin from Environmental Stresses with an Exopolysaccharide Formulation. In *Cosmet. Toiletries* 120, pp. 81–86.

Tian, Bing; Hua, Yuejin (2010): Carotenoid biosynthesis in extremophilic Deinococcus-Thermus bacteria. In *Trends in microbiology* 18 (11), pp. 512–520. DOI: 10.1016/j.tim.2010.07.007.

Timmermans, Marshall L.; Paudel, Yagya P.; Ross, Avena C. (2017): Investigating the Biosynthesis of Natural Products from Marine Proteobacteria. A Survey of Molecules and Strategies. In *Marine drugs* 15 (8). DOI: 10.3390/md15080235.

To, Thi Mai Huong; Grandvalet, Cosette; Tourdot-Maréchal, Raphaëlle (2011): Cyclopropanation of membrane unsaturated fatty acids is not essential to the acid stress response of Lactococcus lactis subsp. cremoris. In *Applied and environmental microbiology* 77 (10), pp. 3327–3334. DOI: 10.1128/AEM.02518-10.

Tomura, Tomohiko; Nagashima, Shiori; Yamazaki, Satoshi; Iizuka, Takashi; Fudou, Ryosuke; Ojika, Makoto (2017): An Unusual Diterpene-Enhygromic Acid and Deoxyenhygrolides from a Marine Myxobacterium, Enhygromyxa sp. In *Marine drugs* 15 (4). DOI: 10.3390/md15040109.

Trautman, Eric P; Crawford, Jason M (2016): Linking Biosynthetic Gene Clusters to their Metabolites via Pathway- Targeted Molecular Networking. In *CTMC* 16 (15), pp. 1705–1716. DOI: 10.2174/1568026616666151012111046.

Velly, H.; Bouix, M.; Passot, S.; Penicaud, C.; Beinsteiner, H.; Ghorbal, S. et al. (2015): Cyclopropanation of unsaturated fatty acids and membrane rigidification improve the freeze-drying resistance of Lactococcus lactis subsp. lactis TOMSC161. In *Applied microbiology and biotechnology* 99 (2), pp. 907–918. DOI: 10.1007/s00253-014-6152-2.

Venter, J. Craig; Remington, Karin; Heidelberg, John F.; Halpern, Aaron L.; Rusch, Doug; Eisen, Jonathan A. et al. (2004): Environmental genome shotgun sequencing of the Sargasso Sea. In *Science (New York, N.Y.)* 304 (5667), pp. 66–74. DOI: 10.1126/science.1093857.

Waditee, Rungaroon; Tanaka, Yoshito; Aoki, Kenji; Hibino, Takashi; Jikuya, Hiroshi; Takano, Jun et al. (2003): Isolation and functional characterization of N-methyltransferases that catalyze betaine synthesis from glycine in a halotolerant

photosynthetic organism Aphanothece halophytica. In *The Journal of biological chemistry* 278 (7), pp. 4932–4942. DOI: 10.1074/jbc.M210970200.

Wang, Mingxun; Carver, Jeremy J.; Phelan, Vanessa V.; Sanchez, Laura M.; Garg, Neha; Peng, Yao et al. (2016): Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. In *Nature biotechnology* 34 (8), pp. 828–837. DOI: 10.1038/nbt.3597.

Weber, Tilmann; Blin, Kai; Duddela, Srikanth; Krug, Daniel; Kim, Hyun Uk; Bruccoleri, Robert et al. (2015): antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. In *Nucleic acids research* 43 (W1), W237-43. DOI: 10.1093/nar/gkv437.

Wei, Jeremy H.; Yin, Xinchi; Welander, Paula V. (2016): Sterol Synthesis in Diverse Bacteria. In *Frontiers in microbiology* 7, p. 990. DOI: 10.3389/fmicb.2016.00990.

Weissman, Kira J. (2009): Introduction to polyketide biosynthesis. In *Methods in enzymology* 459, pp. 3–16. DOI: 10.1016/S0076-6879(09)04601-1.

Widderich, Nils; Czech, Laura; Elling, Felix J.; Könneke, Martin; Stöveken, Nadine; Pittelkow, Marco et al. (2016): Strangers in the archaeal world. Osmostress-responsive biosynthesis of ectoine and hydroxyectoine by the marine thaumarchaeon Nitrosopumilus maritimus. In *Environmental microbiology* 18 (4), pp. 1227–1248. DOI: 10.1111/1462-2920.13156.

Williams, Kelly P.; Kelly, Donovan P. (2013): Proposal for a new class within the phylum Proteobacteria, Acidithiobacillia classis nov., with the type order Acidithiobacillales, and emended description of the class Gammaproteobacteria. In *International journal of systematic and evolutionary microbiology* 63 (Pt 8), pp. 2901–2906. DOI: 10.1099/ijs.0.049270-0.

Williams, Kelly P.; Sobral, Bruno W.; Dickerman, Allan W. (2007): A robust species tree for the alphaproteobacteria. In *Journal of bacteriology* 189 (13), pp. 4578–4586. DOI: 10.1128/JB.00269-07.

Wisecaver, Jennifer H.; Borowsky, Alexander T.; Tzin, Vered; Jander, Georg; Kliebenstein, Daniel J.; Rokas, Antonis (2017): A Global Coexpression Network

Approach for Connecting Genes to Specialized Metabolic Pathways in Plants. In *The Plant cell* 29 (5), pp. 944–959. DOI: 10.1105/tpc.17.00009.

Wood, Janet M. (2006): Osmosensing by bacteria. In *Science's STKE : signal transduction knowledge environment* 2006 (357), pe43. DOI: 10.1126/stke.3572006pe43.

Xu, Wei; Mukherjee, Sumit; Ning, Yu; Hsu, Fong-Fu; Zhang, Kai (2018): Cyclopropane fatty acid synthesis affects cell shape and acid resistance in Leishmania mexicana. In *International journal for parasitology* 48 (3-4), pp. 245–256. DOI: 10.1016/j.ijpara.2017.09.006.

