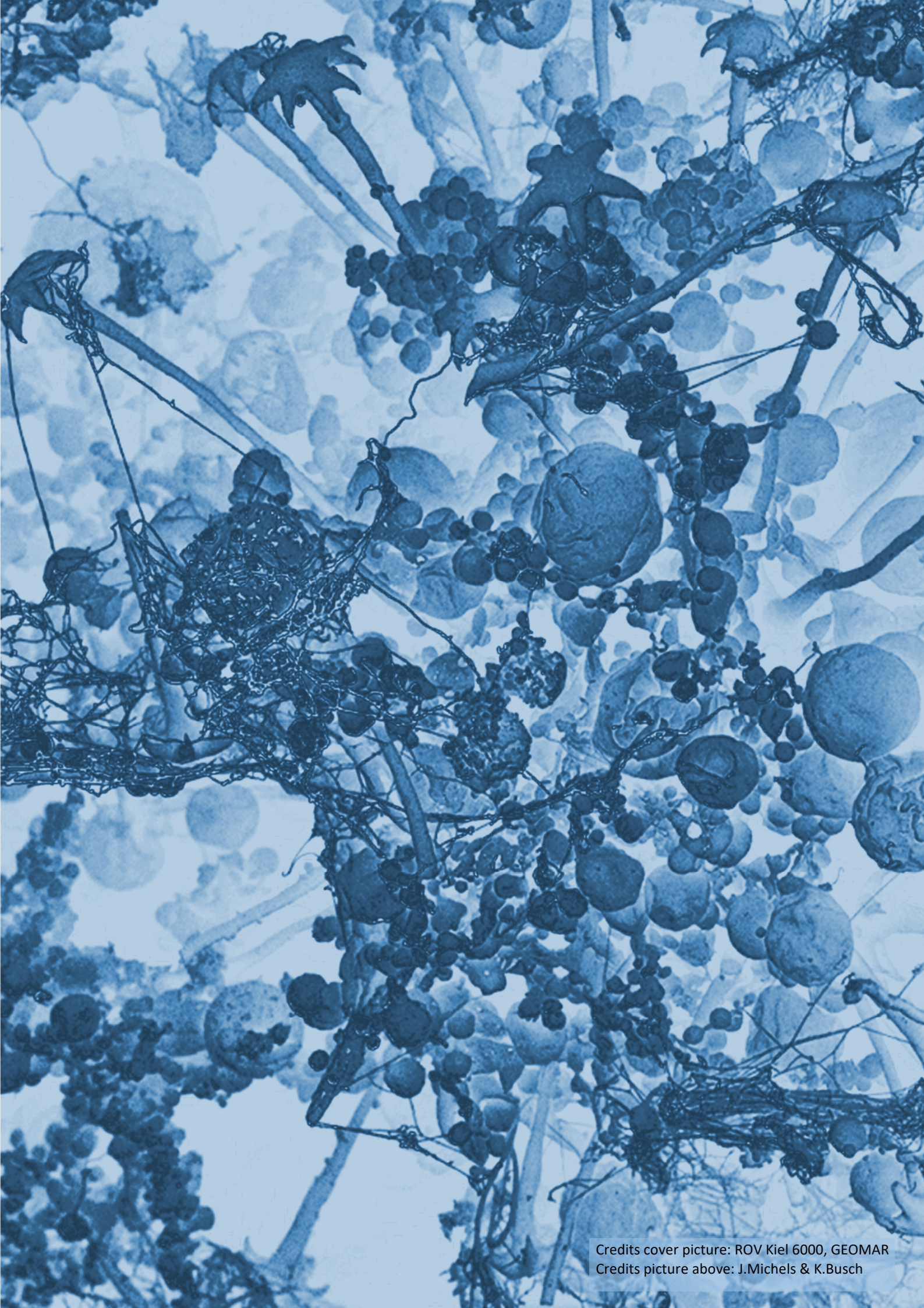


Biodiversity of deep-sea sponge microbiomes
in the ecosystem context

Dissertation

Kathrin Busch





Credits cover picture: ROV Kiel 6000, GEOMAR
Credits picture above: J.Michels & K.Busch

Biodiversity of deep-sea sponge microbiomes
in the ecosystem context

Dissertation

zur Erlangung des Doktorgrades
der Mathematisch-Naturwissenschaftlichen Fakultät
der Christian-Albrechts-Universität zu Kiel

vorgelegt von
Kathrin Busch
Kiel, 2020

Erste Gutachterin: Prof. Ute Hentschel Humeida

Zweiter Gutachter: Prof. Arne Biastoch

Tag der mündlichen Prüfung: 18.12.2020

Table of contents

1. Summary.....	
1.1 English.....	1
1.1 German.....	3
2. Introduction and thesis outline.....	5
3. Methods.....	23
4. Chapters.....	
List of included manuscripts.....	39
4.1 Chapter 1.....	44
4.2 Chapter 2.....	64
4.3 Chapter 3.....	80
4.4 Chapter 4.....	94
4.5 Chapter 5.....	117
4.6 Chapter 6.....	143
4.7 Chapter 7.....	164
4.8 Chapter 8.....	177
5. Discussion.....	186
6. Conclusions.....	200
7. Acknowledgements.....	203
8. References.....	204
9. Appendix.....	214
9.1 Bioinformatic core amplicon analysis.....	214
9.2 Supplementary Material Chapter 7.....	270
10. Declaration.....	334

1

2

3

4

5

6

7

8

9

10

Summary

1.1

The deep-sea is the largest habitat on Earth, but its biodiversity and ecosystem dynamics are still underexplored. Deep-sea sponge grounds (syn. aggregations, gardens) are sponge-dominated ecosystems that are found throughout the world's oceans. They are considered vulnerable marine ecosystems (VMEs) and warrant protection against human intervention. Deep-sea sponge grounds are considered hotspots of diversity and function in the deep ocean. While a significant body of information has been accrued on shallow-water sponges, our understanding of deep-sea sponges and their associated microbiomes at the beginning of this PhD thesis was still very limited. This PhD thesis therefore aims to provide a first comprehensive overview on the diversity, evolution, biogeography, and ecology of deep-sea sponge microbiomes. The overarching aim was to assess whether the concepts obtained in shallow-water sponge microbiology would also hold in the deep-sea. In addition, novel themes such as biogeochemistry, physical oceanography, and trait-based approaches were integrated and further expand the existing theoretical framework in sponge microbiology.

Sampling was conducted during 20 deep-sea expeditions, largely to sponge grounds of the North Atlantic in the context of the EU project "*SponGES: Deep-sea sponge grounds ecosystems of the North Atlantic - an integrated approach towards their preservation and sustainable exploitation*". In total 1077 sponge-associated microbiomes were sampled along with 355 seawater microbiomes and 114 sediment microbiomes from 52 sponge ground locations. Microbial diversity was assessed by 16S rRNA gene amplicon sequencing and host taxonomy was determined by a combination of taxonomic and molecular markers. To this end, a state-of-the-art high-throughput 16S amplicon pipeline was established and corresponding metadata workflows were developed. The resulting data were analysed by six specific case studies (of which all were published) and one overarching meta-analysis (manuscript in preparation). The microbial community composition of deep-sea sponges was explored across different scales, from the ecosystem- and biogeography-level, to individual sponge species and to the microbial taxon (*Amplicon Single nucleotide Variant*; ASV) level. By exploring sponge microbiomes on different levels of integration and by using a nested sampling design, I was able to identify overarching factors, that drive microbiome composition in a statistically proven manner. The main identified environmental drivers of microbial community variability were temperature, salinity, nutrients/oxygen, and depth. It is noteworthy, that these parameters were identified from a total set of 24 environmental parameters. Furthermore, sponge phylogeny, taxonomy, and morphology were found to be related with the microbial community composition. Interestingly, microbial diversity can be predicted based on sponge morphology, which offers exciting opportunities for future studies in respect to imaging or trait-based approaches.

My conclusions on the microbiome composition of deep-sea sponge microbiomes are that each deep-sea sponge harbours an individual set of microbes and a large pool of hidden diversity. Furthermore, deep-sea sponge microbiomes are globally not well connected and rather display heterogeneity on local scales. Interestingly, a deep-sea specific sponge microbiome was discovered. Overall, the results of my thesis suggest a strong nestedness of deep-sea sponge microbiomes within their ecological context.

In the context of this PhD thesis, I established a baseline of deep-sea sponge-associated microbiomes, discovered a large extent of novel diversity and described patterns of specificity, stability and variability. I further identified the environmental and host-related drivers of

sponge microbiome composition. From a methodological point, I have designed and developed a software tool (termed *SVAmPEX*) that allows the archiving and user-friendly accessibility of deep-sea sponge microbiome baseline data. Since microbiome composition is directly related to sponge health, reference baselines are valuable to monitor the integrity and resilience of deep-sea sponges. The collective information gathered in this PhD thesis provides the scientific basis to improve conservation and management strategies of the vulnerable deep-sea sponge ground ecosystems in the long run.

Zusammenfassung

1.2

Die Tiefsee ist der größte Lebensraum auf der Erde, dennoch sind die dort vorkommende Artenvielfalt sowie Ökosystemprozesse weitgehend unbekannt. Tiefsee-Schwammgründe (syn. Schwammaggregationen, Schwammgärten) sind weltweit vorkommende Schwamm-dominierte Ökosysteme. Es handelt sich um sogenannte vulnerable marine Ökosysteme, welche Schutz vor menschlichen Eingriffen bedürfen. Tiefseeschwammgründe sind Diversitäts- und Funktions- Hotspots in den Tiefen des Ozeans. Während bereits eine beachtliche Fülle an Informationen über Flachwasserschwämme angesammelt worden ist, war unser Verständnis von Tiefseeschwämmen und ihren assoziierten Mikroben sehr begrenzt als diese Doktorarbeit begonnen wurde. Meine Dissertation verschafft daher einen umfassenden ersten Überblick über die Diversität, Evolution, Biogeographie, und Ökologie von Tiefsee-Schwamm-Mikrobiomen. Ein übergeordnetes Ziel war es einen Überblick zu erhalten ob die theoretischen Konzepte der Flachwasser-Schwamm-Mikrobiologie auch in der Tiefsee zutreffen. Zusätzlich wurden neue Aspekte aus den Themenbereichen der Biogeochemie und physikalischen Ozeanographie eingebracht und Merkmals-basierte Denkansätze verfolgt. Dies wurde mit dem Ziel getan, den existierenden theoretischen Rahmen der Schwamm-Mikrobiologie zu erweitern.

Probenahmen wurden während 20 Tiefsee-Expeditionen durchgeführt. Hauptsächlich an Schwammgründen im Nordatlantik im Kontext des EU-Projekts *“SponGES: Tiefsee-Schwammgrund-Ökosysteme im Nordatlantik – ein ganzheitlicher Ansatz zum Schutz und einer nachhaltigen Nutzung“*. Insgesamt wurden 1077 Schwamm-assoziierte Mikrobiome, sowie 355 Seewasser-Mikrobiome und 114 Sediment-Mikrobiome an 52 Standorten beprobt. Die mikrobielle Vielfalt wurde mit Hilfe der 16S rRNA Gen Sequenzierung entschlüsselt, wohingegen die Schwamntaxonomie durch eine Kombination von taxonomischen und molekularen Markern bestimmt wurde. In diesem Zusammenhang wurde eine Hochdurchsatz-16S Amplikon Pipeline etabliert und dazugehörige Metadaten-Arbeitsabläufe entwickelt. Die resultierenden Daten wurden in sechs spezifischen Fallstudien (welche alle bereits veröffentlicht wurden), sowie einer übergeordneten Meta-Analyse (hier ist das Manuskript in Arbeit) analysiert. Die Zusammensetzung der mikrobiellen Gemeinschaft in Tiefsee-Schwämmen wurde auf verschiedenen Skalen untersucht, von der Ökosystem- und Biogeografie-Ebene, über individuelle Schwammarten, bis hin zur mikrobiellen Taxon-Ebene (*Amplicon Single nucleotide Variant; ASV*). Durch die Untersuchung von Schwammmikrobiomen auf verschiedenen Integriationsebenen und durch die Verwendung eines verschachtelten Beprobungsdesigns konnte ich mit einem statistischen Ansatz übergeordnete Faktoren identifizieren welche die mikrobielle Gemeinschaftszusammensetzung steuern. Als Umweltfaktoren, welche die mikrobielle Gemeinschaftszusammensetzung hauptsächlich steuern, wurden Temperatur, Salzgehalt, Nährstoffe sowie Sauerstoff, und die Beprobungstiefe identifiziert. Nennenswerterweise wurden diese Parameter aus einem Gesamtset von 24 Umweltfaktoren bestimmt. Zusätzlich stand die Schwammphylogenie, -taxonomie, und -morphologie im direktem Zusammenhang mit der mikrobiellen Gemeinschaftszusammensetzung. Interessanterweise kann die mikrobielle Vielfalt basierend auf der Schwammmorphologie vorhergesagt werden. Dies

bietet spannende Möglichkeiten für weitere Studien mit bildgebenden- oder Merkmalsbasierten Ansätzen.

Basierend auf den generierten Daten ziehe ich die Schlussfolgerung dass jeder Tiefsee-Schwamm eine ganz eigene Zusammensetzung von Mikroben, sowie einen großen Anteil an versteckter mikrobieller Vielfalt besitzt. Desweiteren sind Tiefsee-Schwamm-Mikrobiome auf globaler Eben nicht gut miteinander vernetzt und zeigen eine beachtliche Heterogenität auf der lokalen Ebene. Interessanterweise wurde ein Tiefsee-spezifisches Schwamm-Mikrobiom entdeckt. Zusammengefasst deuten die Ergebnisse meiner Dissertation auf eine starke Einbindung von Tiefsee-Schwamm-Mikrobiomen in den ökologischen Kontext hin.

Im Rahmen dieser Doktorarbeit habe ich eine Datengrundlage zu Tiefsee-Schwamm-assoziierten Mikrobiomen aufgebaut, einen großen Anteil an neuer Diversität entdeckt und Muster der Spezifität, Stabilität und Variabilität beschrieben. Desweiteren habe ich die Umwelt- und Wirts-bezogenen Einflussfaktoren auf die mikrobielle Gemeinschaftszusammensetzung identifiziert. Aus methodischer Sicht habe ich ein Computerprogramm entworfen und entwickelt (die *SVAmPEX*-Software) welches die Archivierung und benutzerfreundliche Anwendbarkeit der generierten Datengrundlage zu Tiefsee-Schwamm-assoziierten Mikrobiomen ermöglicht. Da die mikrobielle Gemeinschaftszusammensetzung im direkten Zusammenhang mit der Gesundheit von Schwämmen steht, sind diese Art von grundlegenden Daten wertvoll zur Überwachung der Intaktheit und Widerstandsfähigkeit von Tiefsee-Schwämmen. Die gesammelten Informationen dieser Doktorarbeit stellen langfristig die wissenschaftliche Grundlage zur Verbesserung der Schutz- und Management-Strategien von gefährdeten Schwamm-Gründen in der Tiefsee dar.

Introduction

The deep-sea ecosystem

More than 90 % of planet Earth's livable volume lies in the deep ocean (Levin and Bris, 2015; **Box 1**). Fascinatingly in a thought experiment, this implies that if one would take a random sample of planet Earth from outer space, probability is high to catch a sample from the deep sea (Boetius, oral communication). Still, the deep sea is largely undersampled today due to its vast space and the need of sophisticated technology to explore it (**Fig.1**). As this technology has only become available to science in recent years, less than 5 % of the deep-sea have been explored and less than 0.001 % have been sampled quantitatively so far (Snelgrove, 1999; Danovaro, 2009; Ramirez-Llodra et al., 2010; Danovaro et al., 2014; Corinaldesi, 2015). These severe sampling constraints are hampering evaluations on the applicability and robustness of well-established shallow water concepts in the deep ocean. Laborious, expensive, and exhaustive sampling efforts are needed to overcome this circumstance.

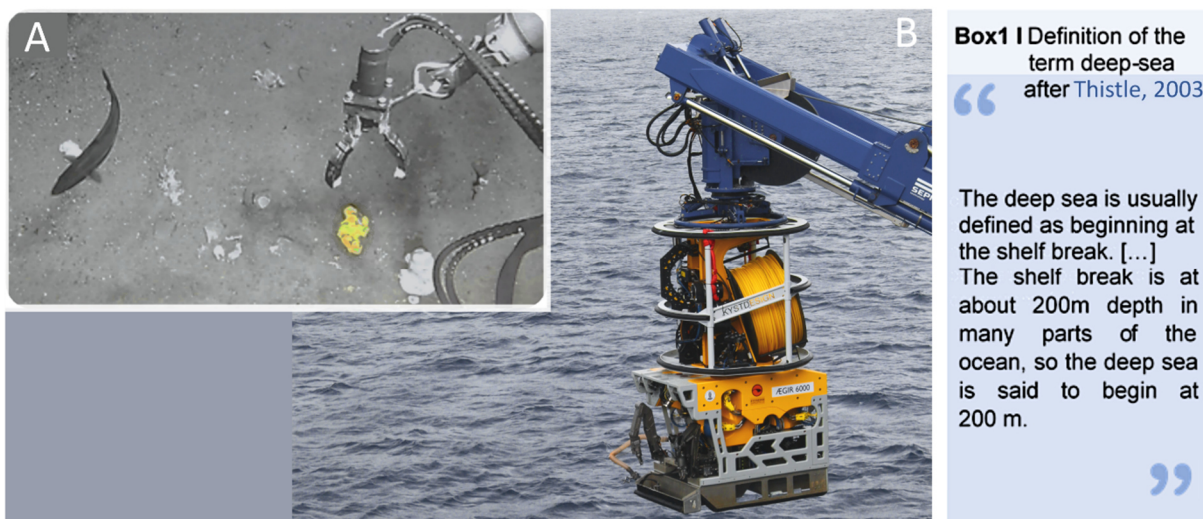


Fig. 1 A) Sampling of a *Styphnus fortis* sponge by remotely operated vehicle (ROV). **B)** Deployment of ROV Ægir 6000 (UiB) from the Norwegian research vessel G.O.Sars.

Most parts of the deep-sea are perceived as nutritionally dependent on the export of organic matter from the surface ocean (Ramirez-Llodra *et al.*, 2010), and characterised by low concentrations of particulate matter. At the same time, dissolved inorganic components (dissolved inorganic carbon, and nutrients such as nitrate, phosphate, and silicate) have comparably high concentrations in the deep ocean due to ongoing remineralisation in the water column. The hydrostatic pressure increases with depth, while temperature, light and oxygen concentrations decrease. Distinct species distribution patterns exist between shallow and deep waters. Species richness generally decreases with depth, although this may strongly depend on the location and taxon studied (Costello and Chaudhary, 2017). Total pelagic bacterial cell numbers typically decrease with depth and the respective microbial community composition may differ between depth layers (Karner *et al.*, 2001; Busch *et al.*, 2017; **Fig.2a**). Microbial

biogeographic patterns are considered to emerge mainly from two processes, which are (i) dispersal limitation, and (ii) environmental selection (Martiny *et al.*, 2006; Fierer, 2008; Hanson *et al.*, 2012). The principle *everything is everywhere but the environment selects* (Baas Becking, 1934) is a popular statement among ecologists and states that microbes are globally distributed, but depending on the environmental conditions only a fraction of them is abundant and active. Critics pointed out that Baas Becking's statement neglects the aspect of dispersal limitation, and that it is impossible to test scientifically whether *everything is everywhere*. However, recent developments in sequencing technology have revealed that a big pool of rare microbes exists (the *rare biosphere*; Pedrós-Alió, 2012). Whether a deep-sea specific microbiome exists in marine sponges will be discussed in this PhD thesis.

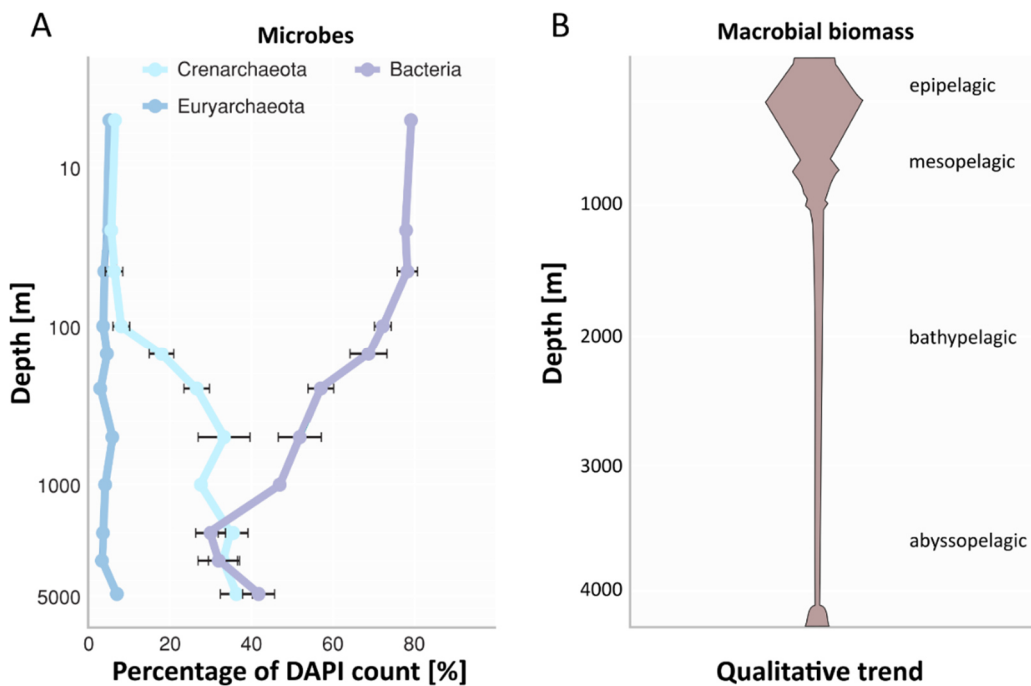


Fig. 2 A) Vertical distribution of pelagic seawater microbes, modified after Karner *et al.* (2001). **B)** Qualitative vertical trend of macrobial biomass, modified after Marshall (1971).

Macrobial biomass is generally lower in the deep-sea compared to shallow depths (Marshall, 1971, **Fig.2b**). Deep-sea species are considered to be relatively widespread and less rich than shallow water species (Costello and Chaudhary, 2017). As population densities in the deep-sea are comparably low, deep-sea organisms need a high dispersal capacity to maintain a high gene flow. Movements of currents connect the ocean, making it the largest interconnected habitat on Earth (Ramirez-Llodra *et al.*, 2011). *Connectivity* is a major determinant of species distributions and biogeographical boundaries and can be defined as the ease with which dispersal occurs across various boundaries and spatial scales (Watling *et al.*, 2013). Connectivity may be interrupted by biophysical barriers, which inhibit gene flow between organisms. Such biophysical barriers can be induced by physical properties (e.g. watermasses) or by geologic features (e.g. seamounts, canyons, banks, etc.). On the contrary, geologic features can also function as stepping stones which promote connectivity between sites. This has for example been suggested for the connectivity of *Bathymodiolus* spp. mussels between hydrothermal vents (Breusing *et al.*, 2016).

The deep-sea is traditionally counted as an extreme habitat (van Dover, 2000), and the presumably most iconic deep-sea habitats are hydrothermal vents. Due to the particularly challenging conditions to maintain basic life functions in extreme environments, microbial symbionts may play a particularly relevant role in adapting their host to challenging environmental conditions (Bang *et al.*, 2018). Three particularly well-studied animal partners in deep-sea chemosymbioses are tube worms (*Riftia pachyptila*; Hinzke *et al.*, 2019), vesicomyid clams (Johnson *et al.*, 2017), and mussels (*Bathymodiolus*; Ansorge *et al.*, 2019). A prominent example of a particularly biodiverse habitat in the deep ocean outside hydrothermal vents are cold-water coral reefs.

Sponge ground ecosystems

Sponge grounds are widely distributed in the world's ocean, but understudied habitats. Existing as either monospecific or multispecific aggregations, sponge grounds increase the structural complexity of the seafloor and thereby provide a shelter for other organisms (**Fig.3**). The deep North Atlantic harbours several known sponge grounds, which include monospecific glass sponge grounds (e.g. of *Vazella pourtalesii* on the Scotian shelf offshore Canada, and of *Pheronema carpenleri* offshore the Azores), as well as multispecific sponge grounds (e.g. the Schulz Bank seamount, the Avilés Canyon System, and the Le Danois Bank). Sponge grounds are considered biodiversity hotspots in the deep ocean because they enhance local species richness and modulate ecosystem dynamics (Maldonado *et al.*, 2017).

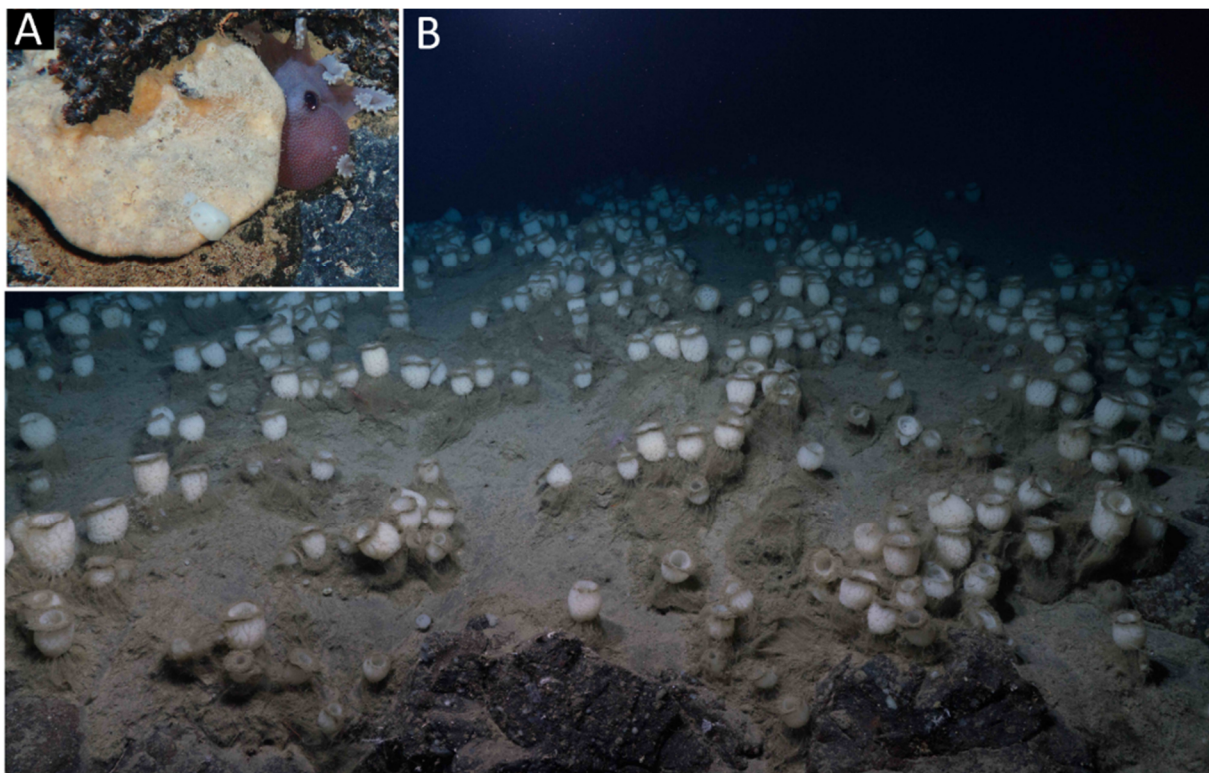


Fig. 3 A) A cephalopod hiding beneath a sponge. Picture taken by ROV Kiel 6000, GEOMAR. **B)** Monospecific aggregations of the glass sponge *Pheronema carpenleri* observed off-shore the Azores. © Rebikoff foundation; Scientific release for IMAR Azores, LULA1000 dive 133 04-07-2018.

Although marine environments are among the most diverse ecosystems on planet Earth (Snelgrove, 2016), only a small part of the world oceans is covered by national legislation. Governance of biodiversity in the open oceans is still in its infancies, and open questions are still to whom the deep-sea ‘belongs’ and what its monetary value is. A recent study was dedicated to assess the monetary loss upon removal of sponges by bottom trawling (Pham *et al.*, 2019). It was estimated that the monetary value of the sponge-driven ecological functions, which are lost upon sponge removal (especially nitrogen and carbon cycling activities via active seawater filtration), is nearly double the market value of the fish catch. A removal of sponges also means that an unknown number of sponge-specific microbial taxa disappear concomitantly. The establishment of conservation areas is therefore fundamental. Recent works have come up with sophisticated ideas to establish a new generation of marine protected areas which should be highly dynamic and interconnected (Fox *et al.*, 2016; Maxwell *et al.*, 2020). Deep-sea sponge grounds are commonly declared as Vulnerable Marine Ecosystem (VME) in need of protection.

Sponges as animal hosts

Today more than 8500 sponge species are formally described, but the true diversity is expected to be much higher (Appeltans *et al.*, 2012; van Soest *et al.*, 2012) (see **Fig. 4** for some examples of existing deep-sea sponge diversity). The phylum *Porifera* (i.e. sponges) consists of four taxonomic classes: *Calcarea* (calcareous sponges), *Demospongia* (demosponges), *Hexactinellida* (glass sponges), and *Homoscleromorpha*. 83.4 % of all known sponge species are demosponges (Steinert, 2015). *Hexactinellida* consist of approximately 500 species globally, and show the greatest diversity between water depths of 300 m to 600 m, whereas only few representatives can be found in shallow water depths (< 15 m) (Leys *et al.*, 2007). Interestingly, the probably oldest animal on Earth is a deep-sea glass sponge (*Monorhaphis chuni*) and was estimated to be 11 000 years old (Jochum *et al.*, 2012). A particularly iconic group of deep-sea (demo-)sponges are carnivorous sponges (*Cladorhizidae*; Vacelet and Boury-Esnault, 1995), which prey on small pelagic animals.

Sponges have a simple morphology and their size ranges from a few millimeters up to the size of a small bus (Bell, 2007; Wagner and Kelley, 2017). Their morphological variety includes massive growth forms, vases, tubes, bowls, balls, branching forms, encrusting ones, and many more. Sponges also come in a large variety of colors (e.g. Freckelton *et al.*, 2012). This rich color palette is often paralleled by the presence and production of bioactive compounds inside the sponge tissue (~ 5300 sponge-derived natural products existed already back in 2009; Lejon *et al.*, 2011). In fact, sponges and their associated microbes are one of the richest marine natural product producers, with over 200 new compounds reported each year. Examples of these bioactive compounds are terpenoids, alkaloids, peptides, and polyketides (Blunt *et al.*, 2006; Han *et al.*, 2019). Whether deep-sea sponges have similarly high secondary metabolism capacities as their shallow water counterparts remains to be tested. From an ecological perspective the need to produce defensive substances may be lower in the deep-sea compared to shallow waters. Sponges also perform biomineralisation to build up their skeleton. Sponge skeletal elements are composed of calcareous (CaCO₃) or siliceous (SiO₂) spicules (Taylor *et al.*, 2007; Hentschel *et al.*, 2012a). Other spicule types are made of chitin or

the collagenous protein spongin. Spicules are commonly used by taxonomists for classification (Uriz *et al.*, 2003; van Soest *et al.*, 2012).