Yamada, Yuuki; Kuzuyama, Tomohisa; Komatsu, Mamoru; Shin-Ya, Kazuo; Omura, Satoshi; Cane, David E.; Ikeda, Haruo (2015): Terpene synthases are widely distributed in bacteria. In *Proceedings of the National Academy of Sciences of the United States of America* 112 (3), pp. 857–862. DOI: 10.1073/pnas.1422108112.

Yang, Jane Y.; Sanchez, Laura M.; Rath, Christopher M.; Liu, Xueting; Boudreau, Paul D.; Bruns, Nicole et al. (2013): Molecular networking as a dereplication strategy. In *Journal of natural products* 76 (9), pp. 1686–1699. DOI: 10.1021/np400413s.

Yeong M. (2016): BiG-SCAPE: exploring biosynthetic diversity through gene cluster similarity networks. Msc. Thesis by BSc. M. Yeong Supervised by dr. MH Medema on the Bioinformatics subdivision of the Wageningen UR.

Zaburannyi, Nestor; Bunk, Boyke; Maier, Josef; Overmann, Jörg; Müller, Rolf (2016): Genome Analysis of the Fruiting Body-Forming Myxobacterium Chondromyces crocatus Reveals High Potential for Natural Product Biosynthesis. In *Applied and environmental microbiology* 82 (6), pp. 1945–1957. DOI: 10.1128/AEM.03011-15.

Zhang, Yu-Qing; Li, Yue-Zhong; Wang, Bing; Wu, Zhi-Hong; Zhang, Cui-Ying; Gong, Xun et al. (2005): Characteristics and living patterns of marine myxobacterial isolates. In *Applied and environmental microbiology* 71 (6), pp. 3331–3336. DOI: 10.1128/AEM.71.6.3331-3336.2005.

Ziemert, Nadine; Lechner, Anna; Wietz, Matthias; Millán-Aguiñaga, Natalie; Chavarria, Krystle L.; Jensen, Paul Robert (2014): Diversity and evolution of secondary

metabolism in the marine actinomycete genus Salinispora. In *Proceedings of the National Academy of Sciences of the United States of America* 111 (12), E1130-9. DOI: 10.1073/pnas.1324161111.

Zinger, Lucie; Amaral-Zettler, Linda A.; Fuhrman, Jed A.; Horner-Devine, M. Claire; Huse, Susan M.; Welch, David B. Mark et al. (2011): Global patterns of bacterial beta-diversity in seafloor and seawater ecosystems. In *PloS one* 6 (9), e24570. DOI: 10.1371/journal.pone.0024570.

9 APPENDIX

Table 9.1 ^{1}H- and ^{13}C-NMR assignments (300 MHz, methanol-d4) for enhygrolide A

	enhygrolide A	
position	δ_C , mult.	δ_H (*J* in Hz)
1	170.6, C	-
2	128.3, C	-
3	153.3, C	-
4	149.7, C	-
5	109.7, CH	6.38, s
6	134.7, C	-
7	131.2, CH	7,82, d (7.5)
8	129.6, CH	7., t (7.5)
9	129.3, CH	7.33, t (7.5)
10	129.6, CH	7.41, t (7.5)
11	131.2, CH	7.82, d (7.5)
12	27.8, CH_2	3.67, s
13	130.4 C	-
14	130.6, CH	7.11, d (8.2)
15	116.4, CH	7.82, d (7.5)
16	156.9, C	-
17	116.2, CH	6.75, d (8.2)
18	130.4, CH	7.11, d (8.2)
19	34.2, CH_2	2.58, d (7.5)
20	30.7, CH	1.97, m
21	22.9, CH_3	0.97, d (6.6)
22	22.9, CH_3	0.97, d (6.6)

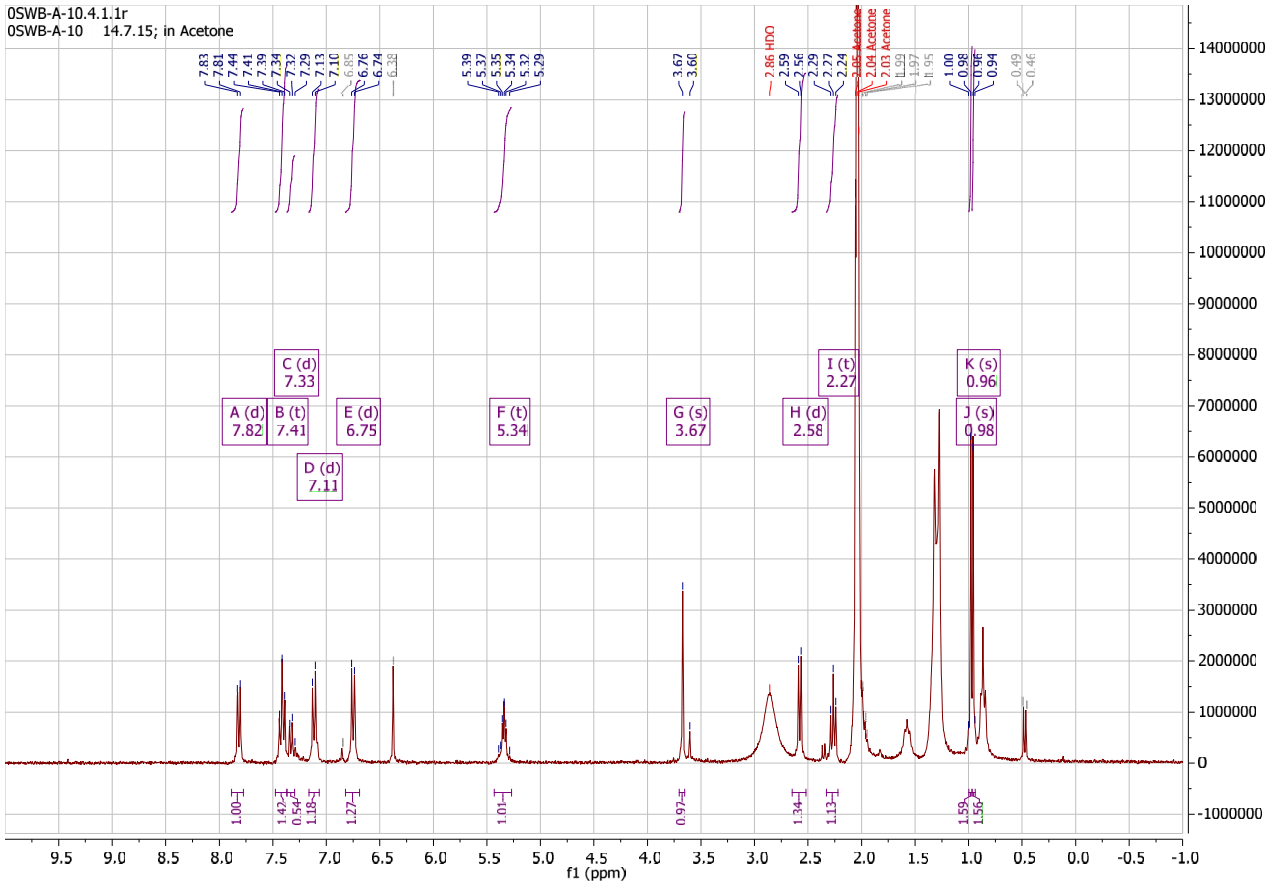


Figure 9.1 ^{1}H NMR spectrum of enhygrolide A in Acetone-d_6 with traces of enhygrolide B with which enhygrolide A is in a dynamic equilibrium.

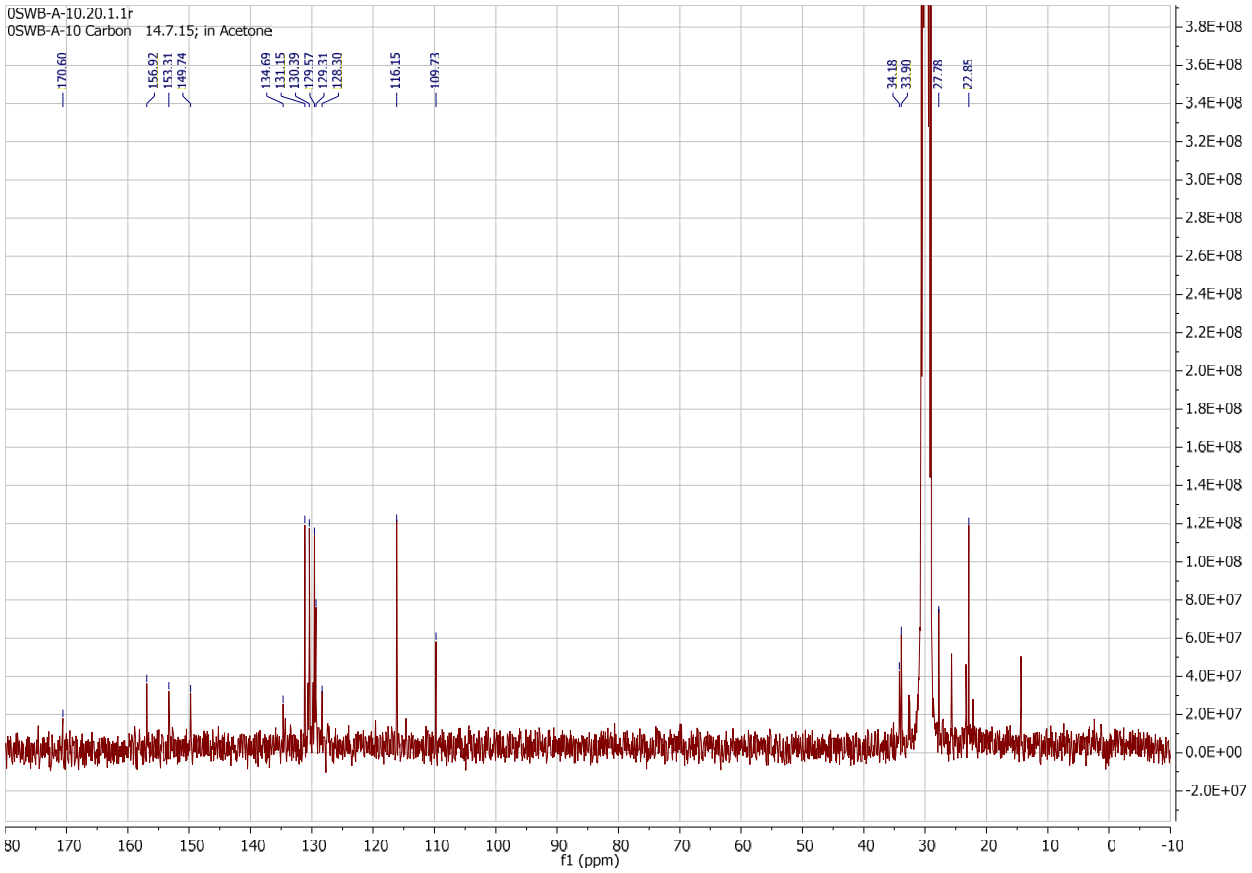


Figure 9.2 ^{13}C NMR spectrum of enhygrolide A in Acetone-*d6* with traces of the stereoisomer enhygrolide B with which enhygrolide A is in a dynamic equilibrium.