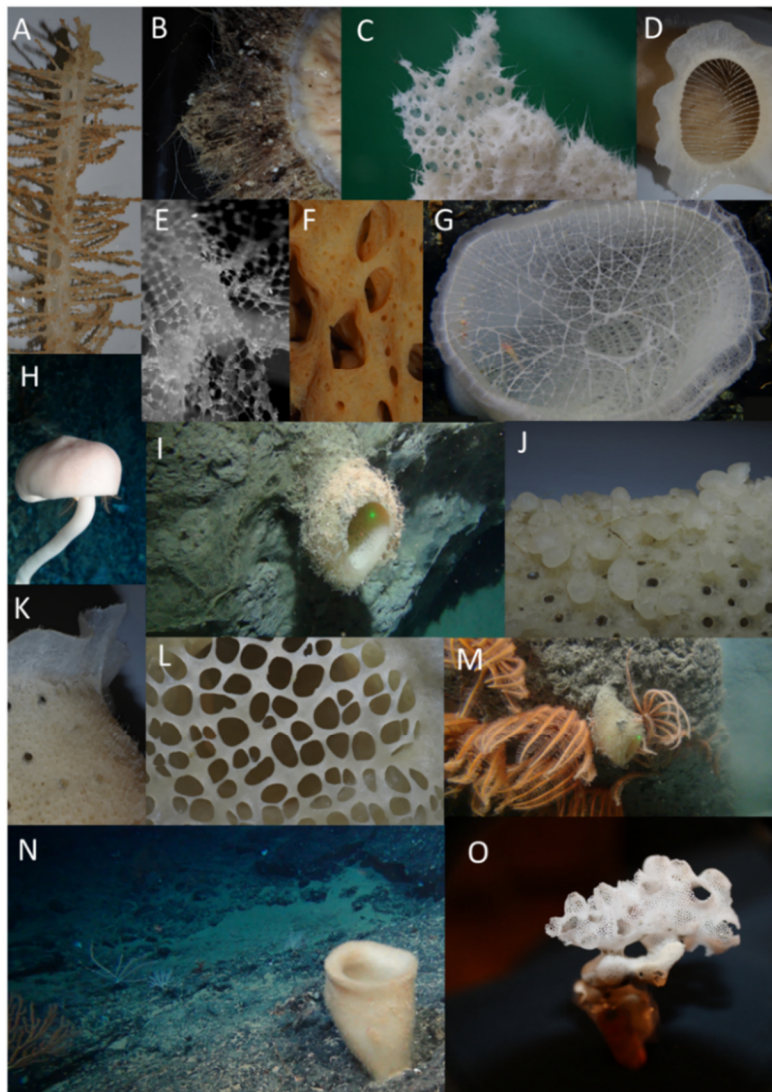


Fig. 4 Examples of deep-sea sponges sampled within the frame of this thesis. **A)** The carnivorous sponge *Walteria leuckarti*. **B)** Cross-section of a *Geodia vaubani* individual. **C)** The glass sponge *Schaudinnia rosea*. **D)** The glass sponge *Regadrella pedunculata*. **E)** *Schaudinnia rosea*, visualised under a binocular. **F)** Detailed view of *Poliopogon* sp. **G)** The glass sponge *Dictyaulus* sp. **H)** A *Caulophacus* individual in its natural habitat. **I)** A glass sponge photographed *in situ*. **J)** Close-up of the glass sponge *Regadrella okinoseana*. **K)** Detailed view of the glass sponge *Regadrella pedunculata*. **L)** Close-up of the glass sponge *Regadrella okinoseana*. **M)** *In situ* footage of the glass sponge *Scyphidium australiense*. **N)** Unidentified, vase-shaped sponge. **O)** The glass sponge *Aphrocallistes beatrix*.

Picture credits:

A,B,D,F,J,K,L: Peter Schupp, ICMB
 C,E,O: Kathrin Busch, GEOMAR
 G,H,I,M,N: ROV Kiel 6000,
 GEOMAR.

Most sponges are highly efficient filter feeding organisms, which can filter up to 24 000 L kg⁻¹ d⁻¹ (Hentschel *et al.*, 2003). The complete sponge bauplan is highly adapted towards efficient filter feeding (**Fig.5a**). It consists of a sophisticated aquiferous system (Leys *et al.*, 2011) coupled with adapted sponge cell types. Choanocytes play a particular relevant role in maintaining the water flow through the animal (Reiswig, 1971a; Vacelet and Donadey, 1977; Weisz *et al.*, 2008). Choanocytes are specialised flagellated cells which are arranged in so called choanocyte chambers, and are equipped with a collar of microvilli that surround the flagella to retain food particles from the inflowing water. Larger food particles may be taken up by archaeocytes before even reaching the choanocyte chambers (Michin, 1900; Simpson, 1984; **Fig.5b**). Exo- and endopinacocytes can also play a role in food uptake in the sponge outer tissue layers (Pourbaix, 1933). Uptake mechanisms of dissolved components include an absorption of nutrients from the water (Püttner, 1909, 1914; Reiswig, 1971b; Yahel *et al.*, 2003). Dissolved oxygen is presumably taken up via diffusion inside the water channels and the choanocyte chambers (Barker Jørgensen *et al.*, 1986).

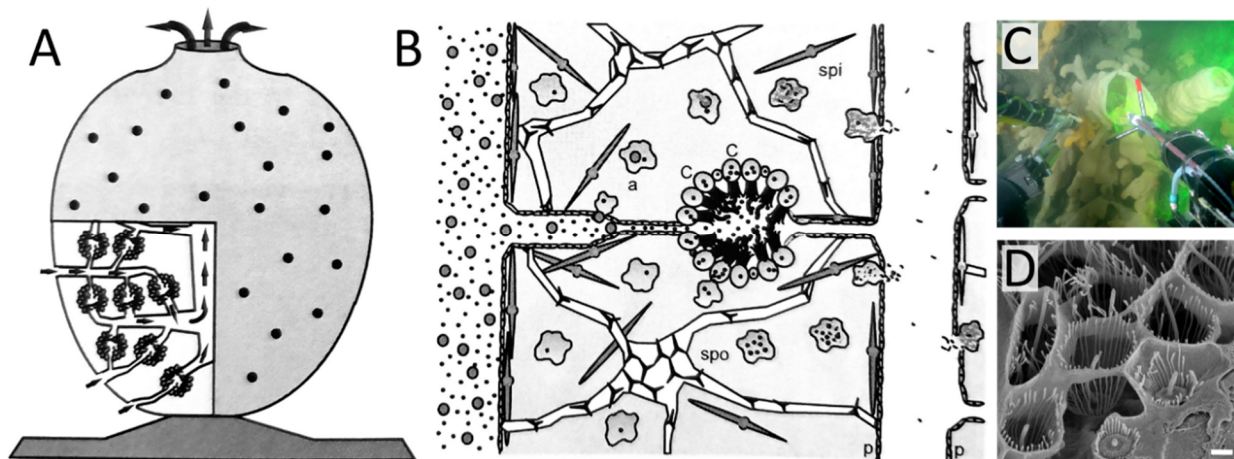


Fig.5 Sponge bauplan. **A)** Habitus of a schematic sponge with several small pores for water inflow (ostia) and one opening for water outflow (osculum). The schematic cross-section depicts the aquiferous system (water channels and choanocyte chambers). The waterflow through the sponge body is indicated by arrows. **B)** Close-up of a water channel and choanocyte chamber (during food uptake; dark grey circles of different sizes). The following cell types and structures are shown: p=pinacocytes, c=choanocytes, a=archaeocytes, spi=spicules, spo=spongin fibres. Figs.A-B were originally published by Osinga *et al.*, 1999; Sipkema, 2004. **C)** Water outflow from the osculum of a deep-sea glass sponge is visualised by green fluorescein dye and recorded by a ROV manipulator arm. Fig.C was originally published by Leys *et al.*, 2011. **D)** Scanning-electron micrograph of a choanocyte chamber (close-up). Scale bar=1 μ m. Fig.D was originally published by Kahn and Leys, 2017.

Due to their high filtering capacity, sponges are considered to play a major role in benthopelagic coupling by mediating the flow of pelagic particulate organic matter (POM) to the benthos (Gili and Coma, 1998; Cattaneo-Vietti *et al.*, 1999; Richter *et al.*, 2001; Ribes *et al.*, 2003, 2005; Lesser, 2006; Pile and Young, 2006; Van Oevelen *et al.*, 2009; Maldonado *et al.*, 2012; Perea-Blázquez *et al.*, 2012). While sponges can filter particles in size from small virus (Hadas *et al.*, 2006) up to larger zooplankton (Vacelet and Boury-Esnault, 1995), they are most efficiently removing particles in the size range of nano- and picoplankton (i.e. 0.2 to 2.0 μ m; retention rates up to 99 %; Pile *et al.*, 1997; Ribes *et al.*, 1999; Kötter and Pernthaler, 2002; Hadas *et al.*, 2009). Although sponges are traditionally assumed to be little selective in their particle uptake (Reiswig, 1971a; Wolfrath and Barthel, 1989; Turon *et al.*, 1997), several studies suggest even a size-independent food selection in some sponges (Frost, 1976, 1980; Van de Vyver *et al.*, 1990; Yahel *et al.*, 2006; Maldonado *et al.*, 2010a). One important component in the uptake of particulate organic carbon (POC) by sponges, is the grazing on bacterial cells. This was also shown for deep-water sponges (Pile and Young, 2006; Yahel *et al.*, 2007; Leys *et al.*, 2018; Bart *et al.*, 2020a), in fact the highest benthic grazing rate of any suspension-feeding community to date was measured on a deep-sea glass sponge ground (Kahn *et al.*, 2015). However, the majority of organic matter in the deep-sea is present in dissolved form. Shallow water sponges are known to play a relevant role in the cycling of dissolved organic matter (DOM; Van Duyl *et al.*, 2008; Ribes *et al.*, 2012), via the *sponge loop* (De Goeij *et al.*, 2013). In the sunlit ocean as much as > 90 % of the total sponge diet is made up by DOM (Yahel *et al.*, 2003; De Goeij *et al.*, 2008a; Mueller *et al.*, 2014a). By extraordinary frequent cell shedding of choanocytes, sponges produce a substantial amount of detritus (POM) (De Goeij *et al.*, 2009;

Alexander *et al.*, 2014, 2015; Maldonado, 2016), which is then fed on by detritivores and passed on to higher trophic levels (De Goeij *et al.*, 2013). As the function is analogous of the *microbial loop* (Azam *et al.*, 1983), this loop is termed *sponge loop* (De Goeij *et al.*, 2013; Mueller *et al.*, 2014a,b; Rix *et al.*, 2016).

Host-microbe symbioses

While many marine organisms live in symbiosis with only a few microbial taxa (e.g. the squid *Euprymna scolopes*; Koch *et al.*, 2020, vesicomid clams and *Bathymodiolus* mussels; Sogin *et al.*, 2020), a single sponge can harbour as much as several dozen bacterial phyla, translating into thousands of bacterial species. That said, sponges are one of the most complex *metaorganisms*. The term *metaorganism* was created to represent both, the animal host and its symbiotic partners (Margulis, 1998; McFall-Ngai *et al.*, 2013; Thompson *et al.*, 2017). However, terminology in the host-microbe symbioses field is not always strictly defined. For example several synonyms exist alongside each other (e.g. *metaorganism* and *holobiont*), and the same term is sometimes used with a different interpretation (e.g. *symbiosis*). **Box 2** includes a glossary of the terms and connotations used in this PhD thesis.

Box2 | Glossary

- **symbiosis:** the intimate association between unlike biological taxa
- **microbes:** microorganisms (here: mainly bacteria and archaea)
- **microbiome:** the entity of microbial symbionts (here: mainly bacteria and archaea)
- **holobiont:** the entity of the host plus its microbial symbionts (syn. metaorganism)
- **metaorganism:** the entity of the host plus its microbial symbionts (syn. holobiont)
- **morphology:** the appearance/shape of an organism
- **metagenomics:** the entire genetic content of all microbes in a sample
- **amplicon sequences:** 16S rRNA gene sequences based on PCR products

Symbiosis (*sensu lato*) is a widespread and advantageous strategy which offers new functions and opportunities to the involved partners. For example organisms may be able to expand their habitat (Márquez *et al.*, 2007; Simonsen *et al.*, 2017) and metabolic (Dubilier *et al.*, 2008; Salem *et al.*, 2017) range. Symbioses may ultimately lead to emergent, complex ecological functions which require an assemblage of (micro-)organisms, rather than isolated individuals. One way to predict interactions between (micro-)organisms is the creation of co-

occurrence networks derived from community composition data; (e.g. Thomas *et al.*, 2016). Another option are predicted models which integrate the metabolic processes of individual microbial cells (e.g. derived from metagenomic binning) and illustrate their metabolic intertwinedness (Moitinho-Silva *et al.*, 2017a; Slaby *et al.*, 2017). Generally, interactions between (micro-)organisms or cells (or other biological units) are multidirectional, and can be of positive, negative, or neutral nature for each involved partner. A compilation of different kinds of interspecific relationships and examples in the sponge context are shown in **Table 1**.

Table 1 Compilation of relationships with and within the sponge holobiont. First two columns after Campbell and Reece, 2009.

Interspecific relationship	Nature of interaction	Example in the sponge context
competition	+/- or -/-	competition between microbes (Esteves <i>et al.</i> , 2017): partner A = bacterium producing antimicrobial activities partner B = pathogenic bacterium
predation	+/-	phagocytic sponge cells (archaeocytes) that capture food bacteria (Leys <i>et al.</i> , 2018)
parasitism	+/-	<i>Cliona</i> boring sponges that drill into coral heads (Riesgo <i>et al.</i> , 2014)
herbivory	+/-	sponge holobionts graze on phytoplankton (Ribes <i>et al.</i> , 1999), but are not herbivores <i>sensu strictu</i>
mutualism (symbiosis)	+/+	sponge host producing ammonia, nitrifying microbes producing nitrate (Bayer <i>et al.</i> , 2014a)
parabiosis (commensalism)	+/0	interactions of the sponge core microbiome (Thomas <i>et al.</i> , 2016)
metabiosis	+/0	spicule mats/sponge grounds providing a habitat; or the putative sponge-induced triggering of the Cambrian explosion (Tatzel <i>et al.</i> , 2017)

Besides the nature of the interaction, also the *interaction strength* is crucial to consider (Sachs, 2004; Sachs and Simms, 2006). This however proves challenging, because the temporal and spatial scale may impact estimates of the interaction strength. Long-term feedbacks, several alternative pathways of effects, and non-linearities may further challenge a mechanistic understanding of interactions (Wootton and Emmerson, 2005). Not many studies exist which mechanistically estimate interaction strength, but existing ones suggest that only a few strong

and many weak interactions exist within communities (Wootton and Emmerson, 2005). Bacterial partners are generally the best studied microbial partners in the host-microbe symbiosis context, while knowledge on archaea is still limited.

Microbial diversity in sponges

Members of the domain *Archaea* occur widely in the world's ocean, despite they were once thought to be found only in extreme environments (Massana *et al.*, 2000; Karner *et al.*, 2001). Pelagic representatives of the *Thaumarchaeota* (previously *Crenarchaeota*; *MarineGroup I*) increase with depth in terms of their absolute and relative abundance (Karner *et al.*, 2001) and are considered to be ammonia-oxidizing chemoautotrophs which have a potential for mixotrophic uptake of organic carbon (DeLong, 1992; Fuhrman *et al.*, 1992). Particularly prominent in this respect are thaumarchaeal members of the family *Nitrosopumilaceae* (Könneke *et al.*, 2005; *Nitrosopumilus maritimus*). The first archaeon described from a marine sponge (the sponge *Axinella mexicana*) was the psychrophilic thaumarchaeon *Cenarchaeum symbiosum* (Preston *et al.*, 1996; which belongs to the *Nitrosopumilaceae*). Similar to seawater-living *Thaumarchaeota*, *Cenarchaeum symbiosum* has been classified as either a strict autotroph (using CO₂ as carbon source) or as a mixotroph (using organic material as carbon sources) (Hallam *et al.*, 2006). A broader association of archaea with sponges has since been demonstrated (Holmes and Blanch, 2007; Lee *et al.*, 2011; Polónia *et al.*, 2014, 2015). Among the few studies on deep-sea sponge microbiomes are three assessments which have proposed archaea to be important members of a potential deep-sea specific sponge microbial community (Kennedy *et al.*, 2014) and which have highlighted an archaeal predominance (Pape *et al.*, 2006; Jackson *et al.*, 2013). Additional studies have identified ammonia-oxidizing archaea (AOA) as the main contributors to nitrification in sponges from cold-water habitats (Hoffmann *et al.*, 2009; Radax *et al.*, 2012a; Li *et al.*, 2014; Tian *et al.*, 2016).

Members of the domain *Bacteria* occur ubiquitarily. Within the global ocean, bacteria are estimated to consist of more than $\sim 2 \cdot 10^6$ species (Curtis *et al.*, 2002). In the pelagic realm SAR11 bacteria (belonging to the *Alphaproteobacteria*) are considered the most abundant group, representing about 25 % of all plankton (Giovannoni, 2017). Proteobacteria (mostly *Alpha*- and *Gammaproteobacteria*) represent one of the most successful bacterial phyla within sponges. Especially uncultivable lineages belonging to the families *Rhodobacteraceae* and *Rhodospirillaceae* are dominant members of sponge-associated microbial communities (Naim *et al.*, 2014; Karimi *et al.*, 2017, 2019). Another prominent bacterial phylum thriving within sponges are *Chloroflexi*. Members of this phylum (especially SAR202) are wide-spread in sponges (Moitinho-Silva *et al.*, 2017b; Bayer *et al.*, 2018) and considered to fulfill important metabolic functions in the sponge holobiont (e.g. cycling of organic matter, Bayer *et al.*, 2018). The SAR202 clade is also known to be abundant in the deep-sea pelagic community, and is considered to play a relevant role in the degradation of recalcitrant organic matter (Landry *et al.*, 2017). Rather restricted to the euphotic water layers are phototrophic *Cyanobacteria*. In the pelagic realm, the free-living cyanobacterial *Synechococcus/Prochlorococcus* clade is considered to be responsible for around 20-40 % of chlorophyll biomass and carbon fixation

(Partensky *et al.*, 1999). Many tropical sponges are inhabited by large abundances of cyanobacterial symbionts, particularly of the genus *Synechococcus* (Unson *et al.*, 1994; Hentschel *et al.*, 2006; Bayer *et al.*, 2014b), and specifically the sponge symbiont *Synechococcus spongiarum* (Steindler *et al.*, 2005; Erwin and Thacker, 2008). These lineages play a relevant role in the nitrogen cycling of the holobiont by fixing nitrogen as ammonia. Several microbial sponge symbionts are known for their involvement in sponge holobiont nitrogen cycling, including processes of nitrogen fixation, nitrification, anaerobic ammonia oxidation, and denitrification (Hoffmann *et al.*, 2009; Fan *et al.*, 2012; Liu *et al.*, 2012; Radax *et al.*, 2012b). *Nitrospira* are nitrite-oxidising bacteria, presumably responsible for the conversion of nitrite to nitrate (i.e. the second step of nitrification) within sponges (Hentschel *et al.*, 2002; Schmitt *et al.*, 2007; Taylor *et al.*, 2007; Off *et al.*, 2010). Nitrifying prokaryotes are known to be slow growing and hard to culture (Off *et al.*, 2010). Actually the majority of environmental microbes remains uncultivated and is therefore described by the term “microbial dark matter” (Rinke *et al.*, 2013). Along this line, an *in vitro* cultivation is presently also impossible for the predominant sponge associated microbial phylotypes, despite considerable efforts and various tested approaches (Sipkema *et al.*, 2011; Lavy *et al.*, 2014; Steinert *et al.*, 2014; Versluis *et al.*, 2017; Gutleben *et al.*, 2020). It was hypothesized that the inability to adequately recreate the sponge microenvironment (i.e. the mesohyl conditions), hamper cultivation success as natural sponge-microbe associations are complex interaction networks involving extensive cross-feeding activities (Pande and Kost, 2017; Gutleben *et al.*, 2020). Cultivation would be particularly valuable for members of two sponge-symbiont phyla: the candidate phylum *Poribacteria* and the *Entotheonella*. *Poribacteria* are equipped with a complex enzymatic repertoire for carbon degradation, several putative symbiosis factors, and were initially considered to be sponge exclusive (Fieseler *et al.*, 2004; Lafi *et al.*, 2009; Hochmuth *et al.*, 2010; Kamke *et al.*, 2013, 2014; Jahn *et al.*, 2016). *Entotheonella* are particularly known for their role in the production of sponge derived natural products and the resulting potential in biotechnological and medical applications (Piel *et al.*, 2004; Freeman *et al.*, 2016; Bhushan *et al.*, 2017; Mori *et al.*, 2018). In terms of taxonomic diversity, 41 microbial phyla have previously been detected within sponges (Thomas *et al.*, 2016). While the microbial community composition is well studied for shallow water sponges, only a handful of studies were available on deep-sea sponge microbiomes when this PhD work was started back in 2016, preventing a comprehensive statement about the prevailing bacterial taxa in this habitat.

Environmental drivers of sponge symbioses

Interactions of the sponge holobiont with the environment are multidirectional: while the sponge holobiont affects biogeochemical cycling of key nutrients like carbon (C), nitrogen (N), phosphorous (P), and silicate (Si), also the ecosystem supposedly influences the microbial community composition. Biogeochemical drivers of the microbial community composition are however not well understood for sponges in general, and in particular not for the deep-sea. In terms of dissolved nutrients, Si processing in sponges is exclusively linked to skeleton formation (Maldonado *et al.*, 2019), while C, N, and P fluxes result from metabolic processes (e.g. feeding, respiration, egestion, excretion; Maldonado *et al.*, 2012). In terms of inorganic C

cycling, autotrophic carbon fixation has been observed for some sponge holobionts, also in the deep-sea (Van Duyl *et al.*, 2008, 2020). A more uncommon nutritional relationship is the the symbiosis between methanotrophic bacteria and cladorhizid sponges in the deep ocean (Vacelet *et al.*, 1995, 1996; Vacelet and Boury-Esnault, 2002; Taylor *et al.*, 2007). The nitrogen cycle is probably the best studied biogeochemical cycle in sponge-microbiology. While sponges can function as a net nutrient source (by OM degradation resulting in ammonium and nitrate; Jiménez and Ribes, 2007; López-Acosta *et al.*, 2019), they can also function as a net nutrient sink (via microbially mediated denitrification and anammox) (Schlappy *et al.*, 2010; Ribes *et al.*, 2012). This controversy seems to also hold true for deep-sea sponges, suggesting that also in the deep-ocean microbially mediated nitrogen cycling plays a relevant and plastic role (Hoffmann *et al.*, 2009; Radax *et al.*, 2012b; Rooks *et al.*, 2020). Vice versa, the influence of biogeochemical parameters on the sponge microbiome had not been investigated for deep-sea sponges at the time when this thesis was started. For shallow-water sponges, sponge microbiome composition was not related with DOC, POC, phosphate and NO_x fluxes, while significant correlations between microbial community structure and NH₄⁺ flux were observed (Gantt *et al.*, 2019). Overall results on the variability of sponges in respect to environmental drivers of microbiome variation were inconsistent between studies. Several studies revealed significant differences in microbial community composition over spatial and temporal scales, geographic location, nutrient concentration or habitat (Anderson *et al.*, 2010; Turque *et al.*, 2010; Fiore *et al.*, 2013; Burgsdorf *et al.*, 2014; Luter *et al.*, 2015; Weigel and Erwin, 2016; Swierts *et al.*, 2018; Cleary *et al.*, 2020; Santos Ferreira *et al.*, 2020). Other studies on the contrary, observed stable microbial community compositions between locations, seasons and depths (Hentschel *et al.*, 2002; Taylor *et al.*, 2005; Erwin *et al.*, 2012, 2015; Björk *et al.*, 2013; Pita *et al.*, 2013a; Simister *et al.*, 2013; Reveillaud 2014). It was suggested, that different host-microbe interaction strengths of different sponge species may explain contrasting results (Lee *et al.*, 2009; Cleary *et al.*, 2013; Thomas *et al.*, 2016). Recently conducted large-scale surveys within the frame of the Earth Microbiome Project (EMP) have identified salinity as a prominent environmental driver for both, host-associated as well as free-living microbes globally (Thompson *et al.*, 2017). A network analysis of the global shallow water sponge EMP has further revealed that the overall structuring of the microbiome is influenced by abiotic factors, while the core community is driven by biotic interactions (Lurgi *et al.*, 2019). In both of these large-scale approaches the number of abiotic/environmental factors included was rather small being limited to depth, temperature and salinity. Information about the relation of microbiome interactions and variations with other physical and biogeochemical parameters was non-existing when this PhD thesis was started. Also no sophisticated attempt of embedding sponge microbiome variability patterns into the oceanographic context had been performed (e.g. translating temperature and salinity into watermasses).

Specificity and stability

The specificity and stability of host-microbe associations is a major theme in symbiosis research. Hentschel *et al.* (2002) were the first to conclude that sponges of different phylogenetic orders and sampling regions harbour a uniform and phylogenetically complex microbial community. The authors defined *sponge-specific clusters* (SSC) as groups which contain at least three sequences that fulfill the following three criteria: “(i) have been recovered from different sponge species and/or from different geographic locations, (ii) are more closely related to each other than to any other sequence from non-sponge sources, and (iii) cluster together independent of the treeing method used” (Hentschel *et al.*, 2002). In recent years, with novel methodological advances at hand, a higher resolution of microbial community data could be achieved and SSC-sequences were retrieved (in low abundances) from other environments than sponges (e.g. Webster *et al.*, 2010; Lee *et al.*, 2011; Taylor *et al.*, 2013). Nevertheless, despite this development, sponge microbiomes are today still considered as overall distinct from environmental reference samples in terms of their richness and composition (Thomas *et al.*, 2016; but note also Cleary *et al.*, 2019). To account for the occurrence of sponge-specific microbes in the environment it was proposed to change the term sponge-specific to sponge-enriched (Moitinho-Silva *et al.*, 2014; Thomas *et al.*, 2016).

Host taxonomy was found to be correlated with the composition of the sponge-associated microbiome, leading to largely species-specific communities (Pita *et al.*, 2013a; Easson and Thacker, 2014; Thomas *et al.*, 2016; Moitinho-Silva *et al.*, 2017c). Schmitt *et al.* (2012) classified the sponge microbial community into three main components, which are the core community, the variable community, and the species-specific community. In comparison to the species-specific community, the core community represents a rather small fraction, consisting of only a few shared abundant bacterial taxa. Subsequent studies have further expanded the knowledge on the core community, revealing that this fraction can also include (permanently) rare symbiont taxa (< 1% relative abundance) (Erwin *et al.*, 2015). The recent study of Thomas *et al.*, 2016 has added the terminology of microbial specialists and generalists onto the stage. It has revealed that shallow water sponge microbiomes consist of generalists and specialists (rather than opportunists), with generalist symbionts comprising the core microbiome. Interestingly this large-scale study has shown that those symbionts which are phylogenetically unique to sponges do not contribute disproportionately to the core microbiome. Further, rather the complexity of the microbiome was shaped by host phylogeny than its composition (Thomas *et al.*, 2016).

The HMA LMA dichotomy

Based on the microbial load inside the tissue, sponges are classified as low microbial abundance (LMA) sponges, or high microbial abundance (HMA) sponges (Vacelet and Donadey, 1977; Reiswig, 1981; Hentschel *et al.*, 2003). This dichotomy comes along with various physiological and ecological differences of the respective sponge types (**Table 2**).

Table 2 Compilation of ecological differences between shallow water HMA and LMA sponges.

	HMA	LMA	reference
Diversity			
diversity and richness of microbial morphotypes/taxa	high	low	Vacelet and Donadey, 1977, Hentschel <i>et al.</i> , 2003, Kamke <i>et al.</i> , 2010, Schmitt <i>et al.</i> , 2012, Giles <i>et al.</i> , 2012, Moitinho-Silva <i>et al.</i> , 2017b
bacterial indicator phyla	Chloroflexi, Acidobacteria, Poribacteria,...	Proteobacteria, Cyanobacteria,...	Fieseler <i>et al.</i> , 2004 Erwin <i>et al.</i> , 2011 Giles <i>et al.</i> , 2012 Blanquer <i>et al.</i> , 2013 Moitinho-Silva <i>et al.</i> , 2014 Wilson <i>et al.</i> , 2014 Bayer <i>et al.</i> , 2014b Moitinho-Silva <i>et al.</i> , 2017b
archaeal indicator taxa	Cenarchaeum	Nitrosopumilus	Hallam <i>et al.</i> , 2006 Polónia <i>et al.</i> , 2018
occurrence of sponge-specific clusters	common	rare	Hentschel <i>et al.</i> , 2002, Taylor <i>et al.</i> , 2007, Erwin <i>et al.</i> , 2011, Giles <i>et al.</i> , 2012
specificity of bacteria and archaea	more distinct from seawater	less distinct from seawater	Moitinho-Silva <i>et al.</i> , 2014 Polónia <i>et al.</i> , 2015 Turon <i>et al.</i> , 2018
specificity of virome	more distinct from seawater	less distinct from seawater	Laffy <i>et al.</i> , 2018, Jahn, 2019
specificity of eukaryotic symbionts	unclear	unclear	Simister <i>et al.</i> , 2012a, Rodríguez-Marconi <i>et al.</i> , 2015, Naim <i>et al.</i> , 2017, De Mares <i>et al.</i> , 2017
Emergence			
transmission mode of symbionts	mixed mode of transmission, but potentially stronger vertical transmission/ weaker horizontal transmission	mixed mode of transmission, but potentially weaker vertical transmission/ stronger horizontal transmission	Schmitt <i>et al.</i> , 2008 Ribes <i>et al.</i> , 2015, Björk <i>et al.</i> , 2019, Rodrigues de Oliveira <i>et al.</i> , 2020
Distribution			
microbial density	high (10 ⁸ to 10 ¹⁰ bacteria/g or mL)	low (10 ⁵ to 10 ⁶ bacteria/g or mL)	Hentschel <i>et al.</i> , 2006, Weisz <i>et al.</i> , 2007, Poppell <i>et al.</i> , 2014, Gloeckner <i>et al.</i> , 2014, Bayer <i>et al.</i> , 2014b
spatial occurrence of symbionts	mostly extracellular in the mesohyl	mostly extracellular in the mesohyl	Vacelet and Donadey, 1977, Wilkinson, 1978, Friedrich <i>et al.</i> , 1999, Taylor <i>et al.</i> , 2007, Burgsdorf <i>et al.</i> , 2019
Interactions			
general nutrition mode	more symbiont-derived	less symbiont-derived	Rix <i>et al.</i> , 2020, Weisz <i>et al.</i> , 2007
difference in microbiome core functional gene prevalence?	no	no	Bayer <i>et al.</i> , 2014a
microbiome core functional gene abundances	N-cylce: higher <i>narG</i> , <i>nosZ</i> , <i>ureC</i> C-cylce: higher <i>mcrA</i> , <i>pcc</i> , <i>RuBisCo</i> , <i>xylA</i> Oxygen stress: higher <i>katE</i>	N-cylce: higher archaeal and bacterial <i>amoA</i> , <i>hao</i> , <i>nirS</i> , <i>nrfA</i> , Phosphate limitation: higher <i>pstS</i>	(when comparing sponges from the same location)

how large is the variability of holobiont acting either as N source or N sink?	higher	lower	Maldonado <i>et al.</i> , 2012, Morganti <i>et al.</i> , 2017, Hoer <i>et al.</i> , 2018
typical role:			
NO_x⁻	source	source	
NH₄⁺	sink	source	
DON/DOM	sink	source	
N₂	sink or source		
nitrification (NO_x⁻ production, NH₄⁺ uptake)	higher	lower	Corredor <i>et al.</i> , 1988, Diaz and Ward, 1997, Jiménez and Ribes, 2007, Southwell <i>et al.</i> , 2008, Maldonado <i>et al.</i> , 2012, Ribes <i>et al.</i> , 2012
diazotrophy	higher (but overall low/neglectable)	lower	Southwell, 2007 Maldonado <i>et al.</i> , 2012, Ribes <i>et al.</i> , 2015
phosphate release	equal (source)	equal (source)	Maldonado <i>et al.</i> , 2012
silicate uptake	unclear	unclear	Maldonado <i>et al.</i> , 2005
oxygen consumption/respiration	higher	lower	Reiswig, 1974, Reiswig, 1981
autotrophy			
photoautotrophy	higher	lower	Maldonado <i>et al.</i> , 2012, Freeman <i>et al.</i> , 2014
chemoautotrophy	higher?	lower?	assumed based on expected presence of nitrifying & mixotrophic symbionts (e.g. Pori-bacteria; Siegl <i>et al.</i> , 2011), van Duyl <i>et al.</i> , 2020 Van Duyl <i>et al.</i> , 2008
direct CO₂ fixation	unclear	unclear	
feeding on DOM	higher	lower	Reiswig, 1974, Ribes <i>et al.</i> , 2012, Maldonado <i>et al.</i> , 2012, Hoer <i>et al.</i> , 2017
	or equal but	or equal but	De Goeij <i>et al.</i> , 2008b De Goeij <i>et al.</i> , 2013 Mueller <i>et al.</i> , 2014a
	higher uptake rate (glucose, amino acids)	lower uptake rate (glucose, amino acids)	Rix <i>et al.</i> , 2020
feeding on POM	lower	higher	Freeman and Thacker, 2011, Reiswig, 1971b, Weisz <i>et al.</i> , 2007 Rix <i>et al.</i> , 2020
upregulated host immune receptors in response to microbial elicitors	G-protein coupled receptor, scavenger receptor cysteine-rich	NLR/NLR-like, LRR containing gene, fibrinogen-like genes	Pita <i>et al.</i> , 2018a, Schmittmann <i>et al.</i> , 2020
other properties			
pumping rate	lower	higher	Vacelet and Donadey, 1977, Boury-Esnault <i>et al.</i> , 1990, Weisz <i>et al.</i> , 2008, Poppell <i>et al.</i> , 2013
* size of choanocyte chambers	smaller	larger	
* density of choanocyte chambers	lower	higher	
* length of water channels	longer	shorter	
* width of water channels	smaller	wider	
mesohyl density	high	low	Vacelet and Donadey, 1977, Weisz <i>et al.</i> , 2008
host genome size	small	large	Ryu <i>et al.</i> , 2016
dichotomy related to host phylogeny?	no	no	Gloeckner <i>et al.</i> , 2014, Moitinho-Silva <i>et al.</i> , 2017b
microbiome stability			
seasonal and interannual dynamics	low	low	Erwin <i>et al.</i> , 2015 and references therein
variability across replicates	higher	higher	

-----Table 2 continued-----

Despite the large knowledge base of the ecological and physiological consequences, the actual drivers of the HMA-LMA dichotomy remain poorly understood. Previous studies have shown (e.g. Moitinho-Silva *et al.*, 2017b) that the HMA-LMA dichotomy in shallow water sponges

cannot clearly be attributed to phylogenetic and evolutionary drivers. For shallow water sponges the pattern of complex HMA microbiomes and reduced LMA microbiomes (which are largely restricted to a proteobacterial and cyanobacterial dominance) is well proven (Moitinho-Silva *et al.*, 2017b). But the transferability of this concept to different ecological contexts remains unknown. Previous studies have beautifully paved the way towards upscaling approaches by training machine learning algorithms to predict the HMA-LMA status of sponges (Moitinho-Silva *et al.*, 2017b). This PhD thesis will pick up these ideas and aims to develop them further.