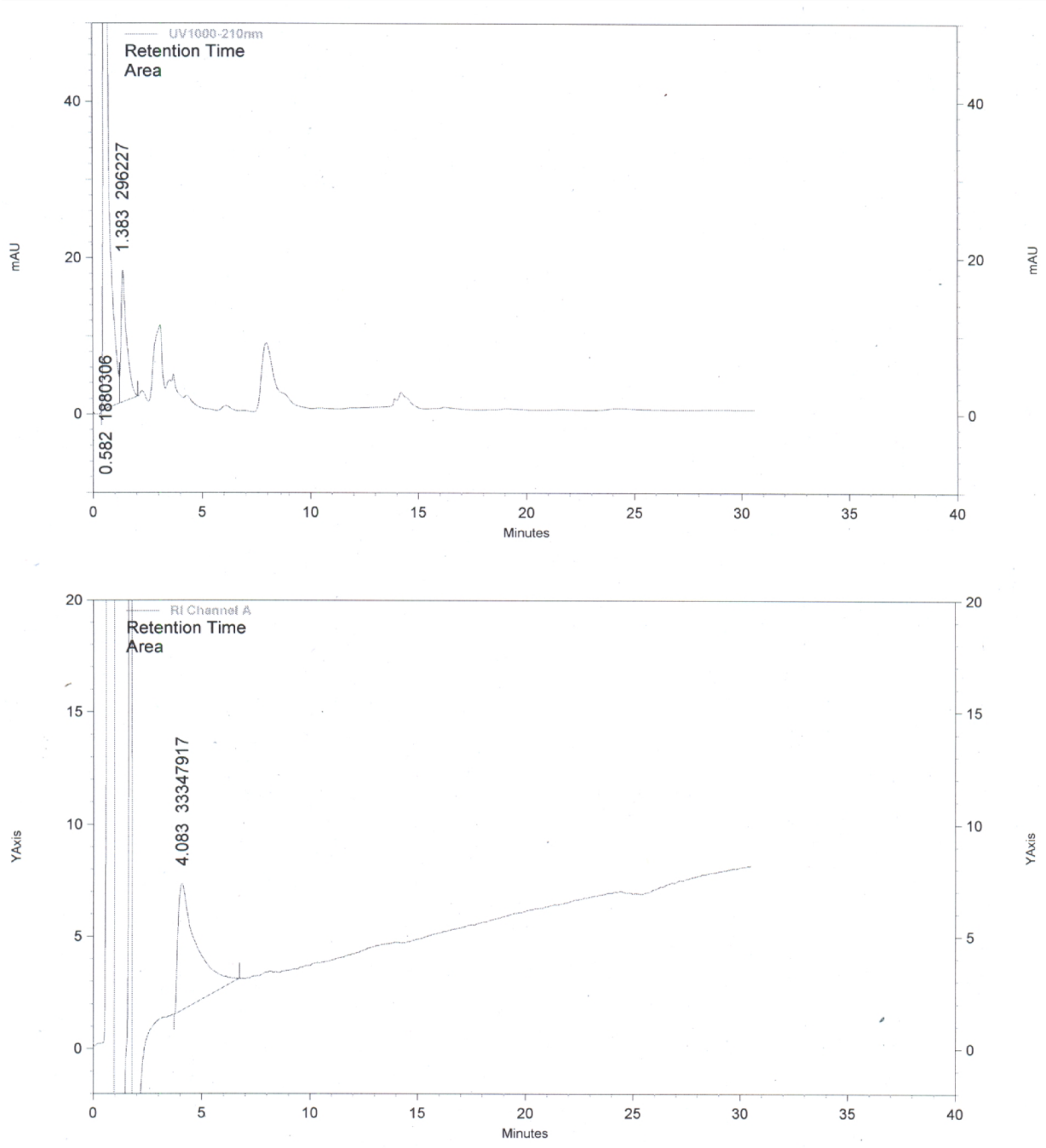


Figure 9.3 HPLC analysis of compatible solute content of *P. pacifica* DSM 14875 at 1% NaCl concentration