Vertical versus horizontal symbiont transmission

In sponges both, symbiont vertical transmission (from parents to early life stages) and horizontal transmission (from the environment) were observed (Schmitt *et al.*, 2008; Ribes *et al.*, 2015; Björk *et al.*, 2019; Rodrigues de Oliveira *et al.*, 2020). Sponges are generally considered as hermaphrodites and are capable of both, asexual (production of fragments, buds, gemules) and sexual reproduction (production of oocytes, sperms, embryos, larvae) (Simpson, 1984; Sipkema, 2004). While a predominance of vertical transmission may result in the formation of stable microbial communities which are linked to host dispersal and evolutionary trajectories (Erpenbeck *et al.*, 2002; Thacker and Starnes, 2003), horizontal transmission may result in biogeographic patterns (Pita *et al.*, 2013b). In shallow-water sponges the most common transmission mode of microbial taxa seems to be the so called leaky vertical transmission, or mixed mode of transmission. This process includes both, vertical and horizontal symbiont transmission (Schmitt *et al.*, 2008; Ribes *et al.*, 2015; Björk *et al.*, 2019; Rodrigues de Oliveira *et al.*, 2020). Whether deep-sea sponges follow similar symbiont transmission strategies as their shallow water counterparts is currently unknown, but evaluated for two deep-sea sponge species in this thesis.

Sponge holobiont responses to stress

Knowledge about the variability of marine communities in response to environmental parameters is key, especially in times of rapid environmental change. Several experiments have been carried out to assess the impact of predicted climate scenarios on shallow water sponge microbiomes. For example Fan *et al.*, 2013 have studied the effects of thermal stress on the sponge-holobiont. They observed that high temperatures caused a direct stress response in both, the host and microbiome. Basic nutritional relationships were disrupted during early heat stress, finally leading to a replacement of sponge-specific symbionts by novel microbial taxa. This effect is commonly known as *dysbiosis* (e.g. reviewed in Pita *et al.*, 2018b; and assessed also in Greenspan *et al.*, 2020). Pronounced effects of thermal stress were found on the sponge-associated bacterial community (Webster *et al.*, 2008a; Simister *et al.*, 2012b; Strand *et al.*, 2017; Ramsby *et al.*, 2018), the host (Pantile and Webster, 2011; Massaro *et al.*, 2012; Webster *et al.*, 2013), and also the virome of sponges (Laffy *et al.*, 2019). Major expected shifts in the future ocean abiotic conditions besides ocean warming are ocean deoxygenation, ocean acidification, and ocean eutrophication. Sponges were observed to show a remarkable plasticity in respect to oxygen concentration tolerance (Hanz *et al.*, 2019). Schuster *et al.*

(2020) assessed the impact of seasonal anoxia on the shallow water sponge microbiome, revealing putative microbial taxa which may confer deoxygenation tolerance. Multiple studies have checked on the impact of ocean acidification on the sponge microbiome (Morrow *et al.*, 2015; Kandler *et al.*, 2018; Botté *et al.*, 2019), and the impact of eutrophication (Luter *et al.*, 2014). Furthermore several experiments were conducted combining multiple stressors, such as warming and acidification combinations (Bennett *et al.*, 2017), or warming and eutrophication (Simister *et al.*, 2012c). In addition to chronic climate change, acute anthropogenic disturbances play a relevant role in the ocean. Among the most pressing topics which have been explored in the sponge microbiome context are oil pollution (Luter *et al.*, 2019), heavy metal exposure (Webster *et al.*, 2001), and bottom trawling (Pineda *et al.*, 2016, 2017a,b). Particularly urging topics for the deep ocean are deep-sea mining and plastic pollution, but when this thesis was started no studies were existing on potential impacts of these two factors on sponge-associated microbial communities.

In a rather conceptual study, Björk *et al.*, 2017 observed a negative relationship between temporal variability of microbial taxa and their mean abundance in a long-term (3 years) experiment on shallow water sponges. They conclude that stability of the microbiome is not related to diversity, but rather to the density of the core microbiome. A similar observation that stability is not directly related to diversity has been described by Glasl *et al.* (2018). Erwin *et al.* (2015) had summed it up nicely: “Understanding the dynamics of host-symbiont interactions is a critical first step in assessing symbiont stability, host resilience and how the interaction of these two factors may determine larger shifts in natural marine communities in the face of growing anthropogenic disturbances“. In addition to targeted experimental studies about the influence of individual parameters, it is crucial to derive at a more systematic evaluation what the actual natural/undisturbed sponge-microbiome looks like. In this respect establishment of baselines is key. The large datasets, which are necessary for baseline establishment, also offer a unique opportunity to define baseline shifts in microbial communities, to detect abnormal shifts and emerging diseases in response to a changing environment (Webster, 2007; Erwin *et al.*, 2015). Habitat destruction, as well as epidemics may severely decimate sponge populations. Numerous previous studies have recorded the outbreaks of sponge epidemics from different geographic locations, including even the deep-sea (Luter *et al.*, 2017), and/or have discussed disease in the sponge context (Webster *et al.*, 2002, 2007, 2008b, 2009; Luter *et al.*, 2010a, 2010b, 2012; Maldonado *et al.*, 2010b; Luter and Webster, 2017; Slaby *et al.*, 2019). As microbes may respond quite rapidly to shifted abiotic conditions, they may serve as early warning of environmental (sublethal) stress (Fan *et al.*, 2013; Glasl *et al.*, 2017, 2019). A recent interesting development is the idea to use sponges as bioindicators for pollutants, such as microplastics (Girard *et al.*, 2020) and toxic metals (Cebrian *et al.*, 2007; Davis *et al.*, 2014). Another recent idea along similar directions is to use sponges (or more precisely the eDNA accumulated in their tissue) for monitoring biodiversity (Mariani *et al.*, 2019). Monitoring biodiversity and protecting the ocean has become a major goal of international politics (Franke *et al.*, 2020).

2.7 Aims & outline of my PhD thesis

Based on the just presented scientific background, I have identified five major research aims within the topics of biodiversity, evolution, biogeography, and ecology, an understanding of which should ultimately lead to improved conservation and management options for deep-sea sponge ground ecosystems:

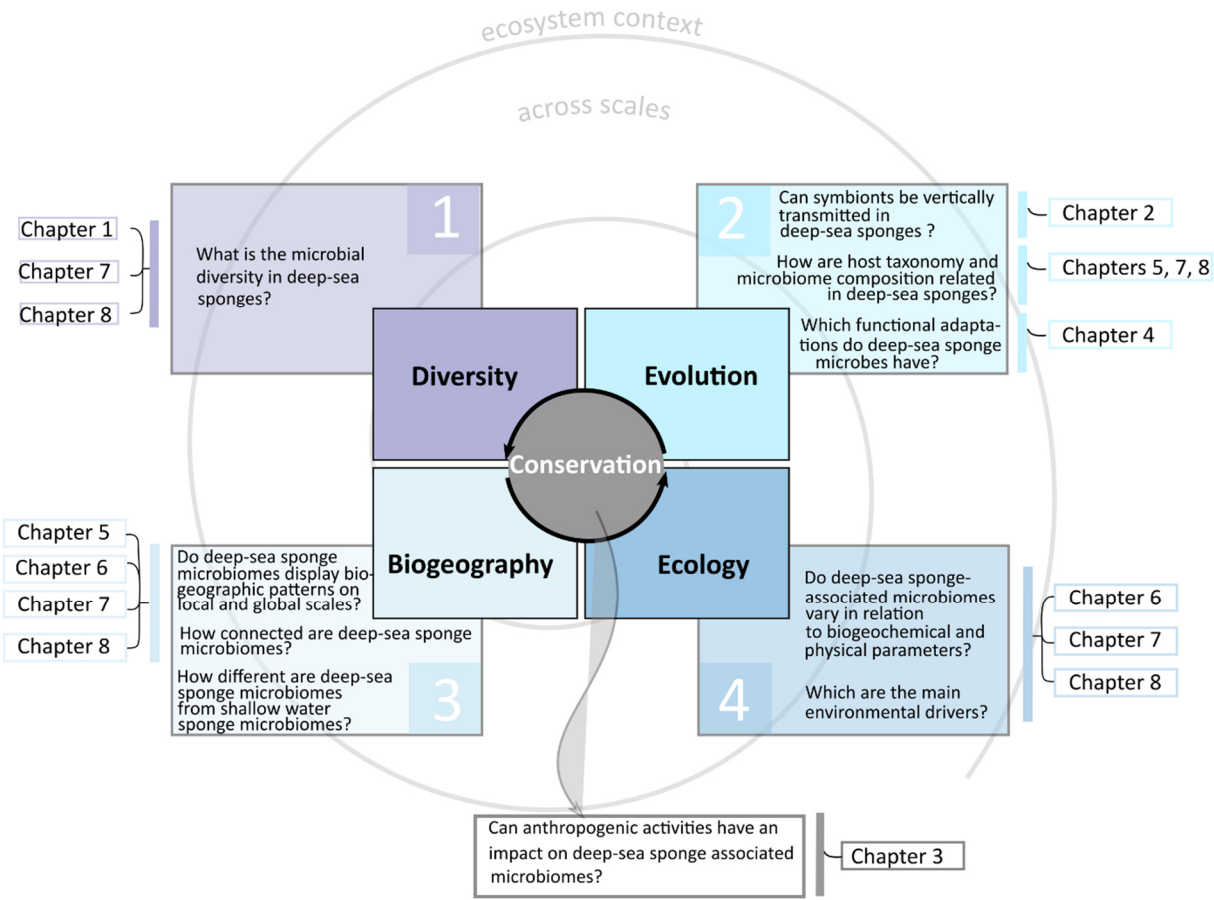


Fig.6 Research aims and questions of this PhD thesis.

Within this thesis, I have worked on different scales, reaching from individual case studies (**Chapters 1-6**) to an integrated global meta-analysis (**Chapters 7-8**). The following specific research aims were defined for the individual chapters:

Chapter 1

Assess whether bacterial community principles also apply to the archaeal communities in deep-sea sponges.

Chapter 2

Evaluate how microbial community composition, diversity, and dominant microbial phyla differ between different deep-sea sponge life stages (adults, embryos, and recruits) for two selected sponge species.

Chapter 3

Examine whether and how deep-sea sponge holobionts respond to the human pressures trawling and marine litter.

Chapter 4

Evaluate the functional adaptations of deep-sea sponge microbes.

Chapter 5

Analyse how well deep-sea sponge holobionts are connected on small spatial scales.

Chapter 6

Examine how natural environmental gradients at seamounts are reflected in the deep-sea sponge and sea-water microbiome.

Chapter 7

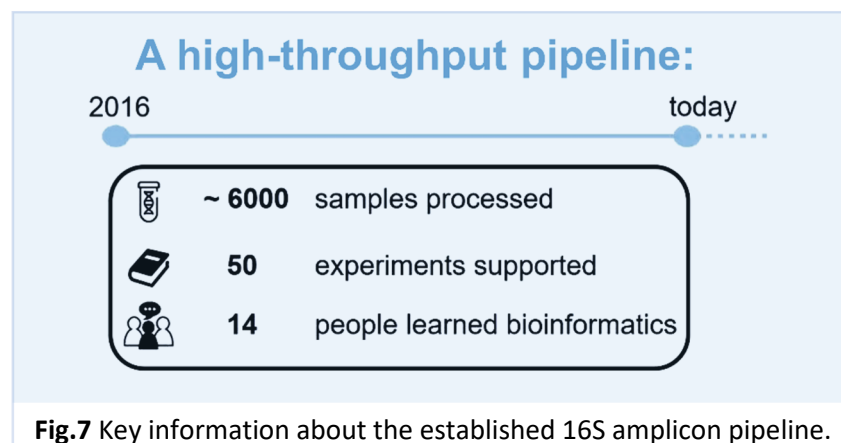
Extract global and universal patterns of deep-sea sponge-associated microbiomes in the ecosystem context.

Chapter 8

Design a software to expand the value of deep-sea sponge-associated microbial baseline data.

This thesis is structured as follows: I will start with a short introduction into the methodological approach (section 3). Afterwards I will continue with a presentation of the main research of this thesis (section 4), where each chapter is either a peer-reviewed publication (**Chapter 1-6**) or a manuscript in preparation (**Chapter 7-8**). An overarching synthesis of the main results of this thesis is presented in sections 5-6.

As the majority of microbial taxa is uncultured today, next-generation sequencing techniques are commonly applied to reveal the present taxonomic biodiversity (by *16S* amplicon sequencing) and function (via the “omics”-repertoire). These technologies are based on the information encoded in genes, transcripts, proteins, or metabolites (Hentschel *et al.*, 2012b). For biodiversity assessments of environmental microbial communities, amplicon sequencing of the *16S* rRNA gene is commonly used. For this technique, the deoxyribonucleic acid (DNA) is first extracted, then a specific region of the *16S* rRNA gene is amplified via polymerase chain reaction (PCR), and sequenced. Sequences are then processed bioinformatically, and taxonomically classified based on reference sequences deposited in public databases. One major advantage is the generation of relative abundance data for uncultured microbial taxa. Since hundreds of samples can be processed simultaneously, this technique allows relatively fast and neat insights into the microbial community composition. Two major technological shifts of recent years include (i) the introduction of amplicon single nucleotide variants (ASVs) which have replaced the traditionally used operational taxonomic units (OTUs) (Callahan *et al.*, 2017), and (ii) a reassessment of the prokaryotic taxonomical framework using full genomes (Parks *et al.*, 2018). Taking the recent developments in the field into account, a standardised, high-throughput *16S* amplicon sequencing pipeline was developed in the frame of this PhD work. The established pipeline was extensively used by a large variety of different projects during the last four years (**Fig.7**).

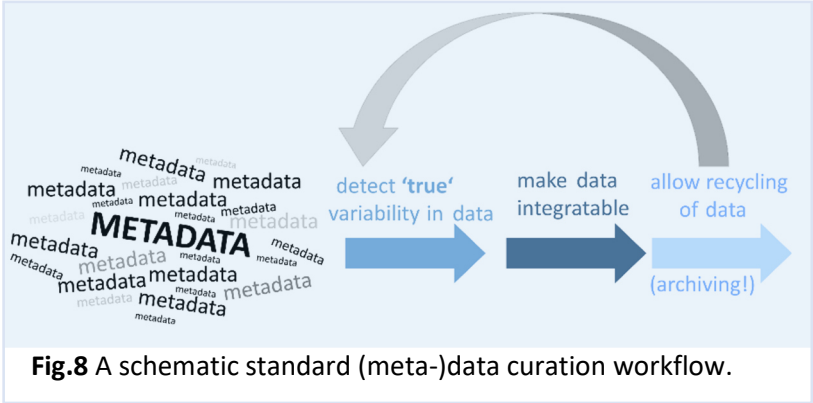


A particular emphasis of the established pipeline was put on extensive quality checks and a high degree of standardisation. A high degree of standardisation is crucial in order to minimise technical variations, which may later on during the analysis infer with true biological variation.

An ocean of data

In order to put the generated data into perspective, ecological metadata are highly valuable. For an efficient data integration and data recycling, standardisation is inevitably necessary, but often done halfheartedly (Yilmaz *et al.*, 2011; Buttigieg *et al.*, 2013; **Fig.8**). As the field of

ocean research is steadily moving more and more towards high digitalisation, it will become increasingly important to generate standardised datasets. Especially advanced bioinformatic techniques (such as artificial intelligence, machine learning) rely strongly on synchronised data formats. For this reason, a sophisticated metadata workflow (from sampling to archiving and visualisation) was established in the frame of this PhD thesis (see **Chapter 8** for details).



Complementary techniques

Amplicon sequencing gains particular strength when being combined with other techniques such as experiments, visualisations (microscopy), and sophisticated statistical approaches. For this PhD thesis, amplicon sequencing was consistently performed for all included studies and chapters. It was then combined with a set of complementary techniques for each case study to answer specific research questions (**Chapter 1-8; Fig.9**).

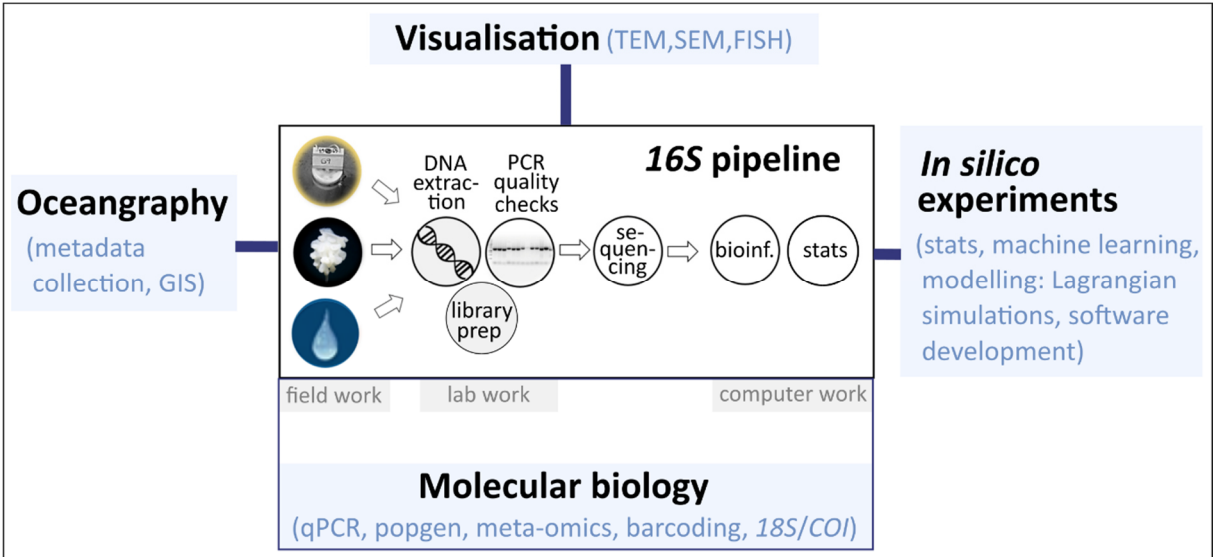


Fig.9 Overview of methodological toolbox used in this thesis.

Interdisciplinary approaches

A holistic understanding of ecosystem dynamics cannot be achieved by a single scientific discipline alone. Instead an interdisciplinary effort is required. This PhD thesis therefore combined approaches from different scientific disciplines spanning marine ecology, microbiology, bioinformatics, biogeochemistry, and physical oceanography (**Fig.10**).

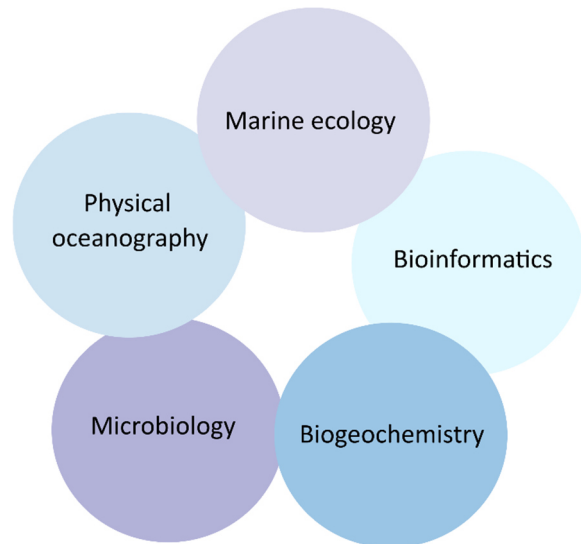


Fig. 10 Overview of scientific disciplines combined in this PhD thesis.

A nested sampling design

Within the frame of this PhD thesis, a major effort was put into field work in order to create (and curate) a large-scale reference database of deep-sea sponge microbiomes. The resulting data set encompasses microbiome data of 1546 deep-sea sponges, seawater, and sediment samples, together with an extensive set of accompanying metadata, from 20 deep-sea expeditions world-wide (**Fig.11**). That said, the overall structure of this PhD thesis is a nested one. The individual **chapters 1-6** represent in-depth studies performed at local case study sites, while **chapters 7-8** represent a global meta-analysis.



Fig.11 Overview of cruises organised within the frame of this PhD thesis.

The importance of biodiversity reference catalogues

Biodiversity reference catalogues are crucial in order to be able to detect, assess and monitor biological resources. Especially standardised large-scale census are valuable in this regard, because missing replication and standardisation typically hamper reliable comparisons among multiple independent small-scale studies, as technical variation may overwhelm biological variation and prevent detection of general patterns or trends (Thompson *et al.*, 2017). Examples of such large-scale efforts with a focus on marine and microbial life, are listed in **Table 3**. My thesis aims to add a new large-scale reference catalogue on the microbial diversity associated with deep-sea sponges.

Table 3 Examples of large-scale coordinated efforts to decipher biodiversity.

initiative	main target organisms
Census of Marine Life (CoML)	all marine organisms
Tara Oceans	planktonic organisms
Malaspina	deep ocean organisms
Ocean Sampling Day (OSD)	marine microbes
Earth Microbiome Project (EMP)	microbes
• Sponge Microbiome Project (SMP)	sponge-associated microbes

Excursion I Box 3

Standardised large-scale initiatives on abiotic parameters in the ocean

- **World Ocean Atlas (WOA):** world ocean climatology
- **Global Ocean Data Analysis Project (GLODAP):** carbon cycle climatology
- **Joint Global Ocean Flux Study (JGOFS):** carbon fluxes
- **World Ocean Circulation Experiment (WOCE):** large parameter set
- **Argo:** conductivity, temperature, depth
- **Geochemical Ocean Sections Study (GEOSECS):** (radio-) chemical, isotopic tracers
- **GEOTRACES:** trace elements and isotopes

Established *16S* pipeline

© 2020 Busch et al.

under embargo at *protocols.io* (Berkeley, USA).



Wet lab SOP of the deep-sea sponge microbiome project ▾

Kathrin Busch¹, Andrea Hethke¹, Ina Clefsen¹, Ute Hentschel¹

¹GEOMAR

Kathrin Busch

[Steps](#) [Materials](#) [Metadata](#)

ABSTRACT

This protocol describes the wet lab standard operating procedures (SOPs) established for the deep-sea sponge microbiome project. It includes the field work procedures, as well as protocols for the isolation and multiplication of microbial genomic DNA for **16S rRNA amplicon sequencing**.

We developed this pipeline with the aim to process a maximum variety of sampletypes. The wet lab part of our amplicon pipeline is modified, but generally based on the guidelines established by the Earth Microbiome Project:

Thompson LR, Sanders JG, McDonald D, Amir A,..., Jansson JK, Gilbert JA, Knight R & The Earth Microbiome Project Consortium (2017) *A communal catalogue reveals Earth's multiscale microbial diversity*. Nature 551:457-463. doi:10.1038/nature24621.

Key characteristics in a nutshell:

	A	B
1	extract_type	DNA
2	library_source	genomic
3	library_strategy	Amplicon (ENA:0000049)
4	lib_layout	paired
5	lib_const_meth	PCR
6	lib_screen	gel electrophoresis
7	target_gene	16S rRNA
8	target_subfragment	V3V4
9	target_main_organism_domain	Bacteria (NEWT:2)
10	seq_meth.platform	Illumina
11	seq_meth.instrument_model	MiSeq
12	seq_chemistry_kitname	MiSeq Reagent Kit v3
13	read_length	2 × 300 bp
14	investigation_type	MIMARKS-survey
15	tissue_fixative	none
16	tissue_store_temp_degrees_c	-80
17	dna_store_temp_degrees_c	-80

BEFORE STARTING

Consider the tab *Materials* for a full list of key resources needed to conduct this protocol.



According to the general principles of molecular work on microbes, we aim to minimise potential cross-contaminations by keeping workspace, devices, and gloves sterile throughout the whole sample handling process.



The deep-sea sponge microbiome project dataset has been collected during 20 **ship expeditions**. For these cruises we established the following standard operating procedures:

A) Metadata collection:

- Gather as much metadata as possible
- Fill in this metadata table for all samples (and add anything else that is worth recording):

 [Example_metadata_sheet.xlsx](#)



FYI:

- We link every collected sample to cruise ids and event ids deposited in the [Pangaea repository](#).
- Usage of the same sample partent ids allows linkage between researchers within the [SponGES project](#).
- We store respective cruise reports in the [OSIS portal](#).

B) Sponge collection:

- Collect four different specimens of the same sponge species per location. Our sampling focus for the deep-sea sponge microbiome project is on adult, healthy looking sponge individuals.

Note: After animal collection, proceed with the preparations as quickly as realistically possible.

- Take a photograph (with size standard and sample name) of every sponge individual.
- From each specimen, aseptically (wear gloves, use scalpel) remove four 5 cm³ pieces of tissue. Our targeted body compartment for tissue sampling is the sponge mesohyl.

Note: Use a fresh scalpel each time between different individuals so that there is no cross-contamination between samples. If you do not have a sufficient amount, then swipe the scalpel with ethanol and a Kimwipe to reduce the carry-over of DNA from one sample to another.

- Rinse dissected pieces by transferring them 3x through sterile filtered seawater. Then gently dip sponge piece on Kimwipe to remove excess liquid and flash-freeze samples. Store samples at -80°C.

Note: Petri plates work well for rinsing of sponges. Technical replicates can be frozen together if tube size allows.

- Flash-freeze additional tissue individually (extra biomass) in a 50 mL falcon tube and store it at -80°C (alternatively -20°C).



All samples are shipped on dry-ice. To prevent a worse case scenario (i.e. loss of precious samples during transportation) we commonly store additional tissue pieces in RNA/later™ Stabilization Solution (ThermoFisher) as backup.

C) Seawater collection:

- Collect 4x 2 L replicates of water (8 L total, 4 Niskins) at the same location where sponges are collected.

Note: Acid-cleaned graduated Nalgene bottles work well for sub-sampling from Niskin bottles.

- Filter each 2 L replicate of water on an individual PVDF filter membrane (Merck Millipore; 0.22 µm pore size, Ø47 mm) using acid-cleaned filter towers.

Note: Rinse filter towers between different samples with MilliQ water.

- Using sterile forceps, remove filter from chimney base, roll up the filter biomass-side in, and place in 15mL falcon tubes. Store in -80° freezer and ship on dry-ice.

Note: To save space in the freezer, filters can also be folded twice in half (similar to a crêpe) and stored in cryo vials.

- Extract and store all available measurements from the CTD cast. Usage of .cnv files facilitates analyses afterwards (.XMLCON +.HEX together work also).

D) Sediment collection:

- Collect 4 sediment push cores from the same location where sponges and water are collected.
- Using sterile spatula (wiped with 70% EtOH), slice off top of core to collect upper 2 cm of sediment.
- Store sediment in individual tubes, freeze in the -80° freezer and ship on dry-ice.



Curation of physical frozen sample collection (sponge tissue, sediment, seawater filters):
The deep-sea collection is situated in an own ½ -80 °C freezer in Prof. Ute Hentschel's lab at GEOMAR, Kiel, Germany. It is currently managed by Kathrin Busch. Storage boxes are labeled with cruise identifiers and tubes are labeled with abbreviated parent ids.

DNA extraction

2



LIMS (Laboratory Information Management System):

For products of the DNA extraction procedure we apply a consecutive labelling system (*WB-xxx* for sponges, *SW-xxx* for seawater, *SED-xxx* for sediment). Further we perform a standardised documentation for each sample processed:

[📎 Example_lab_documentation.xlsx](#)

DNA extraction is performed with the DNeasy Power Soil Kit (Qiagen, Cat. #: 12888-100).
 (the following steps are closely linked to the manufacturers manual)

	A	B
1	Sample input	0.25 g of sponge tissue or sediment; ½ filter for seawater
2	Throughput	1 - 24 samples
3	Time per run or prep	3-4 h



- 2.1 ■ Add 60 µl of Solution C1 and invert several times or vortex briefly.
Note: If Solution C1 has precipitated, heat at 60 °C until precipitate dissolves.