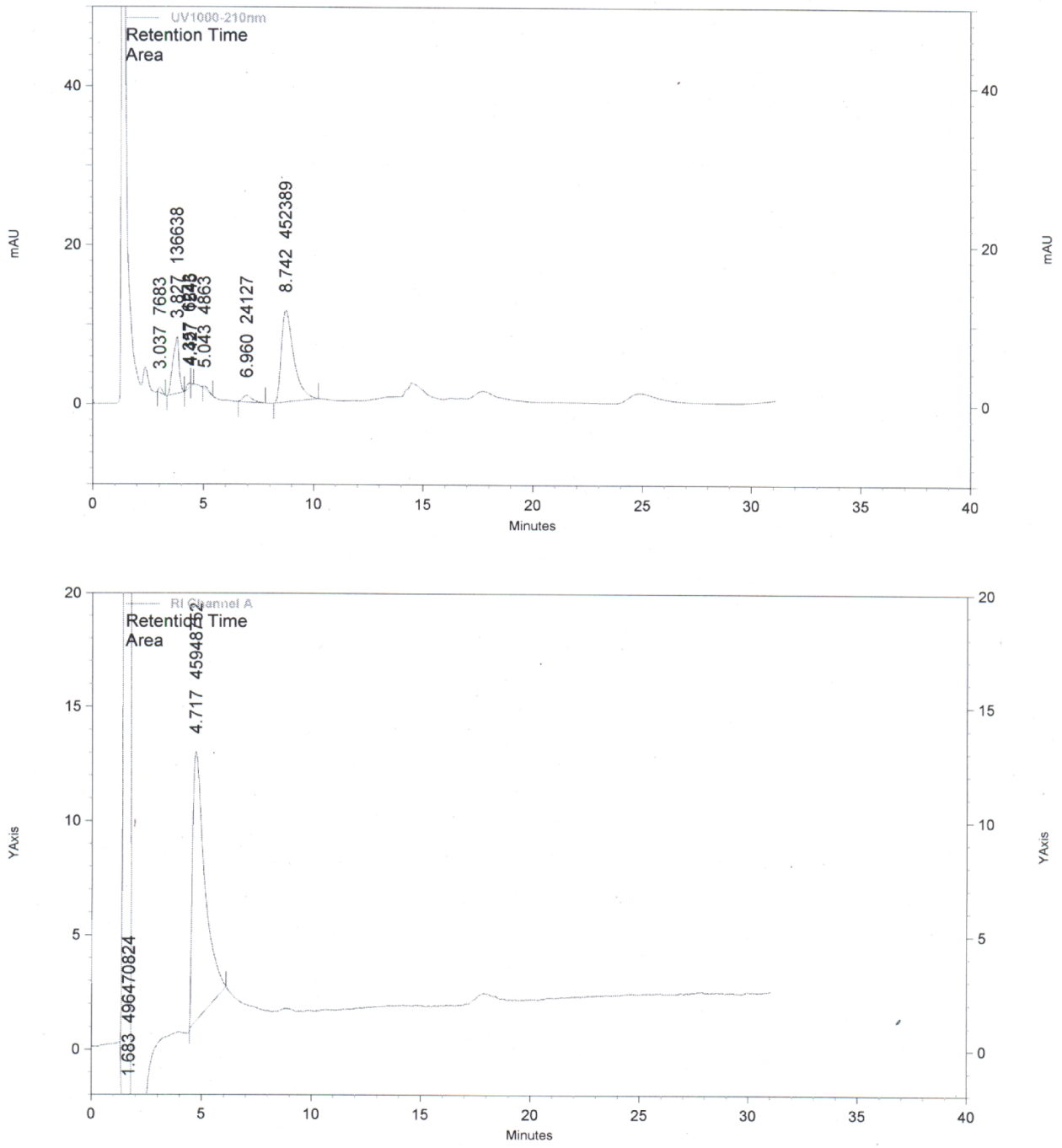


Figure 9.4 HPLC analysis of compatible solute content of *P. pacifica* DSM 14875 at 3.5% NaCl concentration

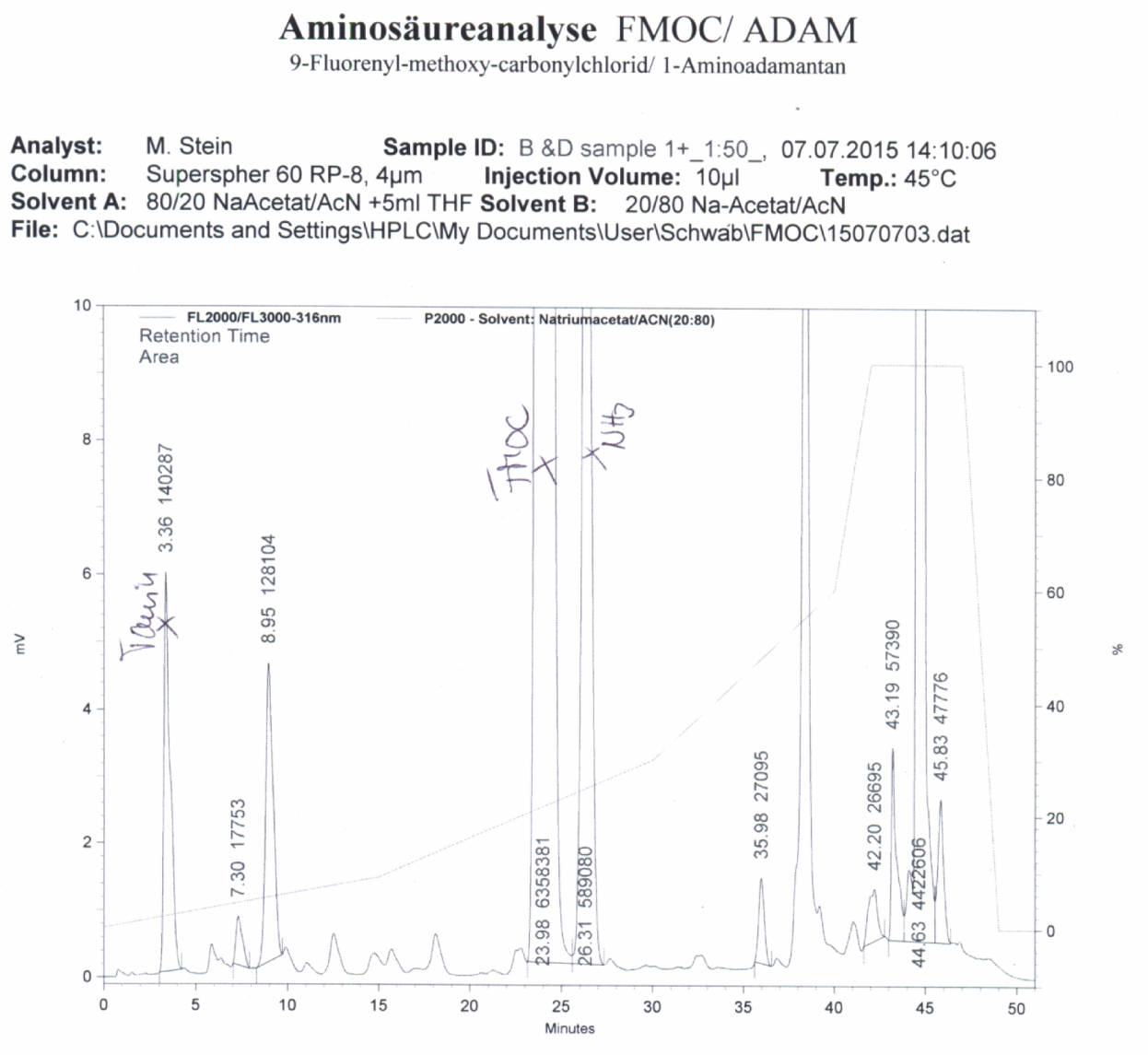


Figure 9.5 HPLC analysis of free amino acids content of *P. pacifica* DSM 14875 at 1% NaCl concentration

Aminosäureanalyse FMOC/ ADAM

9-Fluorenyl-methoxy-carbonylchlorid/ 1-Aminoadamantan

Analyst: M. Stein **Sample ID:** B &D sample 3,5_1:50_, 07.07.2015 15:18:55
Column: Superspher 60 RP-8, 4µm **Injection Volume:** 10µl **Temp.:** 45°C
Solvent A: 80/20 NaAcetat/AcN +5ml THF **Solvent B:** 20/80 Na-Acetat/AcN
File: C:\Documents and Settings\HPLC\My Documents\User\Schwab\FMOC\15070704.dat

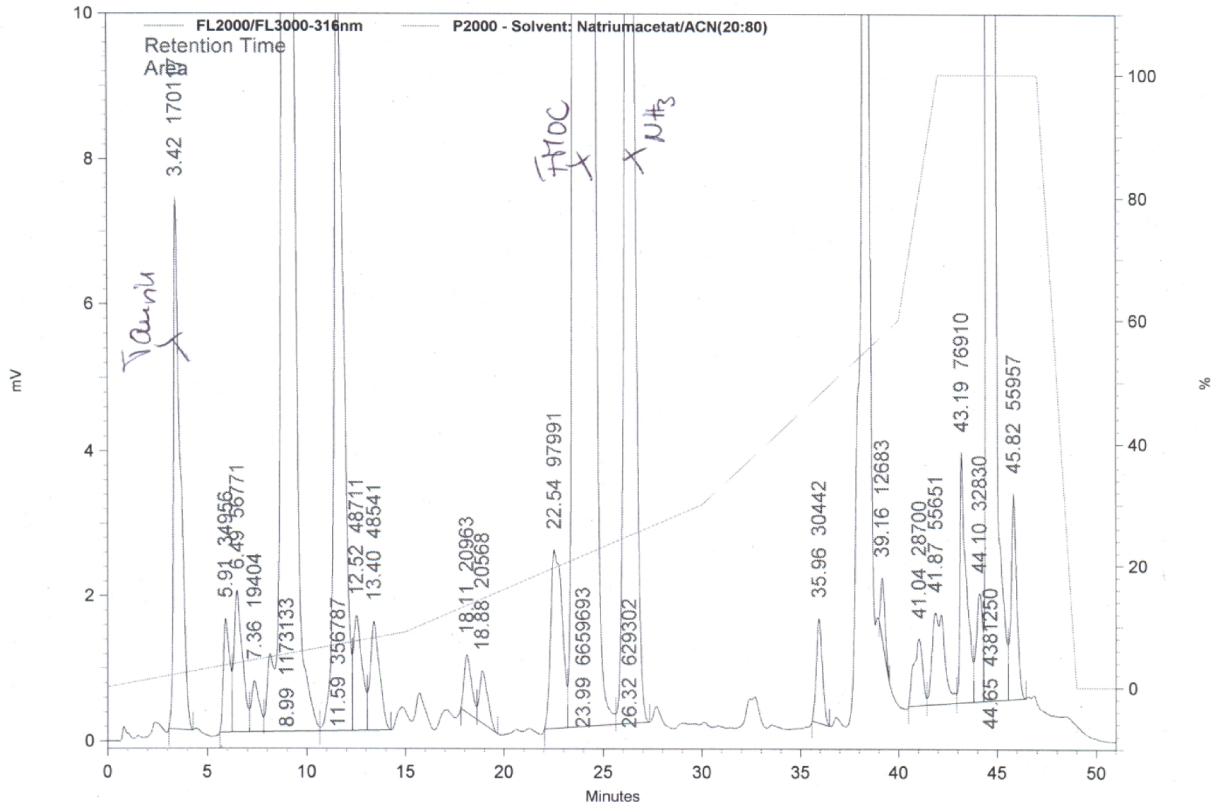


Figure 9.6 HPLC analysis of free amino acids content of *P. pacifica* DSM 14875 at 3.5% NaCl concentration

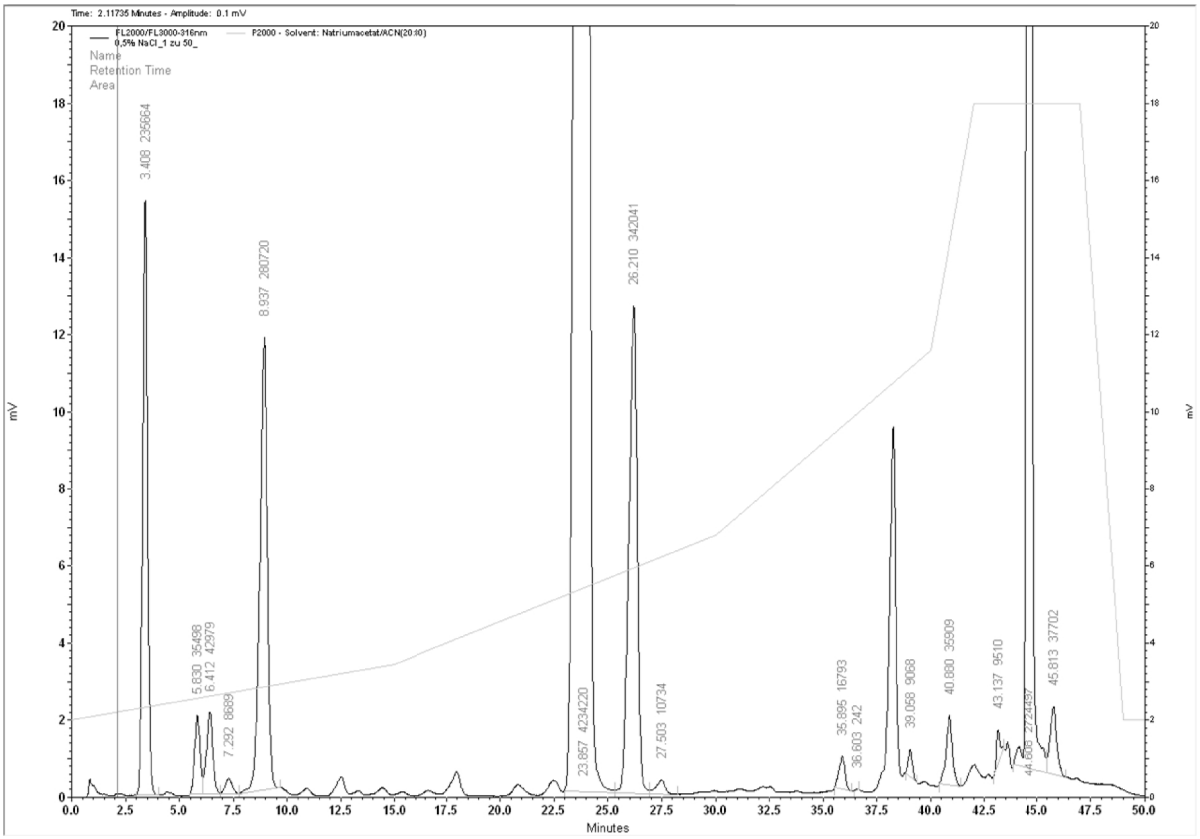


Figure 9.7 HPLC analysis of free amino acids content of *E. salina* SWB007 at 0.5% NaCl concentration