- 2.2 ■ Add sponge tissue, sediment or half of a seawater filter to the PowerBead Tube included in the kit.
 ■ Vortex PowerBead Tubes.
 ■ Secure PowerBead Tubes for bead beating with PowerLyzer® 24 Bench Top Bead-Based Homogenizer (MO BIO Laboratories, #13155) and run program. Settings: 3500 rpm, 2x 30 sec bead beating with 45 sec pause in between.



- 2.3 ■ Centrifuge tubes at 7500 x g for 2 min.
Note: All centrifugation steps in this protocol are performed at room temperature (15-25 °C).
 ■ Transfer the supernatant to a clean 2 mL collection tube.
Note: Expect between 400–500 µL of supernatant.



- 2.4 ■ Add 250 µL of Solution C2 and vortex for 5 sec.



- 2.5 ■ Incubate at 2–8°C for 5 min.



- 2.6 ■ Centrifuge tubes at 10000 x g for 2 min.
 ■ Avoiding the pellet, transfer up to 600 µL of supernatant to a clean 2 mL collection tube.



- 2.7 ■ Add 200 µL of Solution C3 and vortex briefly.



- 2.8 ■ Incubate at 2–8°C for 5 min.



- 2.9 ■ Centrifuge tubes at 10000 x g for 2 min.
 ■ Avoiding the pellet, transfer up to 700 µL of supernatant to a clean 2 mL customer tube.
Note: The total volume will not fit into the collection tubes provided in the kit, it will spill over when you close the lid.



- 2.10
- Shake to mix Solution C4 and add 1200 μL to the supernatant. Invert 8x.



- 2.11
- Load 640 μL onto a MB Spin Column.
 - Centrifuge at 10000 x g for 1 min. Discard flow through.



Repeat step 2.11 twice, until the entire sample has been processed.



- 2.12
- Add 500 μL of Solution C5.



- 2.13
- Centrifuge at 10000 x g for 30 sec.
 - Discard the flow through. Centrifuge again at 10,000 x g for 2 min.
 - Carefully place the MB Spin Column into a clean 2 mL collection tube.

Note: Avoid splashing of Solution C5 onto the column.



- 2.14
- Air dry filter with open lid under the clean bench for 5 min.



- 2.15
- Add 100 μL of prewarmed (37 °C) Solution C6 to the center of the white filter membrane.

Note: For processed seawater filters elute with 50 μL of Solution C6 instead of 100 μL .



- 2.16
- Centrifuge at 10000 x g for 1 min. Discard the MB Spin Column.



- 2.17
- The DNA is now ready for downstream applications.

- Transfer 6 μL of the extract into an extra tube to be used for quality controls (see below).
- Store the rest of the extract in the freezer.

Note: Solution C6 is 10 mM Tris-HCl, pH 8.5 (it does not contain EDTA). We therefore recommend to store the extracted DNA frozen.

- We store the extracts at -20 °C until library preparation. Afterwards we use a rack system for organised long-term storage of extracts at -80 °C.

Remark: We also run kit blanks through our pipeline on a regular basis, as well as two negative and two positive controls for every processed 96-well plate.



Three steps are performed for quality check of the extracted DNA:

	A
1	DNA concentration and purity measurements with a NanoDrop spectrophotometer
2	PCR with universal 16SrRNA primers
3	agarose gel electrophoresis

A) 16S rRNA PCR set up:

	A	B
1	PCR conditions	Specifications
2	27F Forward primer	5'-GAG TTT GAT CCT GGC TCA G-3'
3	1492R Reverse primer	5'-GGT TAC CTT GTT ACG ACT T-3'
4	Template	undiluted sample, 10 ng optimum
5	Template size	1500 bp

To set up parallel reactions and to minimize the possibility of pipetting errors, prepare a PCR master mix by mixing water, buffer, dNTPs, primers and DreamTaq DNA polymerase in order. Prepare enough master mix for the number of reactions plus one or two extra. Aliquot the master mix into PCR-strips/plate and then add template DNA and close with stripe cover/foil.

- Gently vortex and briefly centrifuge all solutions after thawing.
- Place PCR-strips/plate on a cold rack and add the components for a single 20 μ L reaction:

	A	B
1	dH ₂ O	16.4 μ L
2	10x Dream Taq Green Buffer*	2.0 μ L
3	dNTP Mix 10 mM each	0.2 μ L
4	Forward primer 10 mM	0.1 μ L
5	Reverse primer 10 mM	0.1 μ L
6	Dream Taq DNA Polymerase 5 U	0.2 μ L
7	Template DNA 10 pg-1 μ g	1.0 μ L

- Gently vortex the sample stripes and spin down.
- Place the reactions in a thermal cycler. Perform a PCR using following thermal cycling conditions:

	A	B	C	D
1	Step	Temperature [°C]	Time [min]	Number of cycles
2	Initial denaturation	95	3	1
3	Denaturation	95	0.5	34
4	Annealing	56	0.5	34
5	Extension	72	1.5	34
6	Final Extension	72	5	1

B) Agarose gel electrophoresis. Check your DNA on an 1% Agarose gel:

- Mix the agarose with 1xTAE buffer in a 300 mL Erlenmeyer flask.
- Heat up the solution in a microwave until the agarose is completely dissolved.
- Add a magnetic stir bar let it mix and cool down till ~ 45 °C.
- Add Gelgreen (10000x) with a final concentration of 0,5x and mix.
- Carefully transfer the agarose solution into a gel tray with combs.
- After the gel is hardened it can be transferred into the electrophoresis chamber with 1xTAE running buffer.
- Pipette the DNA size (1kb) Ladder and your samples into the slots.
- Start the electrophoresis with a constant power supply of 180 V for 30 min.

We use this template for evaluation of agarose gel electrophoresis:

[Example_gel_evaluation.xlsx](#).

Library preparation



3

Amplification of the V3V4 variable regions of the 16SrRNA gene is done by a one-step PCR. Add the following components for a single 25 µL reaction:

	A	B
1	dH2O	10.25 µL
2	5x Buffer	5 µL
3	dNTPs 10 µmol	0.5 µL
4	Forward primer 341F 0.28 µM	4 µL
5	Reverse primer 806R 0.28 µM	4 µL
6	Phusion Hot Start II High-Fidelity DNA Polymerase 0.5 U	0.25 µL
7	Template DNA	1 µL

We use the primer pair 341F-806R in a dual-barcoding approach. Check out the exact primer composition here: [PrimersV3V4.xlsx](#).

Perform the PCR using the following thermal cycling conditions:

	A	B	C	D
1	Step	Temperature [°C]	Time	Number of Cycles
2	Initial denaturation	98	30 sec	1
3	Denaturation	98	9 sec	30
4	Annealing	55	60 sec	30
5	Extension	72	90 sec	30
6	Final Extension	72	10 min	1
7	Hold	10	∞	-

The amplicon libraries are quality checked by gel electrophoresis, normalised with the SequalPrep Normalization Plate Kit (ThermoFisher Scientific) and pooled equimolarly.

Sequencing



4

Sequencing is performed on a MiSeq platform (MiSeqFGx, Illumina) using v3 chemistry, resulting in 2x 300 bp.

We always run samples in a 384 pool (1 lane) and use a standardised sequencing layout for all runs:

[Example_sequencing_layout.xlsx](#).

A commercial mock community (Zymo Research) is added to every plate and a PhiX sequencing control (20 % spike-in) is utilised. For demultiplexing zero mismatches are allowed in the barcode sequence.

Bioinformatics

5



Our subsequent core bioinformatic pipeline can be found here: [GitHub](#).



You are currently in View mode.



Wet lab SOP of the deep-sea sponge microbiome project ▾

Kathrin Busch¹, Andrea Hethke¹, Ina Clefsen¹, Ute Hentschel¹

¹GEOMAR

Kathrin Busch

Steps **Materials** Metadata

MATERIALS TEXT



Working with deep-sea samples often comes along with large logistic efforts. You can find our standard packing list for a 1-month research cruise here:

[Example_packing_list.pdf](#)

Note: Make sure to watch out for customs, dangerous goods + frozen sample shipping guidelines, and other legal regulations.

Please find a key resource table for Steps 2-4 here:

[Key_resources_table.xlsx](#)

*(the format of this table has been adopted from Cell's STAR*METHODS)*




You are currently in View mode.



Wet lab SOP of the deep-sea sponge microbiome project ▼

Kathrin Busch¹, Andrea Hethke¹, Ina Clefsen¹, Ute Hentschel¹

¹GEOMAR

 Kathrin Busch

Steps Materials **Metadata**

PROTOCOL INFO

Kathrin Busch, Andrea Hethke, Ina Clefsen, Ute Hentschel . Wet lab SOP of the deep-sea sponge microbiome project. **protocols.io**
<https://protocols.io/view/wet-lab-sop-of-the-deep-sea-sponge-microbiome-proj-bkgyktxw>

KEYWORDS

molecular ecology, bacteria, amplicon sequencing, 16S, microbial diversity

IMAGE ATTRIBUTION

K.Busch

WIDGET

Widget code will become available once this protocol is published

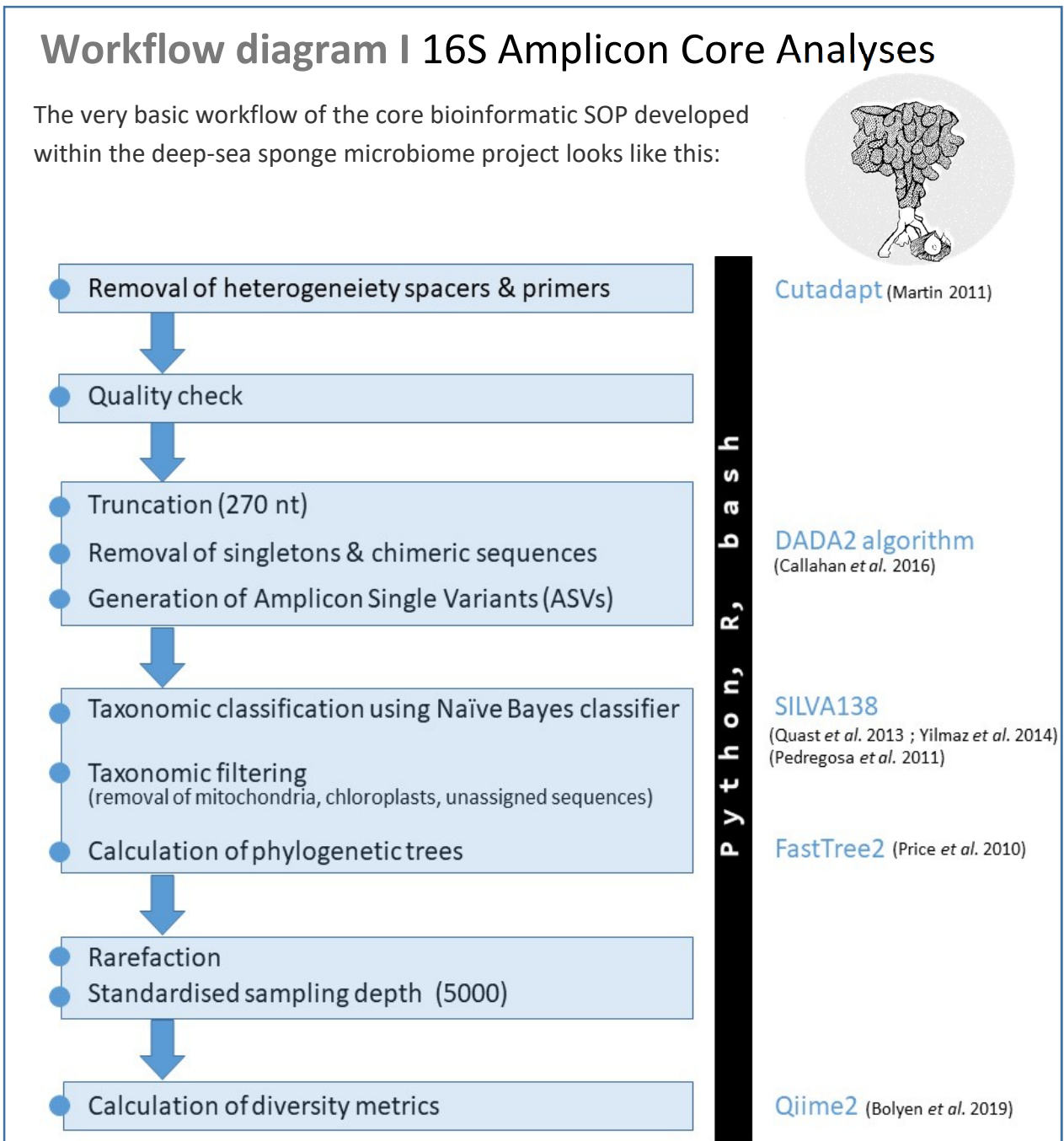
QR CODE



3.2 Established bioinformatic workflow

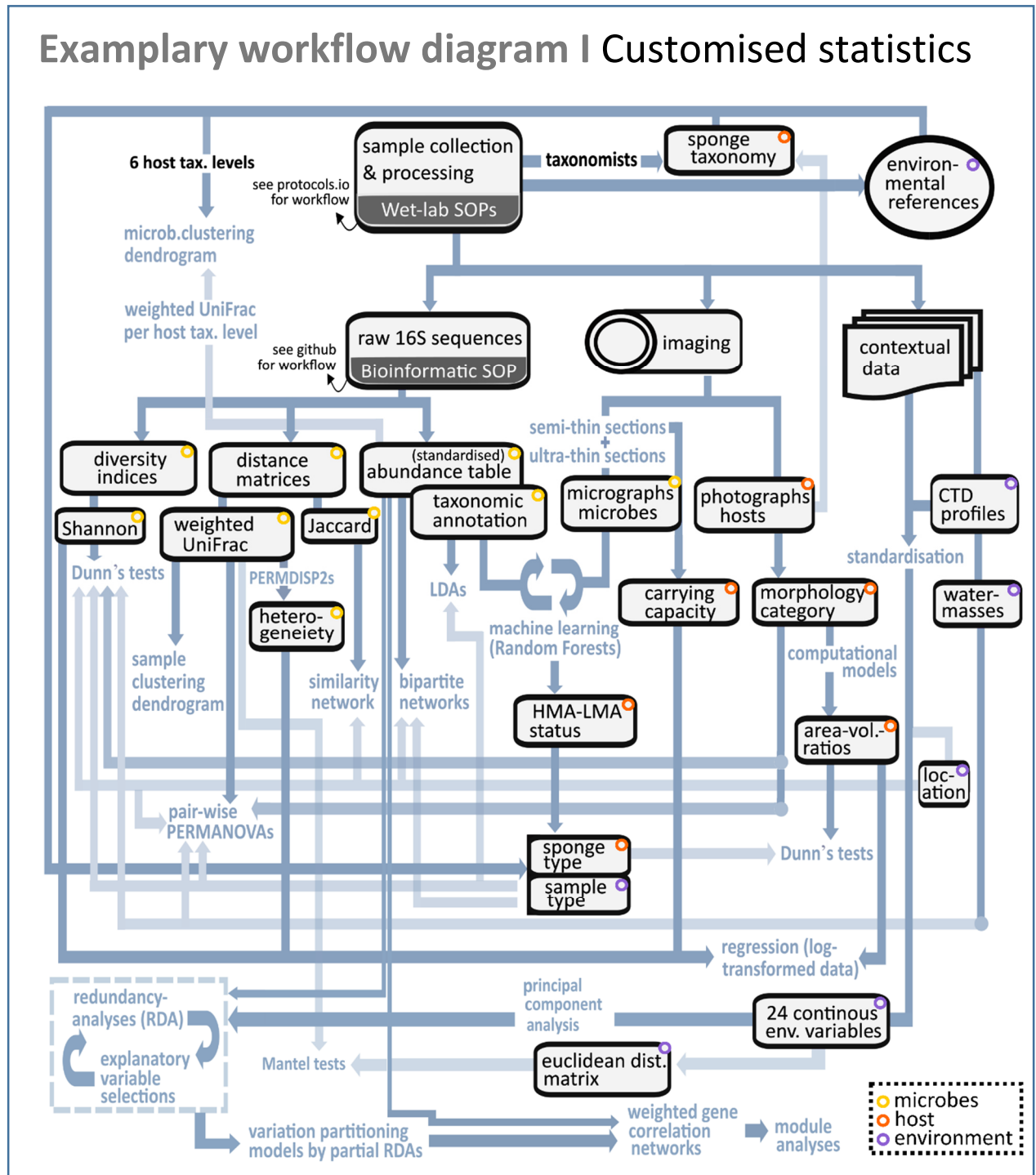
3.2.1 Standard Operating Procedure

The established bioinformatic workflow can be found in the **Appendix** of this PhD thesis and on github (<https://kathrinbusch.github.io/16S-AmpliconCorePipeline/#>; still under embargo).



3.2.2 Customised statistical analyses

One exemplary designed statistical workflow, which has been created and applied for one of the thesis chapters (**Chapter 7**), is shown below.



Main publications and contributions

This thesis is based on the following publications and manuscripts:
(abbreviation **KB** = Kathrin Busch)

1. Steinert, G., **Busch, K.**, Bayer, K., Kodami, S., Arbizu, P. M., Kelly, M., Mills, S., Erpenbeck, D., Dohrmann, M., Wörheide, G., Hentschel, U., Schupp, P. J. (2020) *Compositional and quantitative insights into bacterial and archaeal communities of South Pacific deep-sea sponges (Demospongiae and Hexactinellida)*. *Frontiers in Microbiology*, 11, 716, doi: 10.3389/fmicb.2020.00716.

Participation in	Author initials,	responsibility	decreasing from	left to right
Study design	PS/UH	KB /GS		
Method development	KB /GS			
Sampling	KB /PS			
Completion of experiments	KB /KrB/GW/PS/SK/PA/GS/MD/MK/SM			
Data analysis, interpretation	GS	KB /UH/PS	KrB	
Manuscript writing	GS	KB /UH/PS	KrB	
Manuscript reviewing, editing	KB /GS/GW/KrB/MD/SM/UH/PS			
Data curation	KB /GS			

2. **Busch, K.**, Wurz, E., Rapp, H. T., Bayer, K., Franke, A., Hentschel, U. (2020) Chloroflexi dominate the deep-sea golf ball sponges *Craniella zetlandica* and *Craniella infrequens* throughout different life stages. *Frontiers in Marine Science*, 7, 674, doi: 10.3389/fmars.2020.00674.

Participation in	Author initials,	responsibility	decreasing from	left to right
Study design	KB	UH		
Method development	KB			
Sampling	KB /HTR	EW		
Completion of experiments	KB	KrB	HTR	AF
Data analysis, interpretation	KB	UH	KrB	
Manuscript writing	KB	UH		
Manuscript reviewing, editing	KB /UH/KrB/EW			
Data curation	KB			

3. Busch, K., Beazley, L., Kenchington, E., Whoriskey, F., Slaby, B. M., Hentschel, U. (2020) *Microbial diversity of the glass sponge Vazella pourtalesii in response to anthropogenic activities*. Conservation Genetics. doi: 10.1007/s10592-020-01305-2.

Participation in	Author initials,	responsibility	decreasing from	left to right
Study design	KB	UH	BS	
Method development	KB			
Sampling	LB/EK	FW	KB	
Completion of experiments	KB			
Data analysis, interpretation	KB	UH		
Manuscript writing	KB	UH		
Manuscript reviewing, editing	KB/UH/BS/ LB/EK/FW			
Data curation	KB			

4. Bayer, K. and Busch, K., Kenchington, E., Beazley, L., Franzenburg, S., Michels, J., Hentschel, U., Slaby, B. M. (2020) *Microbial strategies for survival in the glass sponge Vazella pourtalesii*. mSystems, 5 (4), e00473-20. doi: 10.1128/mSystems.00473-20.

Participation in	Author initials,	responsibility	decreasing from	left to right
Study design	BS/KrB	KB/UH		
Method development	KB/BS			
Sampling	LB/EK	KB		
Completion of experiments	KB/BS	SF/JM		
Data analysis, interpretation	KB/KrB/BS/UH			
Manuscript writing	BS/KrB	KB/UH		
Manuscript reviewing, editing	KB/KrB/BS/UH/ EK/LB			
Data curation	KB/BS			

5. Busch, K., Taboada, S., Riesgo, A., Koutsouveli, V., Ríos, P., Cristobo, J., Franke, A., Getzlaff, K., Schmidt, C., Biastoch, A., Hentschel U. (2020) *Population connectivity of fan-shaped sponge holobionts in the deep Cantabrian Sea*. Deep Sea Research Part I: Oceanographic Research Papers. doi: 10.1016/j.dsr.2020.10342.

Participation in	Author initials,	responsibility	decreasing from	left to right
Study design	KB/UH	AR/AB		
Method development	KB/CS/ST	KG		
Sampling	KB/AR/ST/ VK/PR/JC			
Completion of experiments	KB	ST/AR/CS	VK/PR/JC	AF
Data analysis, interpretation	KB/UH	ST/AR	CS/AB	
Manuscript writing	KB/UH	ST/AR	CS/AB	PR/JC/VK
Manuscript reviewing, editing	KB/UH/ST/AR/AB			
Data curation	KB	ST		

6. Busch, K., Hanz, U., Mienis, F., Müller, B., Franke, A., Roberts, E. M., Rapp, H. T., Hentschel, U. (2020) *On giant shoulders: How a seamount affects the microbial community composition of seawater and sponges*. Biogeosciences, 17, 3471-3486. doi: 10.5194/bg-17-3471-2020.

Participation in	Author initials,	responsibility	decreasing from	left to right
Study design	KB/UHe	BM/FM	EMR	
Method development	KB			
Sampling	KB/HTR/UHa/FM/ EMR/			
Completion of experiments	KB	UHa/HTR	AF	
Data analysis, interpretation	KB	UHe/BM/FM		
Manuscript writing	KB/UHe	BM	UHa/FM	
Manuscript reviewing, editing	KB/EMR/BM/UHe/HTR			
Data curation	KB			

7. Busch, K., Slaby, B.M., Beazley, L., Clefsen, I., Colaço, A., Cristobo, J., Federwisch, L., Hethke, A., Kenchington, E., Mienis, F., Riesgo, A., Ríos, P., Pita, L., Schupp, P., Xavier, J.R., Rapp, H.T., Hentschel, U. (in preparation) *How to sum up individuals? Nestedness and connectivity of the global deep-sea sponge microbiome.*

Participation in	Author initials,	responsibility	decreasing from	left to right
Study design	KB/UH			
Method development	KB	IC/AH		
Sampling	KB	BS/LB/AC/ JC/LF/EK/FM/AR /PR/PS/HT		
Completion of experiments	KB	IC/AH		
Data analysis, interpretation	KB			
Manuscript writing	KB	UH		
Manuscript reviewing, editing	to be done by all authors			
Data curation	KB			

8. Busch, K., Schmittmann, L., Hentschel, U. (in preparation) *SVAmPEX – An interactive and interdisciplinary visualisation platform of the global Deep-sea Sponge Microbiome Project database.*

Participation in	Author initials,	responsibility	decreasing from	left to right
Study design	KB	LS	UH	
Method development	KB			
Sampling	-			
Completion of experiments	-			
Data analysis, interpretation	KB			
Manuscript writing	KB			
Manuscript reviewing, editing	KB	UH	LS	
Data curation	KB			

Additional related publications

During the period of this PhD thesis, contributions to multiple additional manuscripts were made which already resulted, or will result, in further publications (with co-authorship of **KB**). As these publications are beyond the scope of this thesis, they are not included in the presented publication list.

Chapter 1

Compositional and quantitative insights into bacterial and archaeal communities of South Pacific deep-sea sponges (Demospongiae and Hexactinellida)

© 2020 Steinert et al.

Original publication by *Frontiers in Microbiology* (Frontiers)

Reprinted under the terms of the Creative Commons Attribution License (CC-BY 4.0).

doi: [10.3389/fmicb.2020.00716](https://doi.org/10.3389/fmicb.2020.00716).

Compositional and quantitative insights into bacterial and archaeal communities of South Pacific deep-sea sponges (Demospongiae and Hexactinellida)

Georg Steinert^{1,2}, Kathrin Busch¹, Kristina Bayer¹, Sahar Kodami³, Pedro Martinez Arbizu³, Michelle Kelly⁴, Sadie Mills⁵, Dirk Erpenbeck^{6,7}, Martin Dohrmann⁶, Gert Wörheide^{6,7,8}, Ute Hentschel^{1,9*} and Peter J. Schupp^{2,10*}

¹ RD3 Marine Symbioses, GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany, ² Institute for Chemistry and Biology of the Marine Environment (ICBM), University of Oldenburg, Oldenburg, Germany, ³ German Center for Marine Biodiversity Research, Senckenberg Research Institute, Wilhelmshaven, Germany, ⁴ National Institute of Water and Atmospheric Research, Ltd., Auckland, New Zealand, ⁵ National Institute of Water and Atmospheric Research, Ltd., Wellington, New Zealand, ⁶ Department of Earth and Environmental Sciences, Paleontology & Geobiology, Ludwig-Maximilians-Universität München, Munich, Germany, ⁷ GeoBio-Center, Ludwig-Maximilians-Universität München, Munich, Germany, ⁸ Bayerische Staatssammlung für Paläontologie und Geologie, Munich, Germany, ⁹ Christian-Albrecht University of Kiel, Kiel, Germany, ¹⁰ Helmholtz Institute for Functional Marine Biodiversity at the University of Oldenburg (HIFMB), Oldenburg, Germany. *Correspondence authors.

In the present study, we profiled bacterial and archaeal communities from 13 phylogenetically diverse deep-sea sponge species (Demospongiae and Hexactinellida) from the South Pacific by 16S rRNA-gene amplicon sequencing. Additionally, the associated bacteria and archaea were quantified by real-time qPCR. Our results show that bacterial communities from the deep-sea sponges are mostly host-species specific similar to what has been observed for shallow-water demosponges. The archaeal deep-sea sponge community structures are different from the bacterial community structures in that they are almost completely dominated by a single family, which are the ammonia-oxidizing genera within the Nitrosopumilaceae. Remarkably, the archaeal communities are mostly specific to individual sponges (rather than sponge-species), and this observation applies to both hexactinellids and demosponges. Finally, archaeal 16s gene numbers, as detected by quantitative real-time PCR, were up to three orders of magnitude higher than in shallow-water sponges, highlighting the importance of the archaea for deep-sea sponges in general.

Keywords: 16S rRNA amplicons, archaea, bacteria, Demospongiae, Hexactinellida, Porifera, quantitative real-time PCR (qPCR), South Pacific Ocean

INTRODUCTION

Marine sponges (Porifera) host a broad range of microorganisms including bacteria, archaea, eukaryotes, and viruses and are therefore considered holobionts (Webster and Thomas, 2016; Pita et al., 2018). Sponges are integral parts of the marine ecosystem as they couple pelagic and benthic ecosystems by virtue of their massive filter-feeding capacities (Vogel, 1977; de Goeij et al., 2017; Pita et

al., 2018). Sponge-associated symbionts perform critical functions for their host, including among others, the provision of nutrients (particularly, for nitrogen and carbon) and chemical defense which affect host health and functioning (Slaby et al., 2019). The microbial consortia of sponges are represented by diverse prokaryotic communities with ≥ 63 phyla having been found in sponges so far (Thomas et al., 2016; Moitinho-Silva et al., 2017b). These prokaryotic

communities show sponge species-specific patterns that differ in richness, diversity, and structure from the prokaryotic seawater communities. The composition of the sponge symbiont consortia is shaped by host taxonomy in that sponge species have species-specific prokaryotic communities (Pita et al.,2013; Easson and Thacker,2014; Thomas et al.,2016; Steinert et al.,2017; Cárdenas et al.,2019). Machine learning provided evidence that the dichotomy between high microbial (HMA) and low microbial abundance (LMA) sponges is a main driver of the sponge-associated community patterns (Moitinho-Silva et al.,2017c). Several prokaryotic phyla and taxa were identified as indicator taxa for either one of the two abundance states, such as Chloroflexi (e.g., SAR202, Caldilineaceae), Acidobacteria (e.g., Solibacteres, PAUC37f, Sva725), Poribacteria or Actinobacteria (e.g., Acidimicrobiia) for HMA sponges, or alternatively, Proteobacteria (e.g., Gammaproteobacteria), Bacteroidetes (e.g., Flavobacteriia), and Planctomycetes (e.g., Planctomycetia) for LMA sponges. The global sponge microbiome data revealed further community features, such as the dominance of specialists and generalists within the symbiont communities, a stable core microbiota, and community structure and functional modularity, with abiotic factors influencing the overall sponge microbiota and biotic factors the prokaryotic core (Thomas et al.,2016; Lurgi et al.,2019).

Most sponge microbiome studies have focused on demosponges that were collected from shallow coastal sites in temperate, subtropical, and tropical sampling locations (e.g., Schmitt et al.,2012; Moitinho-Silva et al.,2014; Naim et al.,2014; Thomas et al.,2016; Steinert et al.,2017; Helber et al.,2019). Considering that the deep-sea is the largest, still relatively underexplored habitat on earth, comparably few studies have been conducted on sponges from remote deep-sea or cold-water locations (e.g., Jackson et al.,2013; Kennedy et al.,2014; Reveillaud et al.,2014; Borchert et al.,2017). Antarctic shallow cold-water demosponges host dominant bacterial taxa that are known to be sponge-associated (Webster et al.,2004;

Rodríguez-Marconi et al.,2015; Cárdenas et al.,2018,2019), and Antarctic deep-water demosponges display high levels of host-specificity (Steinert et al., 2019). In addition, sponges from classes other than Demospongiae, i.e., Hexactinellida, Calcarea, and Homoscleromorpha, are still only poorly covered by 16S rRNA gene sequencing approaches (but see Xin et al.,2011; Tian et al.,2016) and are consequently still underrepresented in global sponge microbiome surveys. Finally, most microbiome studies have used bacterial universal 16S rRNA gene primers, hence the sponge-associated archaeal communities have largely been missed (but see, e.g., Webster et al.,2010; Schmitt et al.,2012; Thomas et al.,2016; Moitinho-Silva et al.,2017a; for exceptions, covering both bacterial and archaeal diversity). In North Atlantic deep-sea sponges, Archaea were previously proposed to be important members of a potential deep-sea specific sponge microbial community (Jackson et al.,2013; Kennedy et al.,2014). Archaeal predominance has also been observed in one Arctic deep-water demosponge species (Pape et al.,2006). Ammonia oxidizing archaea (AOA) were identified as the main contributors of nitrification within the cold-water sponge hosts (Hoffmann et al.,2009; Radax et al.,2012; Li et al., 2014). One recent study underlined the importance of AOA in deep-sea sponges using metagenomic data obtained from one glass sponge (Tian et al.,2016).