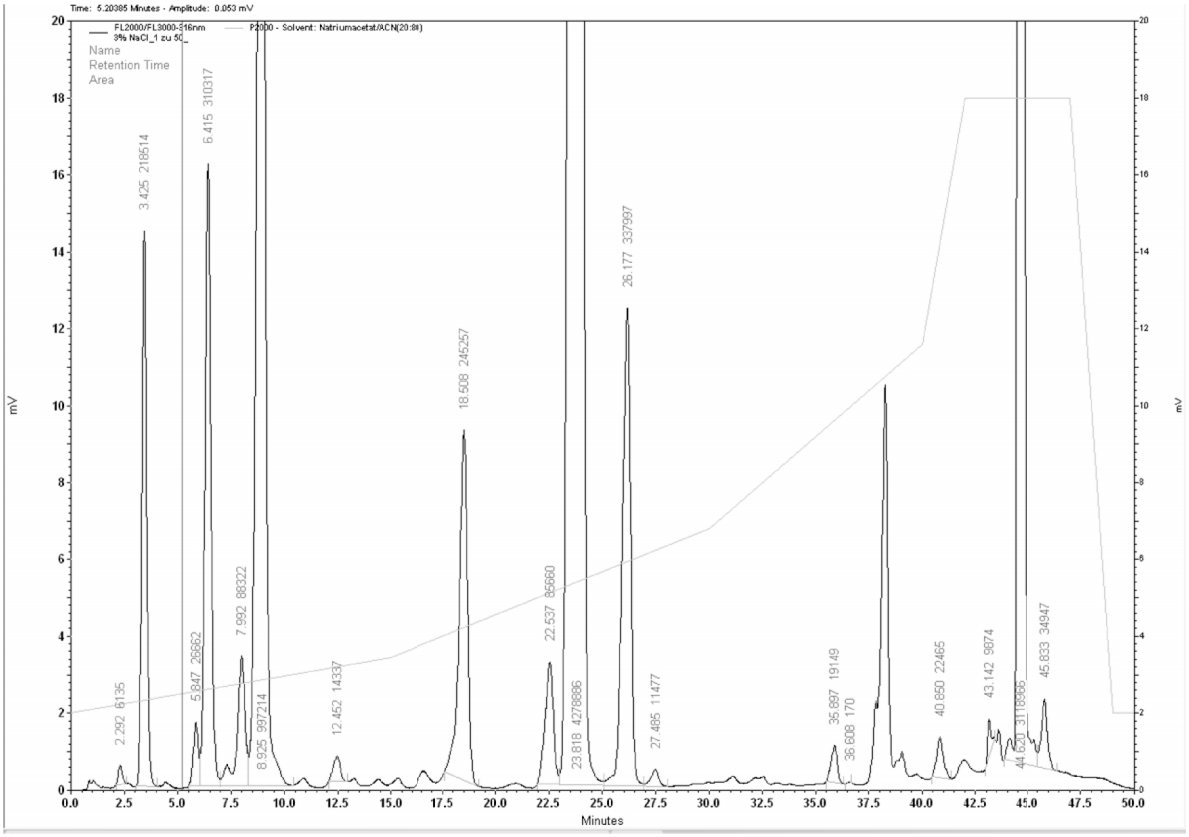


Figure 9.8 HPLC analysis of free amino acids content of *E. salina* SWB007 at 3% NaCl concentration

Table 9.2 Growth of *Enhygromyxa salina* SWB007 under different salt concentrations

NaCl %	0					0,5					1				
	Culture 1	Culture 2	Culture3	MV	SD	Culture 1	Culture 2	Culture3	MV	SD	Culture 1	Culture 2	Culture3	MV	SD
Day 1	1,04	1,1	1,01	1,05	0,05	1,17	0,99	0,86	1,01	0,16	1,1	0,75	1,06	0,97	0,19
Day 2	0,875	0,925	0,845	0,88	0,04	0,765	0,975	0,735	0,83	0,13	1,105	0,825	0,975	0,97	0,14
Day 3	0,815	0,79	0,84	0,82	0,03	0,56	0,615	0,285	0,49	0,18	0,895	0,355	0,705	0,65	0,27
Day 4	0,556	0,676	0,658	0,63	0,06	0,114	0,132	0,098	0,11	0,02	0,312	0,132	0,23	0,22	0,09
Day 5	0,62	0,678	0,706	0,67	0,04	0,046	0,086	0,042	0,06	0,02	0,338	0,044	0,192	0,19	0,15
Day 6	0,656	0,71	0,574	0,65	0,07	0,029	0,033	0,021	0,03	0,01	0,117	0,017	0,033	0,06	0,05
Day 7	0,626	0,74	0,696	0,69	0,06	0	0	0,011	0,00	0,01	0,055	0,006	0,016	0,03	0,03
Day 8	0,9	0,586	0,63	0,71	0,17										
Day 9	0,948	0,8	0,67	0,81	0,14										
Day 10	0,976	0,852	0,794	0,87	0,09										
NaCl %	**1,5**					**2**					**2,5**				
	Culture 1	Culture 2	Culture3	MV	SD	Culture 1	Culture 2	Culture3	MV	SD	Culture 1	Culture 2	Culture3	MV	SD
Day 1	0,75	0,86	1,07	0,89	0,16	0,7	0,79	0,69	0,73	0,06	1	0,81	1,06	0,96	0,13
Day 2	0,75	0,87	1,03	0,88	0,14	0,725	0,73	0,735	0,73	0,01	0,94	0,905	1,08	0,98	0,09
Day 3	0,315	0,575	0,89	0,59	0,29	0,315	0,245	0,27	0,28	0,04	0,84	0,48	0,89	0,74	0,22
Day 4	0,128	0,178	0,674	0,33	0,30	0,068	0,092	0,09	0,08	0,01	0,6	0,12	0,77	0,50	0,34
Day 5	0,036	0,19	0,41	0,21	0,19	0,118	0,13	0,072	0,11	0,03	0,23	0,07	0,63	0,31	0,29
Day 6	0,01	0,088	0,144	0,08	0,07	0,028	0,04	0,031	0,03	0,01	0,12	0,103	0,461	0,23	0,20
Day 7	0,005	0,037	0,13	0,06	0,06	0,022	0,014	0,003	0,01	0,01	0,196	0,071	0,245	0,17	0,09
Day 8											0,122	0,031	0,125	0,09	0,05
Day 9											0,055	0,011	0,135	0,07	0,06
Day 10															

NaCl %	3					3,5					4				
	Culture 1	Culture 2	Culture3	MV	SD	Culture 1	Culture 2	Culture3	MV	SD	Culture 1	Culture 2	Culture3	MV	SD
Day 1	1,07	1,11	1,05	1,08	0,03	1,07	1,07	1,12	1,09	0,03	1,11	1,32	0,93	1,12	0,20
Day 2	1,01	1	0,965	0,99	0,02	1,06	1,015	1,02	1,03	0,02	1,11	1,085	1,02	1,07	0,05
Day 3	0,885	0,94	0,85	0,89	0,05	1	0,99	1,045	1,01	0,03	1,065	1,02	1,03	1,04	0,02
Day 4	0,736	0,816	0,718	0,76	0,05	1,035	1,025	1,08	1,05	0,03	1,14	1,07	1,07	1,09	0,04
Day 5	0,67	0,726	0,606	0,67	0,06	0,995	0,985	1,04	1,01	0,03	1,1	1,1	1,095	1,10	0,00
Day 6	0,424	0,62	0,354	0,47	0,14	0,915	0,935	0,95	0,93	0,02	1,045	1,07	1,03	1,05	0,02
Day 7	0,17	0,496	0,14	0,27	0,20	0,865	0,905	0,97	0,91	0,05	1,115	1,11	1,055	1,09	0,03
Day 8	0,106	0,274	0,121	0,17	0,09	0,925	0,9	1,02	0,95	0,06	1,125	1,115	1,07	1,10	0,03
Day 9	0,112	0,229	0,187	0,18	0,06	0,925	0,87	0,975	0,92	0,05					
Day 10															

NaCl %	4,5					5				
	Culture 1	Culture 2	Culture3	MV	SD	Culture 1	Culture 2	Culture3	MV	SD
Day 1	1,13	1,16	1,19	1,16	0,03	1,11	1,22	1,1	1,14	0,07
Day 2	1,1	1,115	1,09	1,10	0,01	1,075	1,105	1,13	1,10	0,03
Day 3	1,1	1,1	1,135	1,11	0,02	1,055	1,09	1,115	1,09	0,03
Day 4	1,125	1,09	1,17	1,13	0,04	1,085	1,13	1,145	1,12	0,03
Day 5	1,11	1,115	1,125	1,12	0,01	1,075	1,195	1,15	1,14	0,06
Day 6	1,065	1,005	0,885	0,99	0,09	1,025	1,075	1,075	1,06	0,03
Day 7	1,135	1,12	1,16	1,14	0,02	1,12	1,175	1,175	1,16	0,03
Day 8	1,145	1,065	1,215	1,14	0,08	1,12	1,255	1,265	1,21	0,08
Day 9										
Day 10										

Values are the OD600 measurements which were blanked with artificial sea water. MV: mean value; SD: standard deviation

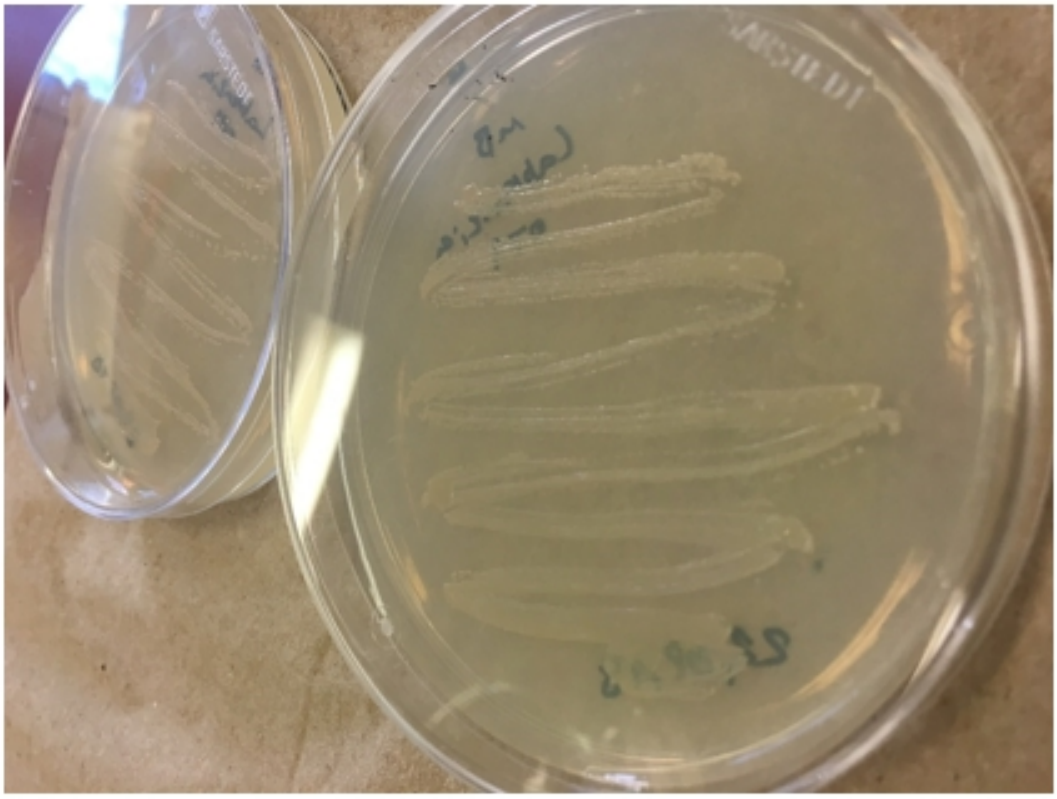


Figure 9.9 Axenic culture of *Labrenzia* sp. 011 on marine broth-based medium.