Our present study aims to contribute to resolving sponge-prokaryote relationships from understudied habitats by exploring both bacterial and archaeal communities in demosponges and hexactinellids, which were collected from meso-, bathypelagic, and abyssal depths in the South Pacific Ocean offshore New Zealand. We investigate whether current sponge-microbiota paradigms hold up for those remotely collected and partially novel sponge species. We address the following questions: (a) are general principles of the Demospongiae microbiota also present in Hexactinellida prokaryotic communities, (b)

do these sponge-bacterial community principles also apply to the archaeal community structures in marine sponges, and (c) can we compare the observed deep-sea sponge microbiota to the tropical/warm water sponge microbiota? In addition to these main questions,

MATERIALS AND METHODS

Sample Processing and Sponge Taxonomy

During the SO254 expedition of the research vessel “*Sonne*” in the South Pacific Ocean south- to northeast around New Zealand in February 2017, over 200 sponge specimens, of 96 sponge taxa comprising the Porifera classes Demospongiae and Hexactinellida, were collected by using a remotely operated vehicle (ROV “*Kiel 6000*”; <http://jlsrf.org/index.php/lssf/article/view/160>). Of these 200 sponges, only species collected in at least triplicate were considered in the microbiome analysis. This criterion reduced the investigated sponge diversity to 45 specimens and 13 species. Depths of the nine collecting sites of the sponges included in the analysis ranged from 472 to 4160 m (**Figure 1**, **Table 1** and Supplementary Table S1).

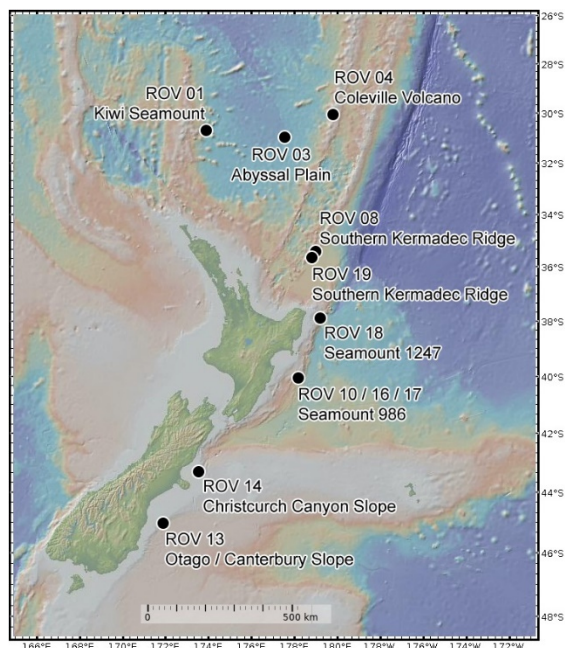


FIGURE 1 | Sampling site map from New Zealand. Dots indicate the collection site with ROV dive number and name of the underwater site.

which we addressed using two high-throughput 16S rRNA gene sequencing data libraries, we also applied quantitative PCR (qPCR) to a subset of the sponge specimens, because quantitative data are frequently missing in sponge microbiome studies.

Sponges were placed in coolers with ice packs upon removal from the bioboxes of the ROV. Briefly, the sponges were photographed and subsequently dissected (Supplementary Figure S1). Small tissue sub-samples were instantly frozen for molecular analyses, while larger tissue samples were stored in ethanol (80%) as vouchers for morphological identification. All samples were stored at -80°C until further processing (see section “Supplementary Material”). In addition, nine triplicate seawater samples were collected at several sponge sampling sites from the same depths using Niskin bottles attached to the ROV. Approximately 2000 ml of each seawater sample was filtered using PVDF filter membranes ($0.22\ \mu\text{m}$ pore size and $\varnothing\ 47\ \text{mm}$) and stored at -80°C until further processing. Sponge species were identified by morphological and molecular taxonomic methods (see section “Supplementary Material and **Table 1**”). Sponge tissue vouchers were stored in three collections, NIWA Invertebrate Collection in New Zealand, ICBM Environmental Chemistry Collection, University of Oldenburg, Germany and Ludwig-Maximilians-Universität, in Germany and are available on request (Supplementary Table S1).

Bacterial and Archaeal 16S rRNA Gene Amplicon Sequencing and Processing

Initially DNA of sponges (three to four replicates per taxon) and seawater samples was extracted using the DNeasy Power Soil Kit (Qiagen) on approximately 0.25 g of sponge tissue or half a seawater filter. After the quality and quantity of the extracts had been checked (by Nanodrop and gel electrophoresis after a PCR with universal 16S rRNA gene primers), a

TABLE 1 | Sponge sample list, including taxonomy from class to species, number of replicates, sampling sites, and sampling depths.

Class	Order	Family	Genus and species	Replicates	Sampling locations	Sampling depths
Demospongiae	Axinellida	Stelligeridae	<i>Paratimea</i> sp. indet.	4	Coleville Volcano	472–532
Demospongiae	Haplosclerida	Halichondriidae	<i>Halichondria</i> sp. indet.	3	Seamount 986	892–899
Demospongiae	Poecilosclerida	Latrunculiidae	<i>Latrunculia morrisoni</i>	3	Otago/Canterbury Slope, Christchurch Canyon Slope	595, 670–706
Demospongiae	Tetractinellida	Geodiidae	<i>Geodia vaubani</i>	4	Southern Kermadec Ridge, Seamount 986	1172–1216, 802
Demospongiae	Tetractinellida	Geodiidae	<i>Penares turmericolor</i>	4	Southern Kermadec Ridge	1187–1191
Demospongiae	Tetractinellida	Pleromidae	<i>Pleroma turbinatum</i>	3	Coleville Volcano	472–497
Hexactinellida	Lyssacosida	Euplectellidae	<i>Bolosoma cyanae</i>	4	Southern Kermadec Ridge	1149–1167
Hexactinellida	Lyssacosida	Euplectellidae	<i>Corbitella plagiariorum</i>	4	Seamount 986	770–802
Hexactinellida	Lyssacosida	Euplectellidae	<i>Regadrella okinoseana</i>	4	Seamount 986	774–896
Hexactinellida	Lyssacosida	Euplectellidae	<i>Saccocalyx tetractinus</i>	3	Seamount 1247, Abyssal Plain	1352–1457, 4160
Hexactinellida	Lyssacosida	Leucopsacidae	<i>Leucopsacus distantus</i>	3	Seamount 986	792–896
Hexactinellida	Lyssacosida	Rosellidae	<i>Lanuginellinae</i> gen. et sp. indet.	3	Seamount 986	802–893
Hexactinellida	Sceptrulophora	Aphrocallistidae	<i>Aphrocallistes beatrix</i>	3	Kiwi Seamount, Seamount 986	759–793

For a detailed list see *Supplementary Table S1*.

one-step PCR was performed for amplification of the bacterial V3 to V4 variable regions (primer pair 341F 5'-CCTACGGGAGGCAGCAG-3' and 806R 5'-GGACTACHVGGGTWTCTAAT-3') and archaeal V4 to V6 variable regions (primer pair Uni519F 5'-CAGCMGCCGCGGTAA-3' and 1000R 5'-GGCCATGCACYWCYTCTC-3') of the 16S rRNA gene (Øvreås et al., 1997; Gantner et al., 2011; Takahashi et al., 2014). A quality check by gel electrophoresis, normalization, and pooling was performed on the amplicon libraries before independent sequencing of the bacterial and archaeal libraries on a MiSeq platform (MiSeqFGx, Illumina) using the v3 chemistry. This resulted in one archaeal and one bacterial 16S rRNA gene amplicon libraries (for detailed amplicon libraries preparation methods see section “Supplementary Material”). The raw data have been deposited in the Sequence Read Archive with BioProject number: PRJNA552490 (bacterial libraries) and PRJNA552540 (archaeal libraries).

Amplicon sequences were processed using QIIME2-2018.11 (Boyle et al., 2018). Bacterial and archaeal libraries were processed and analyzed in parallel applying the same plugin commands if not stated otherwise. Due to quality reasons only the bacterial and archaeal forward amplicon libraries (i.e., single-end) were used in this study. After the import of demultiplexed single-end fastq files via *qiime*

import, primers were trimmed using *qiime cutadapt trim-single*. The QIIME2 plugin DADA2 (*qiime dada2 denoise-single*) was used for the detection and correction of Illumina-generated amplicon sequence data and to generate amplicon sequence variants (ASVs) using the following parameters for both libraries: *-p-trim-left 0* and *-p-trunc-len 250*. Resulting bacterial and archaeal 16S rRNA gene representative ASV sequences were used to calculate phylogenetic ASV trees for subsequent analyses using the *qiime phylogeny align-to-tree-mafft-fasttree* plugin. Bacterial and archaeal primer-specific trained Naive Bayes taxonomic classifiers (<https://docs.qiime2.org/2018.11/tutorials/feature-classifier/>), using the SILVA 132 release files (<https://www.arb-silva.de/download/archive/qiime>), were used to classify the representative ASV sequences (*qiime feature-classifier classify-sklearn*). Before subsequent analyses, the complete amplicon datasets, including all available Demospongiae, Hexactinellida, and seawater sample groups, were divided into additional sample group-specific datasets via the *qiime feature-table filter-samples* plugin.

The generated exact sequence variants, or amplicon sequence variants (i.e., ASVs), are used as substitute for the commonly used operational taxonomic units (OTUs) clusters of sequencing reads by applying the

QIIME2 implemented DADA2 algorithm (Callahan et al.,2016; Nearing et al.,2018). The common pooling of sequences into OTUs limited the possibilities of deep sequencing by preventing fine-scale resolution. Therefore, we chose to generate ASVs instead of OTUs to achieve state-of-the-art fine-scale community data and, in addition, to benefit from the error correction model applied by the DADA2 algorithm. In the following we will use the term *feature*, as introduced by QIIME2, when referring to the microbial ASVs (Boyle et al.,2018).

Quantitative Real-Time PCR (qPCR)

For quantification of the domain-level specific primers (eubacterial and archaeal 16S rRNA genes) we followed the protocol from Bayer et al. (2014). Briefly, 1:5 dilutions of purified PCR products were used as standards, and all standard dilutions were prepared in aqueous tRNA solution (10 ng/ml) (Sigma-Aldrich, Schnellendorf, Germany). The DNA concentration of the highest starting solution of each standard dilution series as well as the diluted template DNAs was measured using the Qubit system (double stranded DNA, high sense kit, Thermo Fisher Scientific, Darmstadt, Germany). Quantitative PCRs were performed in a CFX Connect realtime detection system (Bio-Rad, Munich, Germany) using the SsoAdvanced™ Universal SYBR® Green Supermix (Bio-Rad) following the manufacturer's instructions. For both primer pairs, the reaction conditions previously tested (see Bayer et al.,2014) were used with one exception: the annealing/elongation temperature for the archaeal 16S rDNA gene assay was increased to 66°C. For all qPCR assays, plate reads were taken at the end of each qPCR cycle. All template DNAs from sponges and seawater were tested in triplicates on each plate (technical replicates), whereas the corresponding standards were run in duplicates. The qPCR efficiency and gene copy numbers were calculated using the Bio-Rad CFX MANAGER™ software (version 3.1). Amplification of specific targets was confirmed by analyses of melt curves (in steps of 0.5°C for 5 s, with temperatures ranging from 60 to 95°C). Additionally, PCR

product sizes were checked by electrophoresis on a 1.5% agarose gel (Peqlab now VWR, part of avantor) in 1x TAE buffer with 0.5% GelGreen™ (Biotium, Hayward, CA, United States) for visualization.

Bacterial and Archaeal Community Analyses

Prokaryotic taxonomy tables from domain to species levels were created using the *qiime taxa barplot* plugin. A collective calculation of diversity metrics (both phylogenetic and non-phylogenetic) was applied on available datasets via the *qiime diversity core-metrics-phylogenetic* plugin, using the minimum sampling depth of each dataset for mandatory subsampling. The following alpha diversity indices were considered: Faith's phylogenetic diversity, observed features, Shannon diversity, and evenness. Visualization of beta diversity used the principal coordinates results, which were based on weighted uniFrac distances. Analyses of sample composition in the context of categorical metadata were performed with the *qiime diversity beta-group-significance* plugin (parameters: *permanova* and *permdisp*) again utilizing the weighted uniFrac distances. The *ggpubr* R package was used to calculate sample statistics and plot respective results (i.e., alpha diversity and qPCR bar charts, or PCoA plots; <https://cran.r-project.org/web/packages/ggpubr/>). Ternary plots were created using the Ternary plot maker web tool (<https://www.ternaryplot.com/>). Finally, taxonomy bar plots and bacterial and archaeal features heatmaps were created using a custom R script and the *ggplot2* package (Wickham,2016). For heatmaps the respective feature abundance tables were sub-sampled for sponge samples only in order to avoid the bias introduced by abundant seawater features. The respective seawater samples were added subsequently before heatmap creation. Bacterial and archaeal feature Venn diagrams were created using *mothur* v.1.42.3 (Schloss et al.,2009). The complete QIIME2 pipeline and R script collection from the present study can be accessed online (<https://github.com/marine-moleco/So254Qiime2ASVs>).

Archaeal Taxonomy Note

The applied SILVA 132 database places the family Nitrosopumilaceae within the phylum Thaumarchaeota (class Nitrososphaeria; order Nitrosopumilales). However, a recent prokaryotic taxonomical reassessment using full genomes reorganized the taxonomy of the family Nitrosopumilaceae considerably (i.e., phylum

Crenarchaeota; class Nitrososphaeria; order Nitrososphaerales) (Parks et al., 2018). In terms of reproducibility we decided to use the available SILVA 132 taxonomy without manual changes to certain taxa, such as the family Nitrosopumilaceae. However, the presented taxonomy might be a subject to change in upcoming SILVA releases.

RESULTS

Deep-Sea Sponge Taxonomy

Sponge taxonomic identifications were confirmed using a combination of gene markers, morphology, and spicule analyses. The 45 sponge specimens in this study belong to the two sponge classes Demospongiae and Hexactinellida. The Demospongiae group is composed of six taxa comprising the species *Geodia vaubani* Lévi and Lévi (1983), the novel species *Halichondria* sp. indet., *Latrunculia* sp. nov., the novel species *Paratimea* sp. indet., *Penares turmericolor* Sim-Smith and Kelly (2019), and *Pleroma turbinatum* Sollas (1888). The Hexactinellida group consists of seven taxa comprising the species *Aphrocallistes beatrix* Gray (1858), *Bolosoma cyanae* Tabachnick and Lévi (2004), *Corbitella plagiariorum* Reiswig and Kelly (2018), *Leucopsacus distantus* Tabachnick and Lévi (2004), *Regadrella okinoseana* Ijima (1896), *Saccocalyx tetractinus* Reiswig and Kelly (2018) and one novel species of Rossellidae that belongs to the Lanuginellinae subfamily (Table 1).

Bacterial and Archaeal Features (Amplicon Sequencing)

The sequencing of sponge-associated prokaryotic communities and additional seawater samples yielded 9120 bacterial and 1052 archaeal features in total, whereas the sequencing depth was relatively balanced with 1,960,245 bacterial and 2,090,246 archaeal sequence reads (Table 2). Excluding seawater samples, Hexactinellida exhibited the most

TABLE 2 | Detailed bacterial and archaeal MiSeq library statistics, comprising the sequence data for (a) the whole dataset (i.e., Demospongiae, Hexactinellida, and seawater samples), (b) the Demospongiae subset, (c) the Hexactinellida subset, and (d) the seawater subset.

Metric	Bacteria	Archaea
Number of samples	72	72
Number of features	9,102	1,052
Total frequency	1,960,245	2,090,246
Demospongiae samples	21	21
Demospongiae features	2,084	437
Demospongiae frequency	558,729	770,140
Hexactinellida samples	24	24
Hexactinellida features	4,357	264
Hexactinellida frequency	588,341	535,764
Seawater samples	27	27
Seawater features	4,025	791
Seawater frequency	813,175	784,342

bacterial features ($n = 4357$), compared to the Demospongiae ($n = 2084$) specimens. In contrast, the demosponge samples exhibited a higher archaeal feature count ($n = 437$), compared to the hexactinellids ($n = 264$). Finally, the archaeal seawater feature count ($n = 791$) was higher in comparison with the two sponge classes, whereas bacterial seawater features were almost similar to the hexactinellid counts ($n = 4025$) (Table 2).

Bacterial and Archaeal Taxonomy and Taxon Distribution

In total, 44 bacterial phyla and four archaeal phyla (i.e., Thaumarchaeota, Euryarchaeota, Nanoarchaeota, and Crenarchaeota) were present among all sponge and seawater samples. The demospunges *Paratimea* sp., *P. turmericolor*, *P. turbinatum*, and *G. vaubani*,

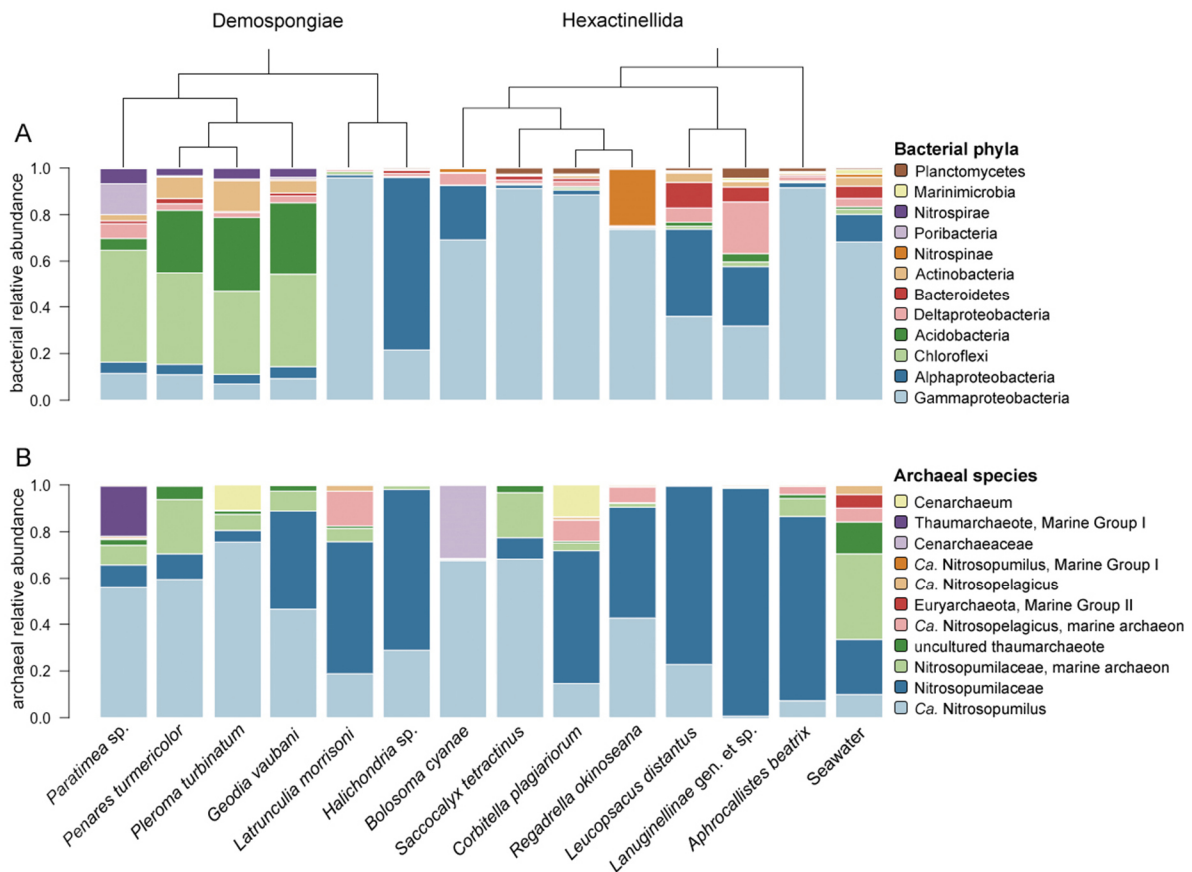


FIGURE 2 | Relative abundance of the most abundant bacterial phyla (A) and archaeal species (B) for all 13 sponges and seawater. The dendrogram is based on phylogenetic relationships between the sponge species. Samples are grouped by sponge class (Demospongiae, Hexactinellida).

contain Chloroflexi, Acidobacteria, Actinobacteria, and Nitrospirae as the most abundant bacterial phyla (Figure 2A). Surprisingly, the *Paratimea* sp. specimens possess a large fraction of poribacterial symbionts (Figure 2A). The Hexactinellida contain bacterial phyla, such as Proteobacteria (Delta-), Bacteroidetes, Nitrospinae, or Planctomycetes, which are relatively more abundant than in the investigated demosponge specimens (Figure 2A).

Due to the Thaumarchaeota dominance (92–100% relative abundance among all sample groups – see Supplementary Table S2) and the overall low archaeal feature richness (compared to the bacterial community – see Supplementary Figure S2), we considered the archaeal species composition as the relevant taxonomic level in our following community description throughout the present study. This is in contrast to the common phylum composition approach that has been used to describe sponge-associated bacterial communities. At

species level a total of 76 archaeal taxa, compared to 1774 bacterial taxa, could be designated to all features present in the sponge and seawater samples. Among those 76 archaeal taxa the subset comprising the most abundant archaea was composed of almost only thaumarchaeotes, except one euryarchaeotal taxon (Thermoplasmata, Marine Group II) (Figure 2B). In addition, all top thaumarchaeotes belong to the family Nitrosopumilaceae (Figure 2B), which could be further identified as several *Candidatus* taxa, such as Nitrosopumilus or Nitrosopelagicus.

Ternary plots were used to examine the distributions of the bacterial and archaeal most abundant taxa among the pooled demosponges, hexactinellids, and seawater samples. As indicated by the ternary plots, certain bacterial phyla are almost exclusive to Demospongiae; such as Acidobacteria, Chloroflexi, Dadabacteria, Nitrospirae, Spirochaetes, Entotheonellaeota, or Poribacteria (Figure 3A).

On the contrary, there is no evidence that the present hexactinellids possess an exclusive phylum. However, Patescibacteria and Nitrospinae seem to be more related to this sponge class as opposed to demosponges or the surrounding seawater. In addition, Marinimicrobia, Cyanobacteria, and Bacteroidetes exhibit a preference for the present seawater samples (Figure 3A). Finally, the group of several abundant proteobacterial classes, and Verrucomicrobia, show an almost even distribution among all three biotopes, however, with a clear tendency for the seawater samples.

On genus level there are some demosponge- (e.g., Caldilineaceae, EC94, SAR202 clade, and Poribacteria) or hexactinellid- (e.g., *Albimonas*, Betaproteobacteriales, Nitrosococcaceae, BD2-7, and Cm1-21) specific taxa that seem to be exclusive for one of the two sponge classes (Figure 3B). In addition, the seawater samples possess certain exclusive genera, such as *Alteromonas*, *Halomonas*, *Pseudoalteromonas*, or *Vibrio*. Similar to the bacterial genera ternary plot, the most abundant archaeal assemblages can be grouped into either Demospongiae-, Hexactinellida-, or seawater-specific taxa. Here, certain taxa are highly specific to one of the available main groups (Figure 3C). For instance, the single euryarchaeotal taxon (Marine Group II) is only present in the seawater samples, whereas several Nitrosopumilaceae-related taxa are highly characteristic for one of the two sponge classes. Moreover, all three abundant *Candidatus Nitrosopelagicus* taxa have a strong preference for the seawater samples (Figure 3C). On the contrary, the *Candidatus Nitrosopumilus* taxa show either a tendency for demosponges and seawater samples or are unique to the Hexactinellida. Finally, a large group composed of not further classified Nitrosopumilaceae features is present almost at the center between all main biotopes, hinting to an even distribution of further Nitrosopumilaceae-related features among the three biotopes (Figure 3C).

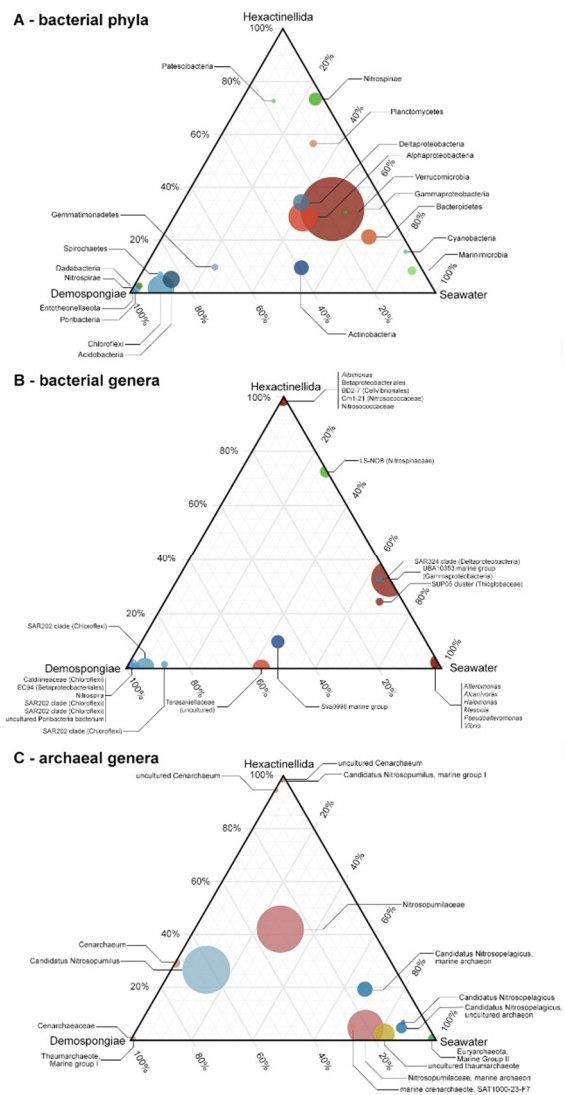


FIGURE 3 | Ternary plots showing the distribution of the most abundant taxa of bacterial phyla (A), bacterial genera (B), and archaeal genera (C) among the pooled Demospongiae, Hexactinellida, and seawater samples. Circle size is indicative of the relative abundance of these phyla.

Alpha Diversity Analysis

Rarefied abundance tables, generated using the individual sample read counts and feature assemblages, were used to calculate the mean bacterial and archaeal feature diversity at a local scale (i.e., alpha diversity indices for demosponges, hexactinellids, and seawater groups). Alpha diversity analyses of the bacterial communities revealed that all four indices (i.e., Faith's PD, observed features, Shannon diversity, and evenness) were significantly different within both sponge classes (Table 3).

TABLE 3 | Bacterial and Archaeal alpha diversity for Demospongiae, Hexactinellida, and seawater.

		Demospongiae		Hexactinellida		Seawater	
		H	<i>p</i> -value	H	<i>p</i> -value	H	<i>p</i> -value
Faith's PD	Bacteria	14.37	0.013	18.17	0.006	15.33	0.053
	Archaea	6.09	0.298	4.50	0.480	11.93	0.155
Features	Bacteria	18.07	0.003	18.14	0.006	14.10	0.079
	Archaea	5.43	0.365	4.00	0.677	15.90	0.044
Shannon	Bacteria	18.97	0.002	18.67	0.005	19.96	0.010
	Archaea	3.71	0.592	6.29	0.391	21.19	0.007
Evenness	Bacteria	18.48	0.002	19.11	0.004	19.15	0.014
	Archaea	4.22	0.518	6.57	0.363	17.40	0.026

Significant *p* values are highlighted in bold.

In contrast, the archaeal feature assemblages showed no significant alpha diversity differences between the species belonging to the two sponge classes. Seawater samples apparently deviated from that overall pattern by exhibiting almost no significant differences for the two present richness indices, except the observed archaeal features ($p = 0.044$), whereas Shannon diversity and evenness showed significant differences among the seawater samples for both prokaryotic domains (Table 3). In addition, the comparisons between the three biotopes showed significant differences among all available alpha diversity indices for bacteria and archaea (Supplementary Figure S2). However, significant *p*-values were lower for the archaeal features compared to the bacterial *p*-values. Moreover, Faith's PD and observed feature values for bacteria and archaea were highest within the seawater samples, followed by demosponges and hexactinellids (Supplementary Figure S2). The same pattern was visible in the bacterial and archaeal Shannon and evenness indices, where the spread was much higher in the sponge related samples compared to the richness indices (Supplementary Figure S2).

Beta Diversity Analysis

The same rarefied abundance tables that were used in the alpha diversity analyses were again utilized to look at several beta diversity aspects (i.e., community differences between samples). First, the bacterial and archaeal abundance and composition information was employed to investigate the differences between the Demospongiae, Hexactinellida, and

seawater samples (Figures 4A,B). Regarding the bacterial community, two main groups are visibly separated by the first axis (34% variance explained) (Figure 4A). The larger group is composed of all Hexactinellida, seawater, and a Demospongiae subset, whereas the second group exclusively consists of demosponges. However, the larger group also exhibits a relatively distinct separation between the seawater, Hexactinellida, and Demospongiae samples. In this group, the second axis separates the seawater- and Demospongiae subsets, while the Hexactinellida prevailing subsets are in-between these two distinct groups (16% variance explained). Nevertheless, a few overlaps between these biotopes exist (Figure 4A). Comparative statistics showed that the group identity has a significant effect on the community composition ($p < 0.001$), while the variances were homogeneous ($p = 0.083$) (Figure 4A and Table 4).

TABLE 4 | Community permanova and permdisp statistics, comprising the sequence data for (a) the whole dataset (i.e., Demospongiae, Hexactinellida, and seawater samples), (b) the Demospongiae subset, (c) the Hexactinellida subset, and (d) the seawater subset.

	Group	Sample size	Groups	Test statistic	<i>p</i> -value
Bacteria	Group permanova	72	3	15.27	0.001
	Group permdisp	72	3	2.53	0.083
	Demo permanova	21	6	40.99	0.001
	Demo permdisp	21	6	30.09	0.013
	Hexac permanova	24	7	7.74	0.001
	Hexac permdisp	24	7	4.50	0.002
	SW permanova	27	9	7.35	0.001
	SW permdisp	27	9	0.58	0.451
	Archaea	Group permanova	72	3	32.52
Group permdisp		72	3	0.09	0.895
Demo permanova		21	6	1.22	0.203
Demo permdisp		21	6	1.82	0.19
Hexac permanova		24	7	1.28	0.106
Hexac permdisp		24	7	0.56	0.496
SW permanova		27	9	20.99	0.001
SW permdisp		27	9	0.47	0.235

Significant *p* values are highlighted in bold.

In case of the archaeal feature composition and abundance, the first axis clearly separates the sponge samples from the seawater samples (55% variance explained) (Figure 4B). Furthermore, seawater samples are then subset by the second axis (16% variance explained) into one large and one small group. The smaller

group is comprised of seawater samples from ROV station 13 (Otago-Canterbury slope) and station 14 (Christchurch slope), which are both located on the continental slope of the South Island within the subtropical front separating the water masses of the South Pacific Subtropical Gyre to the north from the Subantarctic Water Ring to the south (Carter, 2001; **Figure 1** and Supplementary Figure S3, **Table 1**). In comparison, the sponge samples exhibit no clear separation into Demospongiae and Hexactinellida specimens, hence forming a heterogeneous group apart from the seawater samples (**Figure 4B**). However, overall the archaeal composition was still significantly affected by group identity ($p < 0.001$), further supported by homogeneously dispersed groups ($p = 0.895$) (**Figure 4B** and **Table 4**).

Finally, biotope-specific subsets (i.e., Demospongiae, Hexactinellida, and seawater) revealed that the demosponge and/or hexactinellid host-identity has a significant effect on the bacterial community composition ($p < 0.001$ for both sponge classes) (Supplementary Figures S3A,C and **Table 4**), whereas no significant effect of sponge identity could be observed for the archaeal data ($p = 0.203$ and $p = 0.106$ for demosponges and hexactinellids, respectively) (Supplementary Figures S3B,D and **Table 4**). In comparison, sampling location has a significant effect on the bacterial and archaeal community composition of the seawater samples (Supplementary Figures S3E,F and **Table 4**).

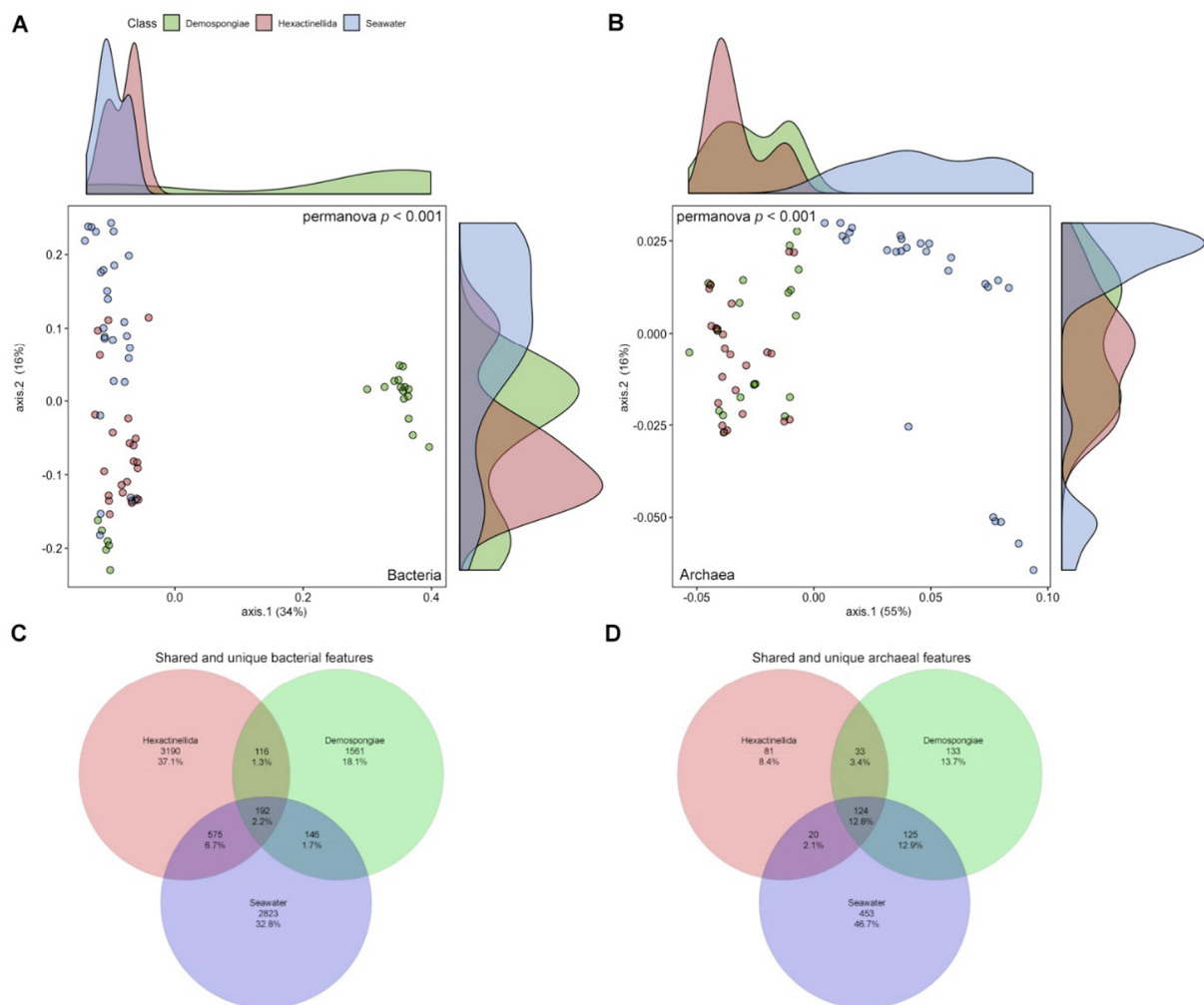


FIGURE 4 | Principal component analysis using bacterial (**A**) and archaeal (**B**) weighted uniFrac distances for Demospongiae, Hexactinellida and seawater samples. Venn diagrams depicting the shared and unique bacterial (**C**) and archaeal (**D**) features as percent for the three sampling groups (Hexactinellida, Demospongiae, seawater).

Bacterial and Archaeal Feature Distribution and Abundance

Venn diagrams visualized the shared and unique bacterial and archaeal features for the three sampling groups (Figures 4C,D). Overall, the percentage of shared bacterial features among all three groups is low and consequently each of the sample groups exhibits a large share of unique features (37.1, 18.1, and 32.8% for Hexactinellida, Demospongiae, and seawater, respectively) (Figure 4C). Similarly, for the archaeal features, the pooled seawater samples hold the largest share of the archaeal features (46.7%) compared to the demosponges

(13.7%) and hexactinellids (8.4%) (Figure 4D). Moreover, some shared features (i.e., 12.9% for seawater/Demospongiae, and 12.8% for seawater/Demospongiae/Hexactinellida) exceed the Hexactinellida-unique features. Despite the large fraction of seawater-unique features the amount of shared features between Hexactinellida and seawater is low (2.1%).

We plotted the most abundant bacterial and archaeal features within the sponge subset as separated relative abundance heatmaps (Figure 5).

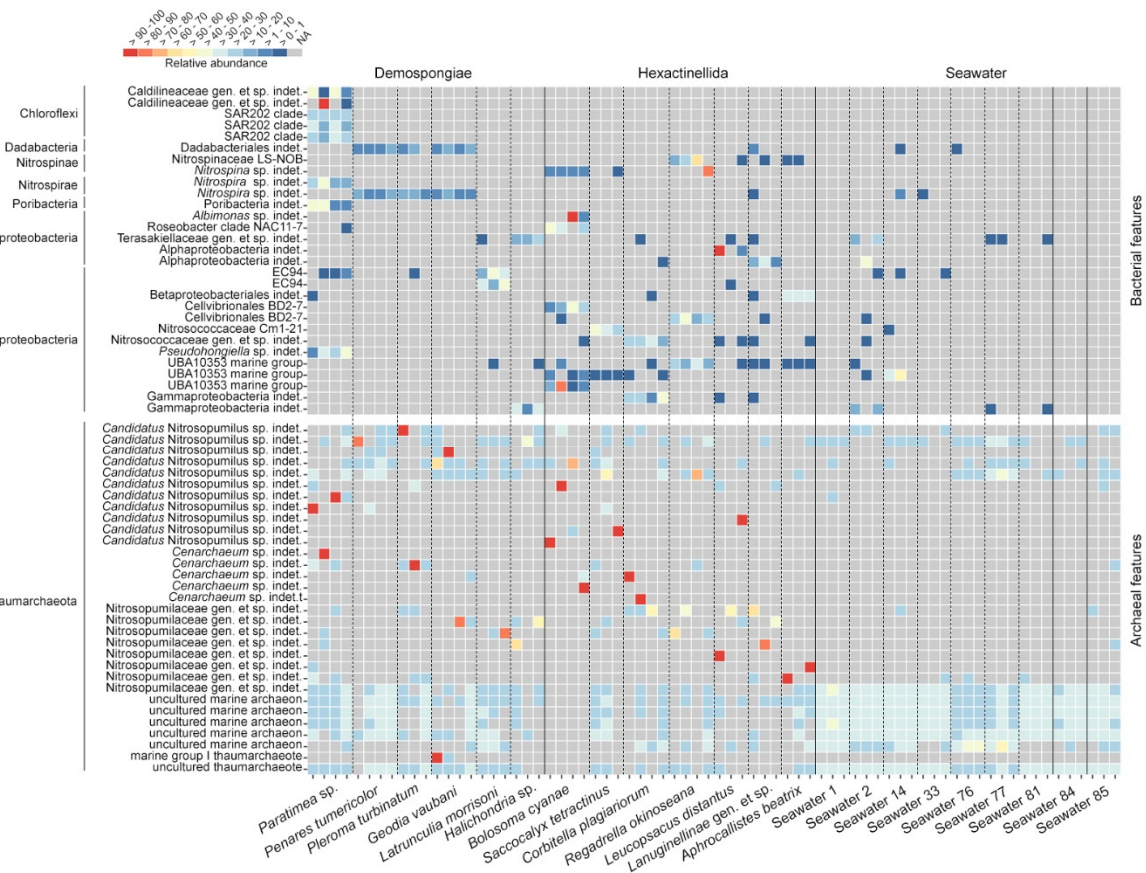


FIGURE 5 | Relative abundance heatmap of the investigated Demospongiae, Hexactinellida, and seawater. Heatmaps are separated between bacterial and archaeal features. Lowest available classification per feature has been used to assign individual taxonomy. Features were subsequently sorted by phylum.

The bacterial heatmap reveals that certain features are either highly host-species specific (e.g., Caldilineaceae gen. et sp. indet., SAR202 clade, *Nitrospira* sp., *Pseudohongiella* sp., *Nitrospina* sp., Roseobacter clade NAC11-7, EC94,

Cellvibrionales BD2- 7, or Nitrosococcaceae) or predominant in the respective sponge classes (Demospongiae: Dadabacteriales, *Nitrospira* sp.; Hexactinellida: UBA10353 marine group). Overall, the distribution among the

hexactinellid samples appears to be more scattered compared to the demosponge samples. Finally, seawater does not possess any of the sponge-specific bacterial features in large abundances and across all samples from the same location (Figure 5). The archaeal heatmap revealed that individual sponges possess single highly abundant features, which all belong to different Nitrosopumilaceae taxa (Figure 5). Contrary to the host-species specific distribution of the bacteria, sponge-associated archaeal symbionts are apparently individual-specific rather than host-specific. In contrast, seawater samples do not possess these singular highly abundant features. Instead, those samples rather exhibit location-specific feature distribution patterns (Figure 5). This corresponds with the archaeal ordination plots (Figure 4 and Supplementary Figure S3) and the respective archaeal community statistics (Table 4). Concerning the sponge-associated archaeal features, there seems to be a tendency that those singular abundant archaea predominantly belong to certain taxa (i.e., *Candidatus Nitrosopumilus* or *Cenarchaeum*).

Quantitative Assignment of Sponge-Associated Bacterial and Archaeal 16S rRNA Genes

For both qPCR assays the published reaction conditions (Bayer et al., 2014) were re-tested with respect to primer concentration, annealing, and elongation temperature and time. Both assays showed good *in silico* coverage and specificity and showed qPCR efficiencies between 95 and 105% with external standards according to MIQE guidelines (Bustin et al., 2009). One exception was that the annealing/elongation temperature for the archaeal assay was increased to 66°C for specificity reasons but without reaction efficiencies loss (Supplementary Table S3). All qPCR results are expressed as gene copy numbers per microgram genomic

DNA. The 16S rRNA gene copy numbers as estimated by qPCR with domain-specific primers for the investigated sponges are reported as follows (Supplementary Table S4). Bacterial 16S rRNA gene copy numbers ranged from $5.05 \times 10^{10} \pm 1.54 \times 10^{10}$ to $2.53 \times 10^{11} \pm 2.55 \times 10^{10}$ in hexactinellid sponges, to $5.62 \times 10^{10} \pm 4.79 \times 10^9$ to $1.81 \times 10^{11} \pm 3.46 \times 10^{10}$ in Demospongiae. Archaeal 16S rRNA gene copy numbers varied from $2.01 \times 10^9 \pm 1.23 \times 10^8$ to $2.71 \times 10^{11} \pm 2.57 \times 10^{10}$ in hexactinellid sponges, to $3.09 \times 10^8 \pm 1.29 \times 10^7$ to $3.24 \times 10^{10} \pm 1.73 \times 10^9$ in Demospongiae (Supplementary Figure S4 and Supplementary Table S4). As the fraction of host DNA within total extracted DNA potentially varies in every sample, we further calculated the ratios of bacteria and archaeal 16S rRNA gene copy numbers. We found that the prokaryotic consortia of hexactinellid sponges *B. cyanae* and *R. okinoseana* seem to have a high proportion of associated Archaea (BAC: ARCH ratios between 0.3 and 3.8), whereas *C. plagiariorum* deviates from that (BAC: ARCH ratios between 14.6 and 16.6). Demospongiae samples showed a higher proportion of Bacteria on average compared to Archaea (BAC: ARCH ratios between 4.0 and 21.9), with one apparent exception, which is *L. morrisoni* (BAC: ARCH ratios between 171.1 and 322.1) (Supplementary Table S4). Finally, the conversion of 16S rRNA gene copy numbers into BAC: ARCH ratios showed that Archaea, compared to the bacterial copy numbers, are significantly ($p = 0.026$) more abundant in hexactinellids than in demosponges (Figure 6A). However, a closer look at the individual sponge species revealed one exception from that initial observation (Figure 6B; *C. plagiariorum* with a BAC: ARC ratio between 14.6 and 16.6). A potential limitation of using the bacteria vs. archaeal ratio could be a possible bias in primer amplification, because of the use of two different 16S rRNA gene primer pairs.

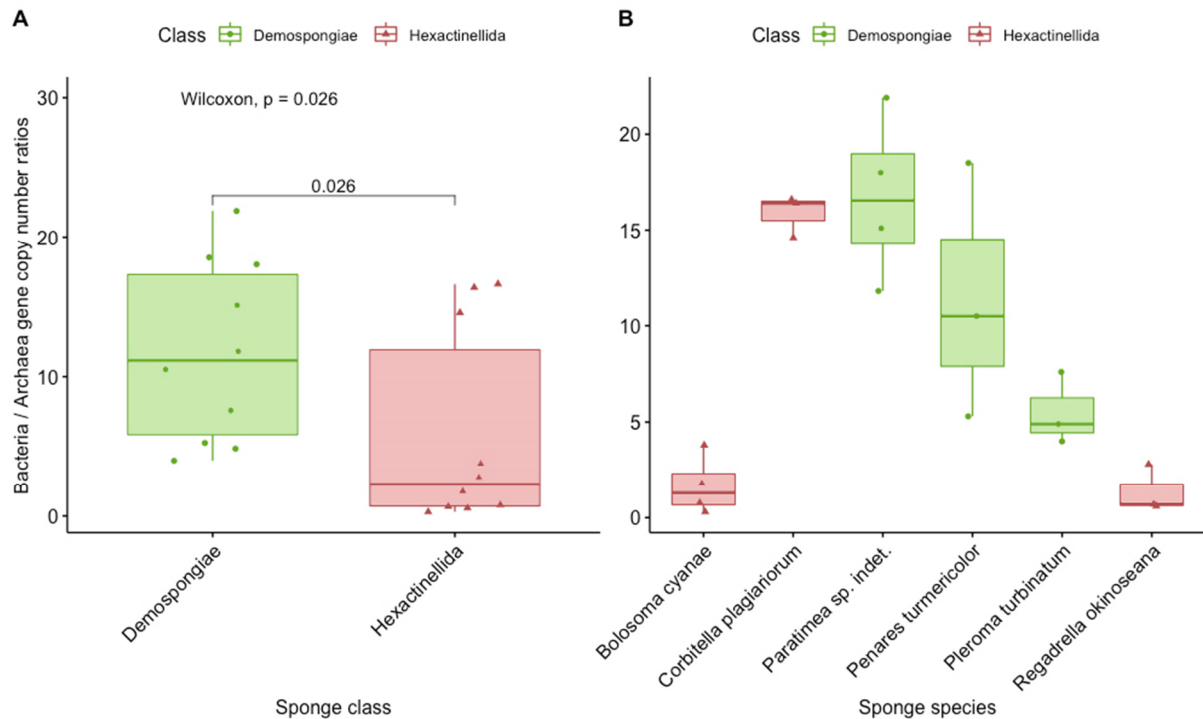


FIGURE 6 | Results of the qPCR experiments for bacterial and archaeal abundance presented as ratio of bacteria and archaea gene copy numbers. **(A)** Representing the overall average for Demospongiae and Hexactinellida. **(B)** Showing the ratio of bacteria and archaea gene copy numbers for three species of Demospongiae and Hexactinellida, respectively.

DISCUSSION

Bacterial Communities in Deep-Sea Demosponges and Hexactinellids

Since the beginning of sponge microbiome research (Vacelet, 1975; Vacelet and Donadey, 1977) the emphasis was placed on sponges from shallow-water marine habitats while sponges from cold-water and/or deep-sea locations are still understudied. Here, we explored the microbial community composition of demosponges and hexactinellids from the South Pacific by constructing independent bacterial and archaeal partial 16S rRNA gene libraries. While collection of sponges during the SO254 expedition included only deep-water specimens, analyses of microbiota patterns between deep-water and shallow-water sponges relied on literature data for shallow-water microbiota data (e.g., Thomas et al., 2016; Moitinho-Silva et al., 2017b).

The six demosponges showed a microbial signature similar to that of shallow-water demosponges with the phyla Proteobacteria

(Gamma-, Alpha-, and Delta-), Chloroflexi, Acidobacteria, Bacteroidetes, Nitrospinae, Nitrospirae, and Poribacteria being most abundant. One apparent difference to shallow-water sponges is the general lack or low abundance of members of the phylum Cyanobacteria, which are typically more abundant in shallow-water sponges (e.g., Erwin and Thacker, 2008; Bayer et al., 2014; Burgsdorf et al., 2014; Thomas et al., 2016). Given their involvement in photosynthesis it is not surprising that the present deep-sea sponges contain reduced numbers of this phototrophic bacterial phylum. The highly sponge-specific phylum Poribacteria appears to be underrepresented in sponges from deep-sea or cold habitats in the present and previous studies (Jackson et al., 2013; Rodríguez-Marconi et al., 2015; Cárdenas et al., 2018; Steinert et al., 2019). Noteworthy, we identified one sponge species (*Paratimea* sp.) with abundant poribacterial features. The overall phylum composition of this sponge matches the typical taxon composition

of high microbial abundance (HMA) sponges as defined by Moitinho-Silva et al. (2017b). Also, three more demosponges in this study (i.e., *P. turmericolor*, *P. turbinatum*, and *G. vaubani*) exhibited these HMA indicator taxa, although Poribacteria were lacking. The two remaining demosponge species (i.e., *Latrunculia* sp. nov. and *Halichondria* sp. indet.), display microbiome characteristics typical of LMA sponges (see **Figure 2**). In summary, with respect to bacterial community composition and HMA/LMA status, the demosponges from this remote deep-sea location largely resemble those of shallow-water collections (Thomas et al., 2016).

Regarding the seven hexactinellid species in the present study there is no literature available for comparisons with other hexactinellids from similar or different habitats. Overall, Gammaproteobacteria dominated the hexactinellid species. While Gammaproteobacteria are common predominant members in sponge bacterial communities, usually those demosponge-related communities exhibit additional dominant taxa regardless of climate zone or sampling depth (e.g., Kennedy et al., 2014; Thomas et al., 2016; Cárdenas et al., 2018; Steinert et al., 2019). Two species, i.e., Lanuginellinae gen. et sp. and *L. distantus*, also possess abundant Proteobacteria (Alpha- and Delta-), Bacteroidetes, and Chloroflexi, thus resembling a bacterial community pattern commonly found in demosponges.

When comparing the two sponge classes, evenness and diversity were indeed lower in hexactinellids compared to demosponges. Besides these alpha-diversity metrics, bacterial community composition and diversity differences are prominent throughout all present analyses. For instance, at least some demosponges possess specific phyla, like the HMA indicators, whereas hexactinellids resemble LMA sponges regarding their bacterial taxon composition. Secondly, demosponges and hexactinellids do not overlap in the ordination analyses, hence hexactinellids possess a rather class-specific microbiota. This is especially apparent when comparing the feature

distribution between LMA-like demosponges and hexactinellids. The Hexactinellida-associated features exhibit a more heterogeneous distribution. Hexactinellida are usually cold water/deep-sea sponges (van Soest et al., 2012). Such habitats imply lower food availability and therefore different metabolic functions in species of this sponge class. Moreover, glass sponges are clearly distinct to other sponge classes by body shape and features such as tissue and spicules (van Soest et al., 2012). We assume that these morphological differences also affect the microbiota composition of this class. Sponges are holobionts (Webster and Thomas, 2016); hence the symbiotic bacterial relationships are adapted to the Pita et al. (2018) host-ecosystem, which could be visible as discernable differences between demosponges and hexactinellid bacterial communities.

Demospongiae and Hexactinellida also share some common sponge-bacteria related characteristics, such as their noticeable difference to seawater samples (e.g., Lee et al., 2010; Taylor et al., 2013; Steinert et al., 2016; Cleary et al., 2018; Helber et al., 2019), which is present in the ordination plots or manifested in the alpha- and beta-diversity results. Especially apparent is the Hexactinellida-related host-specificity (see Supplementary Figure S3C), which is equally consistent in their demosponge counterparts (see Supplementary Figure S3A). Host-specificity is a common feature of the sponge-microbiota relationship, but so far only observed and described in depth for demosponges (Pita et al., 2013; Easson and Thacker, 2014; Thomas et al., 2016; Steinert et al., 2017, 2019). This pattern seems to be similar in hexactinellids, implying analogous bacterial community acquisition and maintenance processes as in their demosponge counterparts.

Archaeal Communities in Deep-Sea Demosponges and Hexactinellids

At high taxonomic archaeal ranks (i.e., from phylum to family level), both Demospongiae and Hexactinellida are dominated or even exclusively inhabited by the phylum Thaumarchaeota, and more specifically, several genera

and species from the family Nitrosopumilaceae. The ecologically important *Candidatus* family Nitrosopumilaceae forms a monophyletic group in the *Candidatus* order Nitrosopumilales based on 16S rRNA and *Candidatus amoA* (encoding for the α -subunit of ammonia monooxygenase) gene sequence analyses (Torre et al., 2016). The Nitrosopumilaceae comprises five genera, three of which were present in our sponges, namely: *Candidatus* Nitrosopelagicus, *Candidatus* Nitrosopumilus, and *Candidatus* Cenarchaeum. Nitrosopumilaceae grow chemolithoautotrophically by acquiring energy from ammonia oxidation and using CO₂ as carbon source. Additionally, some species can utilize urea as a source of ammonia for energy and growth (Torre et al., 2016). Marine demosponges are known hosts of symbiotic thaumarchaeotal members, but the understanding of their functional relationship is still lacking and mostly relies on circumstantial evidence (e.g., Kennedy et al., 2014; Feng et al., 2016, 2018; That et al., 2018; Moeller et al., 2019).

Different nitrogen cycling processes, such as nitrification, denitrification, and anaerobic ammonium oxidation have been observed in different demosponge species (e.g., Bayer et al., 2008; Hoffmann et al., 2009; Schläppy et al., 2010; Radax et al., 2012). Ammonia-oxidizing archaea (AOA) (i.e., thaumarchaeotes) are often abundant and diverse members of the sponge microbiota. AOA have even been detected in sponge larvae indicating vertical transmission and/or early environmental acquisition (Sharp et al., 2007; Steger et al., 2008; Schmitt et al., 2012). Transcription and translation of important functional genes, like the *amoA* gene of thaumarchaeotes, has been observed in different demosponge species (e.g., Bayer et al., 2008; Liu et al., 2012; Radax et al., 2012; Moitinho-Silva et al., 2017a). More complete genomic information about members of the phylum Thaumarchaeota, or more specifically the family Nitrosopumilaceae, has shed light on the metabolic potential and functional relationships of this putatively important

sponge-symbiotic archaeal group. So far, genomic information comprising different taxa within the family Nitrosopumilaceae is available from three demosponges (*Axinella mexicana*, *Cymbastela concentrica*, *lanthella basta*), and even one hexactinellid (*Lophophysema eversa*) (Hallam et al., 2006; Tian et al., 2016; Moitinho-Silva et al., 2017a; Moeller et al., 2019). Given the involvement of the Nitrosopumilaceae family in sponge-related nitrogen cycling processes, it can be assumed that thaumarchaeotal ammonia oxidation is a key functional process in deep-sea demosponges and hexactinellids alike.

Nitrosopumilaceae Are Sponge- but Not Sponge Species-Specific

We performed a high resolution screening of prokaryotic communities in demosponges and hexactinellids to look at the broad archaeal community in greater detail, and to directly compare community patterns and absolute abundances between bacterial and archaeal domains. Therefore, we constructed two independent bacterial and archaeal 16S rRNA gene libraries and quantitative real-time PCR in this study.

The difference between bacterial and archaeal communities is most apparent in sponge host-specificity, either at sponge class rank and/or at sponge species level (**Figures 4A,B** and Supplementary Figures S4A,B). Generally, bacterial demosponge-associated communities exhibit stable host-species specific patterns, which are affected by sponge-host identity and even sponge phylogeny (Easson and Thacker, 2014; Thomas et al., 2016; Steinert et al., 2017). The present archaeal communities do not show any distinct host-relatedness among all sponge species in both classes. Another difference between bacterial and archaeal community relationships is the feature distribution. Unique Nitrosopumilaceae features are randomly distributed among individual specimens and sponge taxa (**Figure 5**). In contrast, dominant sponge-associated bacteria often displayed sponge species-specific distribution patterns

that are generally explained by particular metabolically and functional host–symbiont relationships (Fan et al.,2012; Ribes et al.,2012; Thomas et al.,2016). The present archaeal feature distribution indicates that a random symbiont acquisition process is present among all sponges, which is independent of host phylogeny. Nevertheless, we assume that the latter processes are sponge-specific due to the evident differences between sponge- and seawater-associated archaeal communities (Figures 4B, 5).

In summary, Nitrosopumilaceae acquisition seems to happen randomly across all sponge samples. This implies functional redundancy of the symbiont’s key functional processes within the deep-sea sponge-associated genera of the family Nitrosopumilaceae. Functional redundancy, or evolutionary equivalence, has been hypothesized and observed also for sponge-associated bacteria, but usually at higher taxonomic ranks (Fan et al.,2012; Ribes et al.,2012; Thomas et al.,2016).

Bacterial and Archaeal Abundance

Finally, quantitative real-time PCR using both bacterial and archaeal universal 16S rRNA gene primers support the overall results. Although qPCR is a very sensitive method to assess microbial quantities, there are technical aspects that need to be considered. For example, the amplification of free DNA in sea water and sponge samples cannot be excluded and numbers might be overestimated. One explanation for the relatively high values could be the detection of dead material deposits from the upper water column in all three biotopes (i.e., Demospongiae, Hexactinellida, and seawater). As environmental microbes may encode for more than one 16S rRNA gene per genome, we express microbial abundances as gene copy number per microgram genomic DNA and compare it to literature (Bayer et al.,2014). Bacterial gene copy numbers were within the range found in shallow-water HMA demosponges (Bayer et al.,2014) and no significant difference was found comparing demosponges and hexactinellids from the deep

sea (this study). Archaeal gene copy numbers detected in both deep-sea demosponges and hexactinellids were up to three orders of magnitude higher than in shallow-water counterparts (see Bayer et al.,2014), highlighting the importance of archaeal symbionts for deep-sea sponges in general. Since cold water carries more CO₂ than warmer water (Garrison,2009) the physiological evidence for archaea using CO₂ as a carbon source (Wuchter et al.,2004; Könneke et al.,2014) might explain their higher abundance in deep-sea waters and sponges where ammonium for nitrification purposes might not be limited. Altogether, this suggests a major role of AOA in deep-sea sponge metabolism (in particular for the hexactinellid species), by providing additional metabolites via chemoautotrophy, similar to what diatoms or cyanobacteria contribute via photoautotrophy in shallow- water sponges (see Feng and Li,2019 and references within). Hence, future (meta)-omic’s studies should explore if deep-water sponge microbiomes also contain chemoautotrophic bacteria, which could provide further resources to the host sponges in addition to the AOA.

CONCLUSION

We investigated bacterial and archaeal communities from two sponge classes using independent 16S rRNA gene libraries and quantitative real-time PCR. With regard to bacteria, both demosponges and hexactinellids exhibit community characteristics similar to shallow-water sponges, including the presence of typical sponge symbiont taxa as well as host species-specific microbiomes for all sponge species investigated. In contrast, the archaeal community was taxonomically highly homogeneous and could only be resolved from the Nitrosopumilaceae family level on downward (i.e., from family to individual ASVs). However, the quantitative information hints at three orders higher archaeal gene copy numbers between shallow water and the present deep-water sponges. Hence, it is apparent that AOA are important members of deep-sea sponges in particular, and that different acquisition and

maintenance processes may be involved regarding the archaeal symbionts compared to the bacteria.

DATA AVAILABILITY STATEMENT

Sequences were deposited at NCBI as BioProjects with accession IDs PRJNA552490 and PRJNA552540.

AUTHOR CONTRIBUTIONS

PS, UH, KaB, and GS designed the experiments. GS, KaB, KrB, GW, PS, SK, and PA performed the experiments. GS, DE, KaB, KrB, UH, and PS analyzed the data. GS, KaB, KrB, PS, and UH wrote the manuscript. GS, GW, KaB, KrB, MD, SM, UH, and PS reviewed and edited the manuscript. MD identified hexactinellid specimens. MK identified demosponge specimens. SM provided initial identification and curation of taxonomic vouchers of sponges.

FUNDING

This study was funded by the European Union's Horizon 2020 Research and Innovation Program to UH under Grant Agreement No. 679849 ('SponGES') and by the DFG Collaborative Research Center CRC1182-TP B1 ("Metaorganisms") to UH. PS acknowledges funding by the Federal Ministry of Education and Research (BMBF) for the cruise SO254, grant O3G0254A, PORIBACNEWZ. GW acknowledges funding by LMU Munich's Institutional Strategy LMU excellent within the framework of the German Excellence Initiative.

REFERENCES

- Bayer, K., Kamke, J., and Hentschel, U. (2014). Quantification of bacterial and archaeal symbionts in high and low microbial abundance sponges using real-time PCR. *FEMS Microbiol. Ecol.* 89, 679–690. doi: 10.1111/1574-6941.12369
- Bayer, K., Schmitt, S., and Hentschel, U. (2008). Physiology, phylogeny and in situ evidence for bacterial and archaeal nitrifiers in the marine sponge *Aplysina aerophoba*. *Environ. Microbiol.* 10, 2942–2955. doi: 10.1111/j.1462-2920.2008.01582.x
- Borchert, E., Selvin, J., Kiran, S. G., Jackson, S. A., O'Gara, F., and Dobson, A. D. W. (2017). A novel cold active esterase from a deep sea sponge *Stelletta normani* metagenomic library. *Front. Mar. Sci.* 4:287. doi: 10.3389/fmars.2017.00287
- Boylan, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Gabriel, A., et al. (2018). QIIME 2: reproducible, interactive, scalable, and extensible microbiome data science. *PeerJ Prepr.* 37, 852–857. doi: 10.7287/peerj.preprints.27295v2

ACKNOWLEDGMENTS

This work is dedicated to Hans Tore Rapp, coordinator of the H2020-SponGES project, mentor and friend. We thank Andrea Hethke, Ina Clefsen, and the CRC1182 Z3 team (Katja Cloppenburg-Schmidt, Malte Rühlemann, John Baines) for assistance with the amplicon pipeline. We greatly acknowledge the crew and scientific party of RV *Sonne* cruise SO254, as well as the ROV *Kiel 6000* team for their valuable support at sea. We also thank Sven Rohde, Tessa Clemens and the entire benthic invertebrate team of the RV *Sonne* Cruise SO254 for their assistance in sample collection, processing and cataloging. We thank Henry Reiswig for advice on identification of hexactinellid samples. Sample collection was carried out under the "Application for consent to conduct marine scientific research in areas under national jurisdiction of New Zealand (dated 7.6.2016)." This is publication 68 of Senckenberg am Meer Metabarcoding and Molecular Laboratory. We also thank the reviewers for their comments, which helped to improve this manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2020.00716/full#supplementary-material>.

- Burgsdorf, I., Erwin, P. M., López-Legentil, S., Cerrano, C., Haber, M., Frenk, S., et al. (2014). Biogeography rather than association with cyanobacteria structures symbiotic microbial communities in the marine sponge *Petrosia ficiformis*. *Front. Microbiol.* 5:529. doi: 10.3389/fmicb.2014.00529
- Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., et al. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin. Chem.* 622, 611–622. doi: 10.1373/clinchem.2008.112797
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: high-resolution sample inference from *Illumina* amplicon data. *Nat. Methods* 13, 581–583. doi: 10.1038/nmeth.3869

- Cárdenas, C. A., Font, A., Steinert, G., Rondon, R., and González-Aravena, M. (2019). Temporal stability of bacterial communities in antarctic sponges. *Front. Microbiol.* 10:2699. doi: 10.3389/fmicb.2019.02699
- Cárdenas, C. A., González-aravena, M., and Font, A. (2018). High similarity in the microbiota of cold-water sponges of the genus *Mycale* from two different geographical areas. *PeerJ* 6:e4935. doi: 10.7717/peerj.4935
- Carter, L. (2001). Lessons of the past - currents of change: the ocean flow in a changing world. *Water Atmos.* 4, 15–17.
- Cleary, D. F. R., Polónia, A. R. M., and de Voogd, N. J. (2018). Prokaryote composition and predicted metagenomic content of two *Cinachyrella* morphospecies and water from west papuan marine lakes. *FEMS Microbiol. Ecol.* 94: fix175. doi: 10.1093/femsec/fix175
- de Goeij, J. M., Lesser, M. P., and Pawlik, J. R. (2017). "Nutrient fluxes and ecological functions of coral reef sponges in a changing ocean," in *Climate Change, Ocean Acidification and Sponges*, eds J. L. Carballo and J. J. Bell (Berlin: Springer), 373–410. doi: 10.1007/978-3-319-59008-0_8
- Esson, C. G., and Thacker, R. W. (2014). Phylogenetic signal in the community structure of host-specific microbiomes of tropical marine sponges. *Front. Microbiol.* 5:532. doi: 10.3389/fmicb.2014.00532
- Erwin, P. M., and Thacker, R. W. (2008). Phototrophic nutrition and symbiont diversity of two caribbean sponge-cyanobacteria symbioses. *Mar. Ecol. Prog. Ser.* 362, 139–147. doi: 10.3354/meps07464
- Fan, L., Reynolds, D., Liu, M., Stark, M., Kjelleberg, S., Webster, N. S., et al. (2012). Functional equivalence and evolutionary convergence in complex communities of microbial sponge symbionts. *Proc. Natl. Acad. Sci. U.S.A.* 109, E1878–E1887. doi: 10.1073/pnas.1203287109
- Feng, G., and Li, Z. (2019). "Carbon and nitrogen metabolism of sponge microbiome," in *Symbiotic Microbiomes of Coral Reefs Sponges and Corals*, ed. Z. Li (Dordrecht: Springer), doi: 10.1007/978-94-024-1612-1_9
- Feng, G., Sun, W., Zhang, F., Karthik, L., and Li, Z. (2016). Inhabitancy of active nitrosopumilus-like ammonia-oxidizing archaea and *Nitrospira nitrite-oxidizing* bacteria in the sponge *Theonella swinhoei*. *Sci. Rep.* 6:24966. doi: 10.1038/srep24966
- Feng, G., Sun, W., Zhang, F., Orlic, S., and Li, Z. (2018). Functional transcripts indicate phylogenetically diverse active ammonia-scavenging microbiota in sympatric sponges. *Mar. Biotechnol.* 20, 131–143. doi: 10.1007/s10126-018-9797-9795
- Gantner, S., Andersson, A. F., Alonso-Sáez, L., and Bertilsson, S. (2011). Novel primers for 16S rRNA-based archaeal community analyses in environmental samples. *J. Microbiol. Methods* 84, 12–18. doi: 10.1016/j.mimet.2010.10.001
- Garrison, T. S. (2009). *Oceanography: An Invitation to Marine Science*. Boston, MA: Cengage Learning.
- Gray, J. E. (1858). On aphrocalistes, a new genus of spongiadae from malacca. *Proc. Zool. Soc. Lond.* 26, 114–115. doi: 10.1111/j.1469-7998.1858.tb06352.x
- Hallam, S. J., Konstantinidis, K. T., Putnam, N., Schleper, C., Watanabe, Y., Sugahara, J., et al. (2006). Genomic analysis of the uncultivated marine crenarchaeote *Cenarchaeum symbiosum*. *Proc. Natl. Acad. Sci. U.S.A.* 103, 18296–18301. doi: 10.1073/pnas.0608549103
- Helber, S. B., Steinert, G., Wu, Y.-C., Rohde, S., Hentschel, U., Muhandó, C. A., et al. (2019). Sponges from Zanzibar host diverse prokaryotic communities with potential for natural product synthesis. *FEMS Microbiol. Ecol.* 95:fiz026. doi: 10.1093/femsec/fiz026
- Hoffmann, F., Radax, R., Woebken, D., Holtappels, M., Lavik, G., Rapp, H. T., et al. (2009). Complex nitrogen cycling in the sponge *Geodia barretti*. *Environ. Microbiol.* 11, 2228–2243. doi: 10.1111/j.1462-2920.2009.01944.x
- Ijima, I. (1896). Notice of new hexactinellida from sagami bay. *Zool. Anzeiger* 19, 249–254.
- Jackson, S. A., Flemer, B., McCann, A., Kennedy, J., Morrissey, J. P., O'Gara, F., et al. (2013). Archaea appear to dominate the microbiome of *Inflatella pellicula* deep sea sponges. *PLoS One* 8:84438. doi: 10.1371/journal.pone.0084438
- Kennedy, J., Flemer, B., Jackson, S. A., Morrissey, J. P., O'Gara, F., and Dobson, A. D. W. (2014). Evidence of a putative deep sea specific microbiome in marine sponges. *PLoS One* 9:e091092. doi: 10.1371/journal.pone.0091092
- Könneke, M., Schubert, D. M., Brown, P. C., Hügler, M., Standfest, S., Schwander, T., et al. (2014). Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for CO₂ fixation. *Proc. Natl. Acad. Sci. U.S.A.* 111, 8239–8244. doi: 10.1073/pnas.1402028111
- Lee, O. O., Wang, Y., Yang, J., Lafi, F. F., Al-Suwailem, A., and Qian, P.-Y. (2010). Pyrosequencing reveals highly diverse and species-specific microbial communities in sponges from the Red Sea. *ISME J.* 5, 650–664. doi: 10.1038/ismej.2010.165
- Lévi, C., and Lévi, P. (1983). Epunges tétractinellides et lithistides bathyales de nouvelle-calédonie. *Bull. Mus. Natl. Nat.* 5, 101–168.
- Li, Z.-Y., Wang, Y.-Z., He, L.-M., and Zheng, H.-J. (2014). Metabolic profiles of prokaryotic and eukaryotic communities in deep-sea sponge *Lamellomorpha* sp. indicated by metagenomics. *Sci. Rep.* 4, 1–11. doi: 10.1038/srep03895
- Liu, M., Fan, L., Zhong, L., Kjelleberg, S., and Thomas, T. (2012). Metaproteogenomic analysis of a community of sponge symbionts. *ISME J.* 6, 1515–1525. doi: 10.1038/ismej.2012.1
- Lurgi, M., Thomas, T., Wemheuer, B., Webster, N. S., and Montoya, J. M. (2019). Modularity and predicted functions of the global sponge-microbiome network. *Nat. Commun.* 10:992. doi: 10.1038/s41467-019-08925-8924
- Moeller, F. U., Webster, N. S., Herbold, C. W., Behnam, F., Domman, D., Albertsen, M., et al. (2019). Characterization of a thaumarcahal symbiont that drives incomplete nitrification in the tropical sponge *Ianthella basta* originality- significance statement. *bioRxiv* [Preprint], doi: 10.1101/527234
- Moitinho-Silva, L., Bayer, K., Cannistraci, C. V., Giles, E. C., Ryu, T., Seridi, L., et al. (2014). Specificity and transcriptional activity of microbiota associated with low and high microbial abundance sponges from the Red Sea. *Mol. Ecol.* 23, 1348–1363. doi: 10.1111/mec.12365
- Moitinho-Silva, L., Díez-Vives, C., Batani, G., Esteves, A. I., Jahn, M. T., and Thomas, T. (2017a). Integrated metabolism in sponge-microbe symbiosis revealed by genome-centered metatranscriptomics. *ISME J.* 11, 1651–1666. doi: 10.1038/ismej.2017.25
- Moitinho-Silva, L., Nielsen, S., Amir, A., Gonzalez, A., Ackermann, G. L., Cerrano, C., et al. (2017b). The sponge microbiome project. *Gigascience* 6, 1–7. doi: 10.1093/gigascience/gix077
- Moitinho-Silva, L., Steinert, G., Nielsen, S., Hardoim, C. C. P., Wu, Y.-C., McCormack, G. P., et al. (2017c). Predicting the HMA-LMA status in marine sponges by machine learning. *Front. Microbiol.* 8:752. doi: 10.3389/fmicb.2017.00752
- Naim, M. A., Morillo, J. A., Sørensen, S. J., Waleed, A. A. S., Smidt, H., and Siphema, D. (2014). Host-specific microbial communities in three sympatric North Seas sponges. *FEMS Microbiol. Ecol.* 90, 390–403. doi: 10.1111/1574-6941.12400
- Nearing, J. T., Douglas, G. M., Comeau, A. M., and Langille, M. G. I. (2018). Denoising the Denoisers: an independent evaluation of microbiome sequence error-correction approaches. *PeerJ* 6:e5364. doi: 10.7717/peerj.5364
- Øvreås, L., Forney, L., Daae, F. L., and Torsvik, V. (1997). Distribution of bacteroplankton in meromictic lake Saelenvannet, as determined by denaturing gradient gel electrophoresis of PCR-amplified gene fragments coding for 16S rRNA. *Appl. Environ. Microbiol.* 63, 3367–3373. doi: 10.1128/aem.63.9.3367-3373.1997
- Pape, T., Hoffmann, F., Quéric, N. V., Von Juterzenka, K., Reitner, J., and Michaelis, W. (2006). Dense populations of Archaea associated with the demosponge *Tentorium semisuberites* schmidt, 1870 from Arctic deep-waters. *Pol. Biol.* 29, 662–667. doi: 10.1007/s00300-005-0103-104
- Parks, D. H., Chuvochina, M., Waite, D. W., Rinke, C., Skarshewski, A., Chaumeil, P. A., et al. (2018). A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat. Biotechnol.* 36:996. doi: 10.1038/nbt.4229
- Pita, L., Rix, L., Slaby, B. M., Franke, A., and Hentschel, U. (2018). The sponge holobiont in a changing ocean: from microbes to ecosystems. *Microbiome* 6:46. doi: 10.1186/s40168-018-0428-421

- Pita, L., Turon, X., López-Legentil, S., and Erwin, P. M. (2013). Host rules: spatial stability of bacterial communities associated with marine sponges (*ircinia* spp.) in the western mediterranean Sea. *FEMS Microbiol. Ecol.* 86, 268–276. doi: 10.1111/1574-6941.12159
- Radax, R., Hoffmann, F., Rapp, H. T., Leininger, S., and Schleper, C. (2012). Ammonia-oxidizing archaea as main drivers of nitrification in cold-water sponges. *Environ. Microbiol.* 14, 909–923. doi: 10.1111/j.1462-2920.2011.02661.x
- Reiswig, H. M., and Kelly, M. (2018). The marine fauna of New Zealand. Euplectellid glass sponges (Hexactinellida, Lyssacosida, Euplectellidae). *NIWA Biodiver. Mem.* 130, 1–170.
- Reveillaud, J., Maignien, L., Eren, M., Huber, J. A., Apprill, A., Sogin, M. L., et al. (2014). Host-specificity among abundant and rare taxa in the sponge microbiome. *ISME J.* 8, 1198–1209. doi: 10.1038/ismej.2013.227
- Ribes, M., Jiménez, E., Yahel, G., López-Sendino, P., Díez, B., Masana, R., et al. (2012). Functional convergence of microbes associated with temperate marine sponges. *Environ. Microbiol.* 14, 1224–1239. doi: 10.1111/j.1462-2920.2012.02701.x
- Rodríguez-Marconi, S., De la Iglesia, R., Díez, B., Fonseca, C. A., Hajdu, E., and Trefault, N. (2015). Characterization of bacterial, archaeal and eukaryote symbionts from antarctic sponges reveals a high diversity at a three-domain level and a particular signature for this ecosystem. *PLoS One* 10:e0138837. doi: 10.1371/journal.pone.0138837
- Schläppy, M. L., Schöttner, S. I., Lavik, G., Kuypers, M. M. M., de Beer, D., and Hoffmann, F. (2010). Evidence of nitrification and denitrification in high and low microbial abundance sponges. *Mar. Biol.* 157, 593–602. doi: 10.1007/s00227-009-1344-1345
- Schloss, P. D., Westcott, S. L., Ryabin, T., Hall, J. R., Hartmann, M., Hollister, E. B., et al. (2009). Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl. Environ. Microbiol.* 75, 7537–7541. doi: 10.1128/AEM.01541-1549
- Schmitt, S., Tsai, P., Bell, J., Fromont, J., Ilan, M., Lindquist, N., et al. (2012). Assessing the complex sponge microbiota: core, variable and species-specific bacterial communities in marine sponges. *ISME J.* 6, 564–576. doi: 10.1038/ismej.2011.116
- Sharp, K. H., Eam, B., John Faulkner, D., and Haygood, M. G. (2007). Vertical transmission of diverse microbes in the tropical sponge *Corticium* sp. *Appl. Environ. Microbiol.* 73, 622–629. doi: 10.1128/AEM.01493-1496
- Sim-Smith, C., and Kelly, M. (2019). Review of the sponge genus *penares* (demospongiae, tetractinellida, astrophorina) in the New Zealand EEZ, with descriptions of new species. *Zootaxa* 4638, 1–56. doi: 10.11646/zootaxa.4638.1.1
- Slaby, B. M., Franke, A., Rix, L., Pita, L., Bayer, K., Jahn, M. T., et al. (2019). “Marine sponge holobionts in health and disease,” in *Symbiotic Microbiomes of Coral Reefs Sponges and Corals*, ed. Z. Li (Dordrecht: Springer), 81–104. doi: 10.1007/978-94-024-1612-1
- Sollas, W. J. (1888). Report on the *Tetractinellida* collected by H.M.S. challenger, during the years 1873-1876. report on the scientific results of the voyage of H.M.S. challenger during the years 1873–76. *Zoology* 25, 1–44.
- Steger, D., Ettinger-Epstein, P., Whalan, S., Hentschel, U., De Nys, R., Wagner, M., et al. (2008). Diversity and mode of transmission of ammonia-oxidizing archaea in marine sponges. *Environ. Microbiol.* 10, 1087–1094. doi: 10.1111/j.1462-2920.2007.01515.x
- Steinert, G., Rohde, S., Janussen, D., Blaurock, C., and Schupp, P. J. (2017). Host-specific assembly of sponge-associated prokaryotes at high taxonomic ranks. *Sci. Rep.* 7:2542. doi: 10.1038/s41598-017-02656-2656
- Steinert, G., Taylor, M. W., Deines, P., Simister, R. L., de Voogd, N. J., Hoggard, M., et al. (2016). In four shallow and mesophotic tropical reef sponges from guam the microbial community largely depends on host identity. *PeerJ* 4:e1936. doi: 10.7717/peerj.1936
- Steinert, G., Wemheuer, B., Janussen, D., Erpenbeck, D., Daniel, R., Simon, M., et al. (2019). Prokaryotic diversity and community patterns in antarctic continental shelf sponges. *Science* 6, 1–15. doi: 10.3389/fmars.2019.00297
- Tabachnick, K. R., and Lévi, C. (2004). “Lyssacosida du Pacifique sud-ouest (Porifera: Hexactinellida),” in *Tropical Deep-Sea Benthos 23. Mémoires du Muséum National d’Histoire Naturelle*, Vol. 191, eds B. A. Marshall and B. Richer de Forges (Dordrecht: Springer), 11–71.
- Takahashi, S., Tomita, J., Nishioka, K., Hisada, T., and Nishijima, M. (2014). Development of a prokaryotic universal primer for simultaneous analysis of bacteria and archaea using next-generation sequencing. *PLoS One* 9:e105592. doi: 10.1371/journal.pone.0105592
- Taylor, M. W., Tsai, P., Simister, R. L., Deines, P., Botte, E., Ericson, G., et al. (2013). ‘Sponge-specific’ bacteria are widespread (but rare) in diverse marine environments. *ISME J.* 7, 438–443. doi: 10.1038/ismej.2012.111
- That, T., Dat, H., Steinert, G., Thi, N., Cuc, K., Smidt, H., et al. (2018). Archaeal and bacterial diversity and community composition from 18 phylogenetically divergent sponge species in Vietnam. *PeerJ* 6:e4970. doi: 10.7717/peerj.4970
- Thomas, T., Moitinho-Silva, L., Lurgi, M., Björk, J. R., Easson, C., Astudillo-García, C., et al. (2016). Diversity, structure and convergent evolution of the global sponge microbiome. *Nat. Commun.* 7:11870. doi: 10.1038/ncomms11870
- Tian, R. M., Sun, J., Cai, L., Zhang, W. P., Zhou, G. W., Qiu, J. W., et al. (2016). The deep-sea glass sponge *Lophophysema eversa* harbours potential symbionts responsible for the nutrient conversions of carbon, nitrogen and sulfur. *Environ. Microbiol.* 18, 2481–2494. doi: 10.1111/1462-2920.13161
- Torre, D., Forest, L., Wa, P., Kobelt, J. N., Wa, S., Stahl, D. A., et al. (2016). *Candidatus Nitrosopumilaceae*. Hoboken, NJ: Wiley. doi: 10.1002/9781118960608.fbm00262
- Vacelet, J. (1975). Etude en microscopie électronique de l’association entre bactéries et spongiaires du genre *Verongia* (Dictyoceratida). *J. Microsc. Biol. Cell* 23, 271–288.
- Vacelet, J., and Donadey, C. (1977). Electron microscope study of the association between some sponges and bacteria. *J. Exp. Mar. Biol. Ecol.* 30, 301–314. doi: 10.1016/0022-0981(77)90038-7
- van Soest, R. W. M., Boury-Esnault, N., Vacelet, J., Dohrmann, M., Erpenbeck, D., de Voogd, N. J., et al. (2012). Global diversity of sponges (Porifera). *PLoS One* 7:35105. doi: 10.1371/journal.pone.0035105
- Vogel, S. (1977). Current-induced flow through living sponges in nature. *Proc. Natl. Acad. Sci. U.S.A.* 74, 2069–2071. doi: 10.1073/pnas.74.5.2069
- Webster, N. S., Negri, A. P., Munro, M. M. H. G., and Battershill, C. N. (2004). Diverse microbial communities inhabit Antarctic sponges. *Environ. Microbiol.* 6, 288–300. doi: 10.1111/j.1462-2920.2004.00570.x
- Webster, N. S., Taylor, M. W., Behnam, F., Lückner, S., Rattei, T., Whalan, S., et al. (2010). Deep sequencing reveals exceptional diversity and modes of transmission for bacterial sponge symbionts. *Environ. Microbiol.* 12, 2070–2082. doi: 10.1111/j.1462-2920.2009.02065.x
- Webster, N. S., and Thomas, T. (2016). The sponge hologenome. *mBio* 7:e00135-16. doi: 10.1128/mBio.00135-116
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York: Springer-Verlag.
- Wuchter, C., Schouten, S., and Coolen, M. J. L. (2004). Temperature-dependent variation in the distribution of tetraether membrane lipids of marine crenarchaeota: implications for TEX86 paleothermometry. *Paleoceanography* 19:A4028. doi: 10.1029/2004PA001041
- Xin, Y., Kanagasabhapathy, M., Janussen, D., Xue, S., and Zhang, W. (2011). Phylogenetic diversity of gram-positive bacteria cultured from antarctic deep-sea sponges. *Pol. Biol.* 34, 1501–1512. doi: 10.1007/s00300-011-1009-y

Chapter 2

***Chloroflexi* dominate the deep-sea golf ball sponges *Craniella zetlandica* and *Craniella infrequens* throughout different life stages**

© 2020 Busch et al.

Original publication by *Frontiers in Marine Science* (Frontiers)

Reprinted under the terms of the Creative Commons Attribution License (CC-BY 4.0).

doi: 10.3389/fmars.2020.00674.

Chloroflexi dominate the deep-sea golf ball sponges *Craniella zetlandica* and *Craniella infrequens* throughout different life stages

Kathrin Busch¹, Erik Wurz², Hans Tore Rapp^{3†}, Kristina Bayer¹, Andre Franke⁴ and Ute Hentschel^{1,5*}

¹GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany, ² Marine Animal Ecology Group, Wageningen University and Research, Wageningen, Netherlands, ³ Department of Biological Sciences, K.G. Jebsen Centre for Deep Sea Research, University of Bergen, Bergen, Norway, ⁴Institute of Clinical Molecular Biology (IKMB), Kiel, Germany, ⁵ Faculty of Mathematics and Natural Sciences, Christian-Albrechts-Universität zu Kiel, Kiel, Germany. † Deceased March 07, 2020. *Correspondence author.

Deep-sea sponge grounds are underexplored ecosystems that provide numerous goods and services to the functioning of the deep-sea. This study assessed the prokaryotic diversity in embryos, recruits, and adults of *Craniella zetlandica* and *Craniella infrequens*, common and abundant representatives of deep-sea sponge grounds in the North Atlantic. Our results reveal that symbiont transmission in the two *Craniella* sponge species likely occurs vertically, as highly similar microbial consortia have been identified in adults, embryos, and recruits. Moreover, transmission electron microscopy revealed high abundances of sponge-associated microorganisms, among which *Chloroflexi* (SAR202) were identified as common representatives by amplicon sequencing and fluorescence *in situ* hybridization (FISH). Equal diversity metrics, a similar overall prokaryotic community composition and a distinct dominance of the phylum *Chloroflexi* within all life stages are the key findings of our analyses. Information such as presented here provide understanding on the recruitment of deep-sea sponge holobionts which is needed to develop integrated management tools of such vulnerable marine ecosystems.

Keywords: vulnerable marine ecosystems, sponges, *Craniella*, *Chloroflexi*, symbiosis, early life stages, amplicon sequencing, fluorescence *in situ* hybridization

INTRODUCTION

A fundamental process influencing the fitness of many organisms is the establishment of host-microbial colonization during development (Mortzfeld et al.,2016). Microbial consortia enable the respective animal host to expand its range of nutrition (Nicholson et al.,2012), defense (Flórez et al.,2015), immunity (Eberl,2010), and development (Koropatnick et al.,2004). Host-microbe interactions may be regulated from early developmental stages and transmitted through generations (reviewed in Rosenberg and Zilber-Rosenberg,2011; McFall-

Ngai et al.,2013). In general, beneficial microbial symbionts are expected to be transmitted vertically (from parents to offspring) in order to assure an establishment of a compatible partnership (Björk et al.,2019 and references therein). Further, probability of vertical transmission is expected to increase with increasing dependence of the host on its microbial partners (Ewald,1987; Bull et al.,1991; Yamamura,1993; Douglas,1994; Thompson,1994; Herre et al.,1999; Wilkinson,2001; Sachs et

al.,2011). In the marine environment symbioses are common and host-microbe interactions have been assessed at different developmental stages for a variety of model organisms, including cnidarians (Apprill et al.,2009; Mortzfeld et al.,2016), echinoderms (Galac et al.,2016; Carrier et al.,2019), cephalopods (Montgomery and McFall-Ngai,1994), molluscs (Endow and Ohta,1990; Wentrup et al.,2014), and annelids (Nussbaumer et al.,2006; Schimak et al.,2015).

Sponges (Porifera) are evolutionarily ancient animals which first appeared between 650 and 700 million years ago (Love et al.,2009; Zumberge et al.,2018). Today, more than 9,000 species (according to the World Porifera Database) are classified into four classes (Demospongiae, Calcarea, Homoscleromorpha, and Hexactinellida) and populate virtually all marine benthic habitats from shallow tropical reefs to the polar seas and the deep ocean (van Soest et al.,2020). Sponges are prominent examples of marine holobionts, consisting of diverse prokaryotic communities with more than 63 prokaryotic phyla detected within sponge hosts (Thomas et al.,2016; Moitinho-Silva et al.,2017a). The prokaryotic communities of sponges are host species-specific and distinct from seawater prokaryotic communities (Hentschel et al.,2002; Thomas et al.,2016). The transmission of microbial sponge symbionts to the next generation occurs vertically or horizontally (De Caralt et al.,2007; Maldonado,2007; Schmitt et al.,2008; Webster et al.,2010; Gloeckner et al.,2013). Microbial communities have been detected in reproductive stages, such as oocytes, embryos, larvae, and juveniles of various shallow water species (Ereskovsky et al.,2005; Schmitt et al.,2007; Sharp et al.,2007; Gloeckner et al.,2013; Björk et al.,2019) indicating that vertical transmission is an important mechanism for the establishment and maintenance of sponge-microbe associations over evolutionary time scales.

In comparison to sponges from shallow and often warm-water habitats, much less is known about sponges and their microbiomes from deep and cold environments. Sponges are

ubiquitous and abundant components of deep-sea benthic communities and can form structurally complex and diverse ecosystems, termed “sponge grounds” (Klitgaard and Tendal,2004; Xavier et al.,2015; Maldonado et al.,2016). Sponge grounds are known to strongly support biodiversity, either directly by provision of shelter or indirectly via food webs, and further perform a variety of ecosystem functions, among others via participation in benthic-pelagic coupling and biogeochemical cycles (reviewed in Maldonado et al.,2016; Thompson and Fuller,2020). Deep-sea sponge grounds (syn. “sponge aggregations” or “sponge gardens”) have been flagged as vulnerable marine ecosystems (VMEs) that are currently under pressure through fishing and other human activities (Pham et al.,2019; Murrillo et al.,2020). The species *Craniella infrequens* and *Craniella zetlandica* are common members of such sponge grounds and have further been recorded in shelf-areas around the globe (**Figure 1**). *Craniella* sponges are spherical in shape and grow up to the size of a golf ball and larger. This study aims to describe the prokaryotic community composition and diversity of two *Craniella* species and their early life stages from the deep North Atlantic Ocean. In particular we sought to address whether prokaryotic community composition, diversity, and dominant microbial phyla differ between different life stages (adults, embryos, and recruits) of the two sponge species. Knowledge about deep-sea sponge holobiont recruitment and plasticity across life stages is crucial to enable, in the long term, an adequate conservation of sponge populations from vulnerable marine ecosystems.

MATERIALS AND METHODS

Sampling

Sponges of different developmental stages (adults, embryos, recruits) were collected during two research cruises onboard the Norwegian research vessel RV G.O.Sars (“GS2017110” and “GS2018108”) to the North Atlantic Ocean (Schulz Bank seamount at

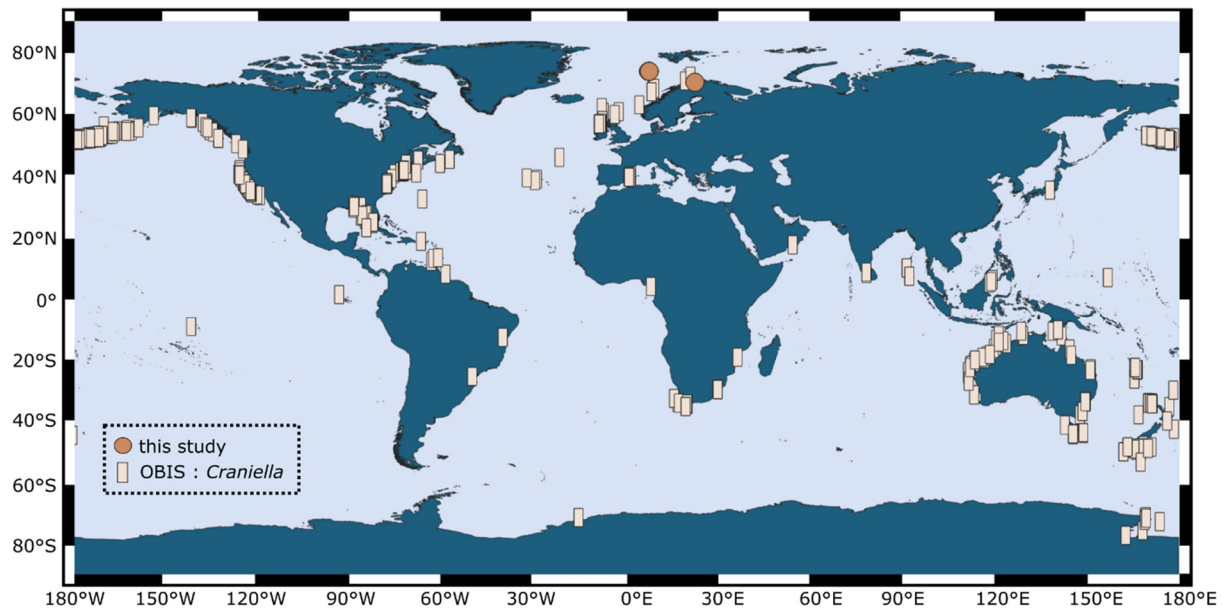


FIGURE 1 | Global distribution of the genus *Craniella*, as compiled from records in the Ocean Biogeographic Information System (OBIS; squares) in March 2020. Sampling locations of individuals from our study are indicated by points.

73.8°N, 7.5°E; and Stjærnsund at 70.3°N, 22.5°E) in July 2017 and August 2018. Samples were collected by remotely operated vehicle (ROV *AEgir*, University of Bergen) or Agassiz trawl at an average depth of 600 m below surface. Tissue chunks of adult sponges and recruits were subsampled onboard ship with sterile scalpels from the mesohyl, rinsed, flash-frozen and stored at -80°C for molecular analyses. Additional pieces of tissue were subsampled with biopsy punches and preserved onboard ship for microscopy. Embryos were carefully picked out of the parent sponges (**Figure 2**) with sterile spring steel forceps using a stereomicroscope. After rinsing, 10 whole embryos were each pooled per adult sponge to account for the small biomass, fixed in RNAlater™ Stabilizing Solution (Thermo Fisher Scientific, Waltham, MA, United States), stored at 4°C overnight, transferred to -20°C on the next day and subsequently to -80°C. In addition, intact embryos were preserved onboard ship for microscopy. In terms of sample numbers, we sampled eight adult *C. infrequens* sponges, of which four were brooding embryos, and nine *C. infrequens* recruits. During the cruise we further sampled eight adult *C. zetlandica* sponges

of which one was brooding embryos. In addition to our cruise activities, we sampled nine *C. zetlandica* recruits in September 2018 from an aquarium system (Bergen, Norway). The respective parent sponges had been maintained since summer 2017 in 45 L tanks with sea-water flow through (1 L min⁻¹) pumped from the adjacent fjord at 105 m depth below surface. Seawater was treated with a 20 µm drum filter prior to flowing through the holding tanks. Under these conditions spawning was documented frequently and the new recruits had settled in the tanks. *C. zetlandica* recruits from the aquarium system were smaller (and thus most likely younger) than the *in situ* sampled *C. infrequens* recruits and therefore flash-frozen as a whole. In our case the new sponge recruits had the following sizes: 1 cm mean diameter for *C. infrequens*; and 0.3 cm maximal diameter for *C. zetlandica*. Adult sponges of both sponge species had the size of a fist. In summary, for this study in total 39 sponge individuals of the two sponge species were collected: *C. zetlandica* (eight adults, of which one was brooding embryos, and nine recruits) and *C. infrequens* (eight adults, of which four were brooding embryos, and nine recruits). In addition to the sponge samples, we also sampled 18

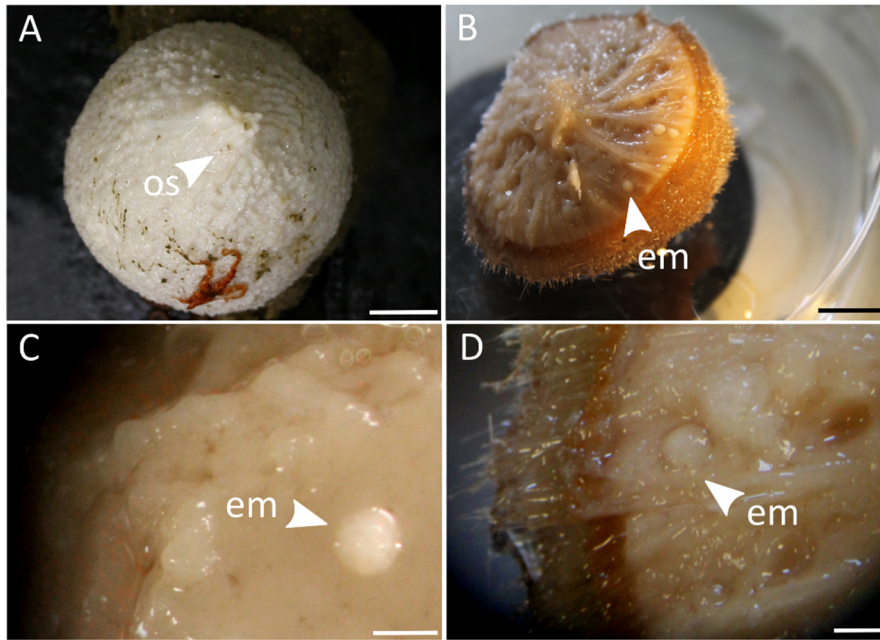


FIGURE 2 | (A) Photograph of an adult *C. zetlandica* individual with an attached red brittle star at the bottom. Scale bar = 1 cm. (B) Photograph of an adult *C. infrequens* individual that was cut in half, showing numerous embryos inside the sponge tissue. Scale bar 1 cm. (C) Stereomicroscopic magnification of *C. zetlandica* tissue. em, embryo. Scale bar = 1 mm. (D) Stereomicroscopic magnification of *C. infrequens* tissue. em, embryo(s) brooding in the mesohyl. Scale bar = 1 mm.

seawater microbial communities during the two cruises (at all sponge sampling locations) and from the aquarium system as environmental reference samples. Two liters of seawater sample were filtered onto polyvinylidene fluoride (PVDF) filter membranes (with 0.22 μm pore size; Merck Millipore) and stored at -80°C .

16S Amplicon Sequencing

DNA extraction was performed on either ~ 0.25 g of tissue or bulk of 10 pooled embryos using the DNeasy Power Soil kit (Qiagen, Venlo, Netherlands). For seawater samples DNA was extracted of one half of a PVDF filter membrane. Subsequent DNA quantity and quality controls were performed with a NanoDrop spectrophotometer and by polymerase chain reaction with universal 16S primers (27F-1429R; Weisburg et al.,1991), followed by gel electrophoresis. For amplicon sequencing, the V3-V4 variable regions of the 16S rRNA gene were amplified in a one-step PCR with the primer pair 341F-806R (dual barcoding approach; Kozich et al.,2013). The primer pair has the following nucleotide sequences: 5'-CCTACGG-GAGGCAGCAG-3' (Muyzer et al.,1993) and

5'-GGACTACHVGGGTWTCTAAT-3' (Caporaso et al.,2011). Verification of PCR-products was performed with gel electrophoresis, followed by normalization and pooling of samples. Sequencing was conducted on a MiSeq platform (MiSeqFGx, Illumina, San Diego, CA, United States) using v3 chemistry. Zero mismatches were allowed in the barcode sequence for demultiplexing.

Raw reads were quality-checked and processed within the QIIME2 environment (version 2018.11; Bolyen et al.,2019). Based on forward reads (truncated to 270 nt), Amplicon Sequence Variants (ASVs) were generated using the DADA2 algorithm (python-based; Callahan et al.,2016). Data were analyzed based on ASVs instead of classical operational taxonomic units (OTUs), in order to resolve fine-scale community patterns and to use a state-of-the-art method which profits from the error correction model conducted by the DADA2 algorithm. Singletons and chimeric sequences were removed during this process.

Phylogenetic trees were calculated based on the resulting ASVs via the FastTree2

plugin. Taxonomic classification of representative ASVs was achieved with the help of a primer-specific trained Naive Bayes taxonomic classifier, based on the Silva 132 99% OTUs 16S database (Quast et al.,2013). Taxonomic filtering was applied on the dataset, meaning that mitochondria, chloroplasts and unassigned sequences were removed. Further, a sampling depth of 3,900 was applied to standardize the number of reads across samples. Diversity metrics were calculated, such as Shannon indices (alpha diversity) and weighted UniFrac distances (beta diversity). Non-metric multidimensional scaling (nMDS) was conducted to evaluate sample separation in ordination space (based on weighted UniFrac distances). We performed PERMANOVAs (999 permutations) on weighted UniFrac distances, and Tukey *Posthoc* tests on Shannon indices following ANOVA type III tests, to assess if the prokaryotic community differed significantly between sample groups. To evaluate further on the differential abundance of microbial phyla present in aquarium sponges (i.e., *C. zetlandica* recruits) in comparison to *in situ* sampled sponges (and in aquarium water vs *in situ* water), we ran the Linear Discriminant Analysis (LDA) Effect Size (LEfSe) algorithm (Segata et al.,2011). Visualizations were created using R (version 3.0.2 with the ggplot2 package loaded; R Development Core Team,2008; Wickham,2016) and Inkscape (version 0.92.4; Harrington and the Inkscape Development Team,2005). The ternary plot was created using the Ternary plot maker web tool (Blom,2019; <https://www.ternaryplot.com/>).

In order to provide an independent proof-of-concept accompanying our 16S amplicon sequencing results we performed qualitative microscopic analyses. These included light microscopy, transmission electron microscopy (TEM), and fluorescence *in situ* hybridization of the most dominant microbial phylum detected by amplicon sequencing. Especially for the embryo-based DNA libraries we aimed at confirming by microscopy that they contained sequences which were indeed present

within the embryos and not merely contaminants.

Light Microscopy and TEM

Tissue samples for light microscopy and TEM were fixed in a cocktail of glutaraldehyde (2.5%) with sodium chloride (0.34 M) and phosphate buffered saline (PBS; 0.4 M). Samples were stored at 4°C. Back on land, samples were washed 3x with buffer (at 4°C; 15 min per washing step), post-fixed in 2% osmiumtetroxide for 2 h and afterward washed three more times with buffer (at 4°C; 15 min per washing step). Subsequently partial dehydration was conducted with an ascending ethanol series, from 30% EtOH (2 x 15 min) to 50% EtOH (1 x 15 min), to 70% EtOH (2 x 15 min, including an overnight storage at 4°C), to 90% EtOH (1 x 15 min) up to 100% EtOH (2 x 15 min). After dehydration, the samples were gradually infiltrated with LR-White resin (at room temperature). The infiltration series covered the following steps: from 1 x 1 h 2:1 EtOH:LR-White, to 1 x 1 h 1:1 EtOH:LR-White, to 1 x 1 h 1:2 EtOH:LR-White, up to 2 x 2 h pure LR-White. Samples were transferred to gelatine embedding capsules filled with fresh LR-White resin and subsequently incubated for polymerization at 57°C for two days. The resulting resin blocks were manually pre-trimmed. Semithin sections (0.5 µm) were cut for light microscopy using an ultramicrotome (Reichert-Jung ULTRACUT E; Leica Camera AG, Wetzlar, Germany) which was equipped with a diamond knife (DIATOME, Biel, Switzerland). After staining with a Richardson solution (after Mulisch and Welsch,2015), semi-thin sections were mounted onto Super-Frost Ultra Plus® microscopy slides and visualized with the help of an Axio Observer.Z1 microscope (Zeiss, Göttingen, Germany).

For TEM, ultrathin sections (70 nm) were cut with the ultramicrotome, mounted on pioloform coated copper grids (75 mesh) and contrasted with uranyl acetate (20 min incubation and subsequent washing steps) as well as with Reynold's lead citrate (3 min incubation, followed by washing steps). Visualization of the ultra-thin sections was performed on a Tecnai

G2 Spirit Bio Twin transmission electron microscope (FEI Company, Hillsboro, United States) at an acceleration voltage of 80 kV.

Fluorescence in situ Hybridization

Sponge tissue samples were fixed overnight (at 4°C) onboard with 4% paraformaldehyde (PFA) in 1xPBS (pH was adjusted to approximately 6.9 with the help of NaOH and HCl). For preparation of all FISH solutions, we used filtered (0.22µm membrane) diethylpyrocarbonate (DEPC)-treated water (Thermo Fisher Scientific, Waltham, United States) to inactivate RNase enzymes, and performed autoclaving for sterilization of buffers and solvents onboard the ship. After fixation with PFA, five washing steps (3 min) were performed with 1xPBS (4°C). Subsequently quenching of free aldehyde groups was conducted by incubation with glycine (50 mM) in 1xPBS for 15 min (room temperature). The quenching step was followed by five additional washing steps (3 min) in DEPC water (at 4°C). Dehydration was performed onboard in an ascending ethanol series (molecular grade, RNase free), covering the following steps: 30% EtOH in 1xPBS (2 x 15 min at 4°C), 50% EtOH in 1xPBS (2 x 30 min). Samples were subsequently stored at 20°C. Back on land, dehydration was continued in following steps: 50% EtOH in MilliQ water (1 x 15 min), 60% EtOH in MilliQ water (2 x 15 min), 70% EtOH in MilliQ water (2 x 15 min), 90% EtOH in MilliQ water (1 x 15 min) up to 100% EtOH (2 x 15 min). The samples were gradually infiltrated with LR-White resin, polymerized, pre-trimmed and cut into semithin sections (0.5 µm) with an ultramicrotome, in the same manner as described in section 2.3. As embryos were embedded in whole, cutting with the microtome was run for a substantial time, until the central part of the embryo was reached. At the area with widest diameter, sections were received and fixed onto two separate SuperFrost Ultra Plus® microscopy slides. While one half of the slides was directly stained with a Richardson solution and mounted using Biomount medium (Plano, Germany; see description in section “Light Microscopy and TEM”). The other half was only fixed

onto microscopy slides, without any staining or covering in Biomount medium. Following the procedures described in Bayer et al. (2018), these latter samples were afterward labeled with *Chloroflexi* clade-specific probes. In particular the following probes were used: Cal825 (5'-[Cy3]-ACACCGCCACACCTCGT-[Cy3]-3'; *E. coli* binding positions 825 to 843) for *Caldilineae*; Ana1005 (5'-[Alexa Fluor 647]-TCCGCTTTCGCTTCCGTA-[Alexa Fluor 647]-3'; *E. coli* binding positions 1005 to 1023) for *Anaerolineae*; and SAR202-104 (5'-[Alexa Fluor 488]-GTTACTCAGCCGTCTGCC-[Alexa Fluor 488]-3'; *E. coli* binding positions 104 to 122) for the *SAR202* group. A double labeling approach at 5' and 3' ends was performed for all probes (Sigma-Aldrich, Germany).

Hybridization was performed in Sylgard chambers (GEOMAR in-house production) which were located inside an equilibrated humid chamber (at 46°C for 3 h). The hybridization buffer consisted of 900 mM NaCl, 20 mM Tris-HCl (pH = 7.4), 10% formamide, 0.01% sodium dodecyl sulfate and 20% dextransulfate. All *Chloroflexi* clade-specific probes were co-hybridized, in addition counterstaining of bacterial nucleic acids and sponge cell nuclei was performed with pre-warmed DAPI in hybridization buffer (1 ng/µL), with an incubation time of 20 min at 46°C. After hybridization, the slides were incubated in pre-warmed wash buffer (450 mM NaCl, 20 mM Tris-HCl, 0.01% sodium dodecyl sulfate) at 48°C for 25 min. The washing step was repeated twice. Subsequently, the slides were carefully rinsed with ice cold molecular grade water, dried and mounted in Mowiol medium (Mowiol® 40–88, Kuraray Europe GmbH, Tokyo, Japan). The same procedure was conducted for a negative control for non-specific binding (i.e., hybridization buffer without *Chloroflexi*-specific probes, but with an oligonucleotide complementary to the probe EUB338 (NON-338-probe; Wallner et al., 1993), to prevent falling into the trap of false-positive results. Fluorescence signals at the four specific wavelength ranges were detected using an Axio Observer.Z1 microscope equipped with

AxioCam 506 and Zen 2 software (version 2.0.0.0; Zeiss Germany).

RESULTS

In this study, the prokaryotic community composition of 39 *Craniella* sponge individuals was analyzed. A visual overview of the denoising stats is compiled in Supplementary Figure S1. The complete dataset after quality filtering consisted of 614,928 reads, which is 82% of the initial input read count. On average each sample had 15,767 reads, but as several samples contained less reads, a sampling depth of 3,900 was applied on the complete dataset for normalization. Four adult *C. infrequens* individuals and one adult *C. zetlandica* individual were found brooding embryos inside their tissue (Figure 2). For each of these brooding individuals, ten embryos had been picked and pooled. In *C. infrequens*, the embryos were more numerous compared to *C. zetlandica* and were scattered throughout the mesohyl of the parental sponge. Prokaryotic diversity (as estimated by Shannon index) was similar between the developmental stages within each sponge species (Figure 3 and Supplementary Table S1).

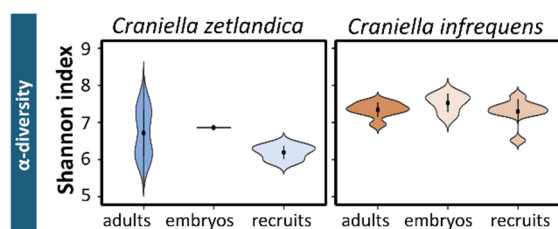


FIGURE 3 | Prokaryotic alpha-diversity for the sampled *Craniella* species. Shannon indices calculated for the different developmental stages of *C. zetlandica* and *C. infrequens*.

Other alpha diversity indices (Pielou's evenness, Faith's PD and number of observed ASVs) led to a similar result (Supplementary Figure S2). Average Shannon indices were slightly lower in *C. zetlandica* than in *C. infrequens*, despite not statistically significantly (Figure 3 and Supplementary Table S1). The prokaryotic community compositions of both sponge species were significantly different from seawater (Figure 4A

and Supplementary Table S2). Further, *C. infrequens* and *C. zetlandica* displayed species-specific microbiomes, as shown by their distinct clustering based on weighted UniFrac distances and on PERMANOVA analyses (Figure 4B, with some exceptions; Supplementary Table S2).

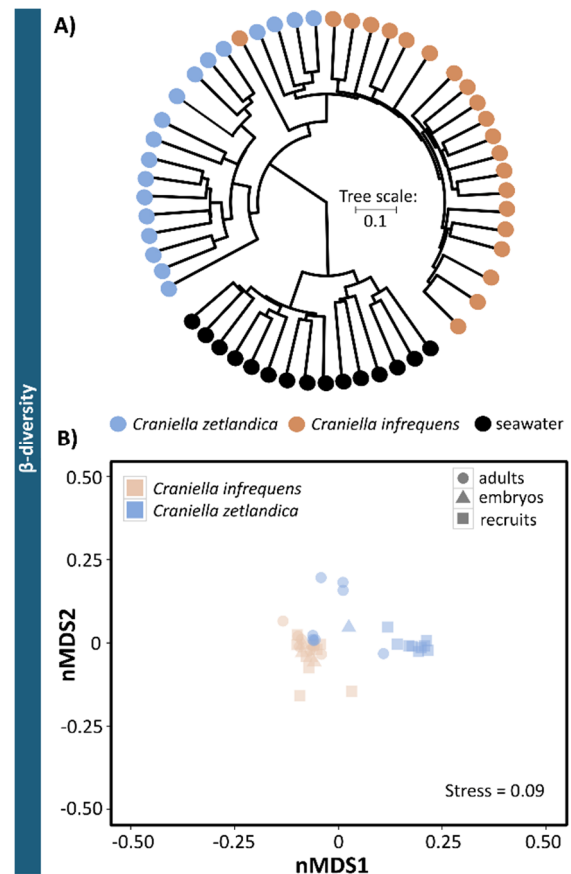


FIGURE 4 | Prokaryotic beta-diversity across the sampled *Craniella* species and reference samples. (A) Clustering dendrogram based on weighted UniFrac distances for all samples of seawater and both sponge species. (B) Non-metric multidimensional scaling (nMDS) plot based on weighted UniFrac distances. The sponge species *C. zetlandica* and *C. infrequens* are marked by different colors, the different developmental stages (adult, recruit, and embryo) are indicated by different symbols.

Overall *C. zetlandica* samples showed a slightly higher variability than *C. infrequens* samples (please note in this regard that while *C. infrequens* specimens were all sampled at the same geographic location, *C. zetlandica* specimens analyzed in this study originate from different geographic locations).

With respect to the prokaryotic community composition, most bacterial phyla were

shared across all life stages, and were not particularly enriched in any specific life stage (Figure 5A). *Chloroflexi* were the most abundant prokaryotic phylum in both sponge species and across all life stages and SAR202 was the most abundant clade, constituting on average 73%

of all recorded *Chloroflexi* (Table 1). *C. zetlandica* recruits obtained from the aquarium system were a noticeable exception that were dominated by *Proteobacteria* and significantly enriched in this phylum compared to *in situ* sampled sponges (Figure 5B and Supplemen-

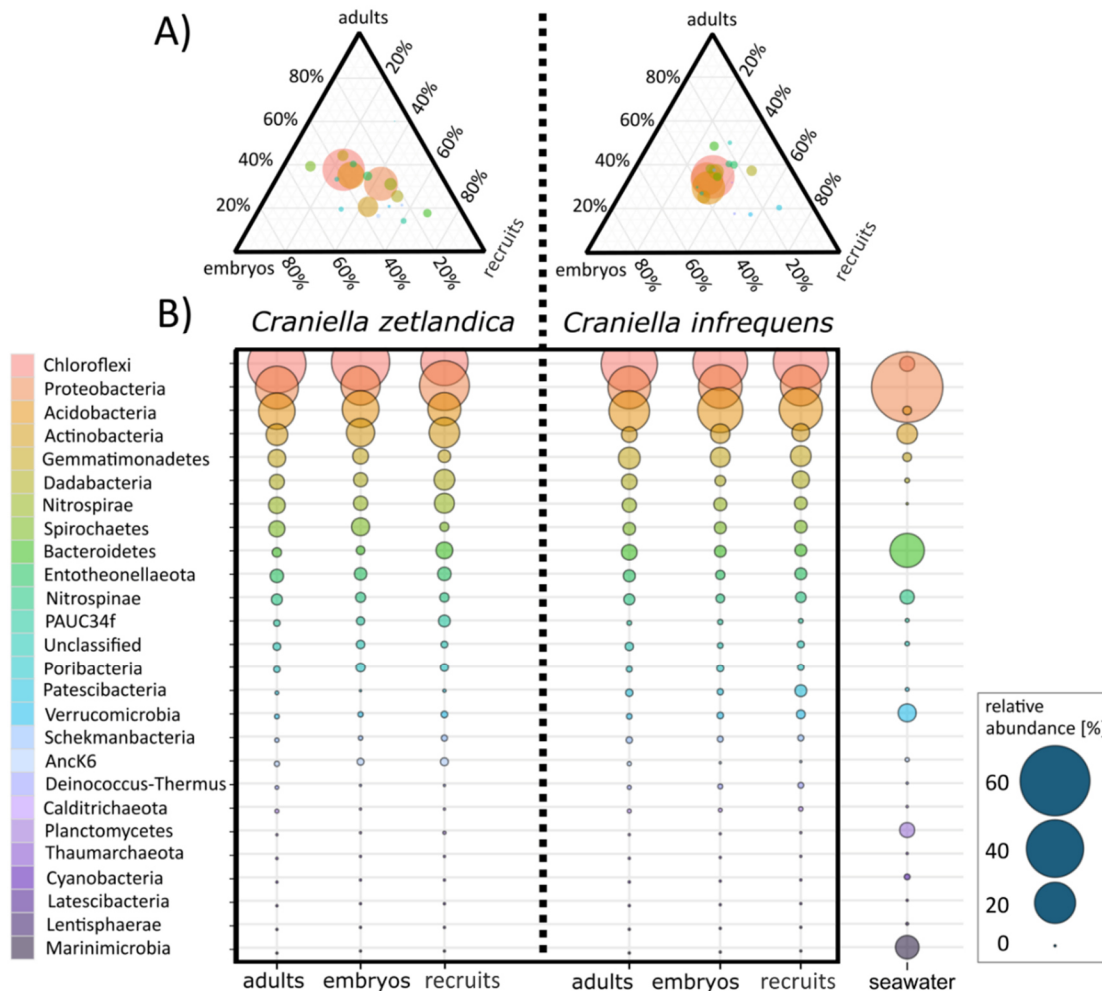


FIGURE 5 | (A) Ternary plot of average prokaryotic phyla distributions for embryos, recruits, and adults of *C. zetlandica* and *C. infrequens*. Colors of phyla are the same in both figures: (A,B). Circle size indicates relative abundance of respective phyla. **(B)** Average prokaryotic community composition of *C. zetlandica* and *C. infrequens* developmental stages as well as seawater. Microbial phyla are sorted in descending order by mean relative abundances. Circle size indicates relative abundances (for exact relative abundance values consider the scale presented at the right bottom).

tary Figure S3A). *Proteobacteria* were however not significantly enriched in the aquarium water compared to *in situ* sampled seawater (Supplementary Figures S3B C). Neither were *Chloroflexi* significantly depleted in the aquarium water compared to *in situ* sampled seawater, although *Chloroflexi* were significantly en-

riched in the *in situ* sampled sponges compared to the aquarium sampled sponges. A total of 390 *Chloroflexi* ASVs were detected in the 39 *Craniella* samples, of which 171 were present in $\geq 50\%$ of the samples per category (adult, embryo, recruit) (Figure 6A). Of these 171 common *Chloroflexi* ASVs, 16 were present in all life stages and both species, while being absent

from seawater reference samples. These 16 ASVs are particularly abundant and account for 20– 40% of the total *Chloroflexi* prokaryotic community in each sponge sample. Seven ASVs

of the 16 ASVs belong to the SAR202 clade, while the remainder belongs to the classes JG30-

TABLE 1 | Average relative abundances of *Chloroflexi* (*Anaerolineae*, *Caldilineae*, SAR202, others) as determined by 16S amplicon sequencing. mean, and standard deviation per group are indicated.

		<i>Anaerolineales</i>	<i>Caldilineales</i>	SAR202 clade	Other <i>Chloroflexi</i>
<i>C. zetlandica</i>	Adult	0.1 ± 0.1	6.9 ± 3.4	77.5 ± 8.2	15.5 ± 7.3
	Embryo	0 ± NA	7.9 ± NA	71.4 ± NA	20.7 ± NA
	Recruit	0 ± 0	15.6 ± 6.0	52.0 ± 14.2	32.4 ± 9.1
<i>C. infrequens</i>	Adult	0.2 ± 0.1	6.0 ± 4.1	85.2 ± 4.4	8.7 ± 2.0
	Embryo	0.2 ± 0.2	7.1 ± 0.8	78.9 ± 3.7	13.7 ± 3.5
	Recruit	0.5 ± 0.7	11.2 ± 6.1	70.0 ± 11.5	18.2 ± 8.1

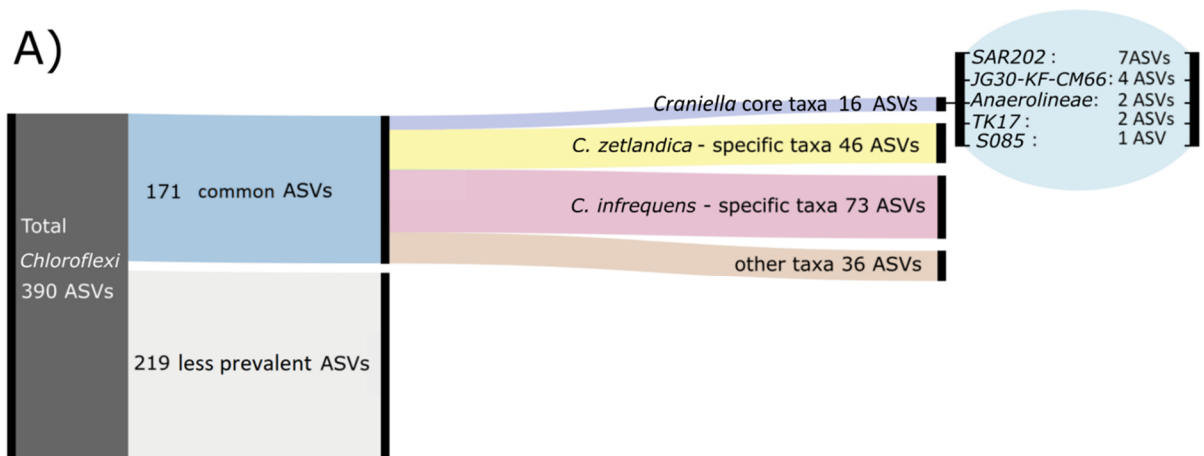
KF-CM66 (4 ASVs), *Anaerolineae* (2 ASVs), TK17 (2 ASVs), and the order S085 (1 ASV).

We next explored whether the 171 common *Chloroflexi* ASVs are species-specific or developmental stage-specific (Figures 6B,C). We determined 46 ASVs as being *C. zetlandica* specific (absent from *C. infrequens*), and 73 ASVs as being *C. infrequens* specific (absent from *C. zetlandica*). For *C. zetlandica* 26 ASVs were exclusively occurring in adults, 12 in embryos, and 10 in recruits. For *C. infrequens* four ASVs were occurring only in adults, 18 in embryos, and four in recruits. Interestingly, one *Chloroflexi*-ASV, occurred only in embryos of both sponge species. This ASV was affiliated with the *Chloroflexi*-class TK17, that could only be classified to class-level. Upon blasting against a sponge microbiome reference database (Moitinho-Silva et al., 2017a), this particular ASV was found in 134 sponges samples, of which 88 samples were identified as *Geodia barretti*. In summary, the major fraction of the detected *Chloroflexi* ASVs were sponge species-specific but not developmental-stage specific. Further information on the relative abundances of the ASV groups is presented in Supplementary Figure S4.

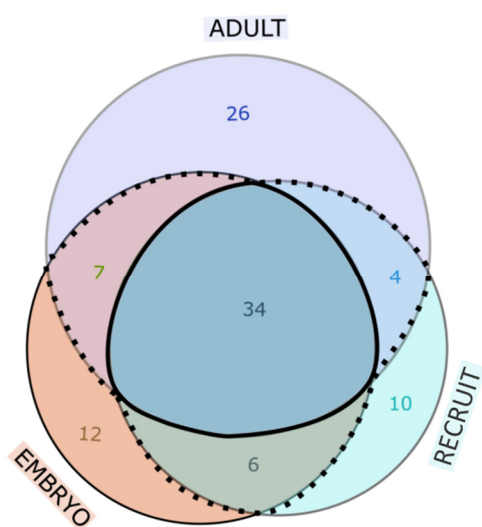
Light microscopy revealed that bacterial cells were numerous inside *C. infrequens*. In the embryos, the microbial cells were restricted to a narrow band at the peripheral rim (Figures 7A,B), while the interior was devoid of

cellular material or bacterial morphotypes. An artifact resulting from cutting through the possibly more fluidic embryo interior can however not be excluded. Fluorescence *in situ* hybridization confirmed high abundances of *Chloroflexi*, in particular SAR202, within the “bacterial band” at the embryo rim (Figure 7C). The negative control for non-specific binding was negative (Supplementary Figures S5I–L). Bacterial cells were further found to be abundant and widely distributed in the mesohyl in adult *C. infrequens* (Supplementary Figure S6A). FISH confirmed high abundances of SAR202 also for adult individuals (Supplementary Figure S6B). Microbial cells within the embryos displayed diverse morphologies and were found in dividing stages (see Figure 7D for examples).

For *C. zetlandica* we observed an *in situ* (stress-induced) release of ripe propagules, visible by ROV live-footage in August 2018, during the sampling campaign “GS2018108” in Stjernsund (Supplementary Video S1: doi.pangaea.de/10.1594/PANGAEA.918218). We suspect a similar timing of spawning in summer for those sponges maintained in the aquarium system in Bergen, as high numbers of settled recruits were found in October 2018 (Supplementary Figure S7). These observations point to a synchronized reproductive cycle in aquarium and *in situ* thriving *C. zetlandica* with a spawning in late summer.



B) *Craniella zetlandica*



C) *Craniella infrequens*

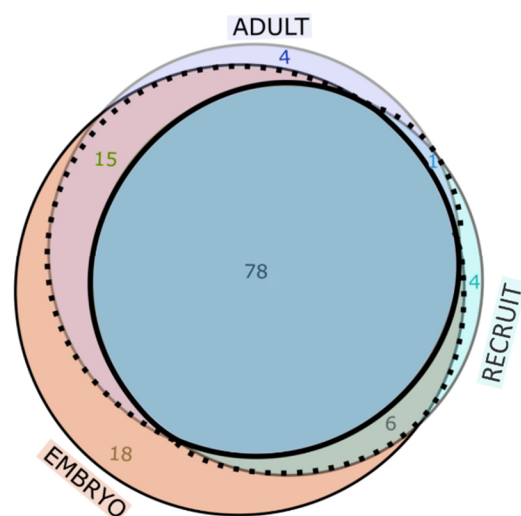


FIGURE 6 | (A) Alluvial diagram showing grouping scheme of *Chloroflexi* ASVs. **(B,C)** Area proportional Venn-diagram showing counts of common *Chloroflexi* ASVs being shared or specific across different life stages for **(B)** *C. zetlandica* and **(C)** *C. infrequens*.

DISCUSSION

Deep-sea sponge grounds are underexplored ecosystems that provide numerous goods and services to the functioning of the deep-sea (Maldonado et al., 2016), such as the provision of food, habitat and refugia for commercially relevant fish species. The need to preserve these vulnerable marine ecosystems (VMEs) is thus contradictory to the high demand for socio-economic processes such as fishing in North Atlantic waters. In this study we assessed the prokaryotic diversity of two adult deep-sea sponge species of the genus *Craniella* and their developmental stages. These sponges are

prominent although understudied members of deep-sea sponge grounds and have further been recorded in shelf-areas around the globe. While there is one report on the *Proteobacteria*-dominated microbiome of *Craniella australiensis*, which is a shallow water species in the Indo-Pacific and South China Sea (Anbuhezian and Li, 2014), the microbiome of *C. infrequens* and *C. zetlandica* have—to our knowledge—not been characterized to date.

Our results reveal that *C. zetlandica* and *C. infrequens* are typical high microbial abundance (HMA) sponges (Gloeckner et al.,

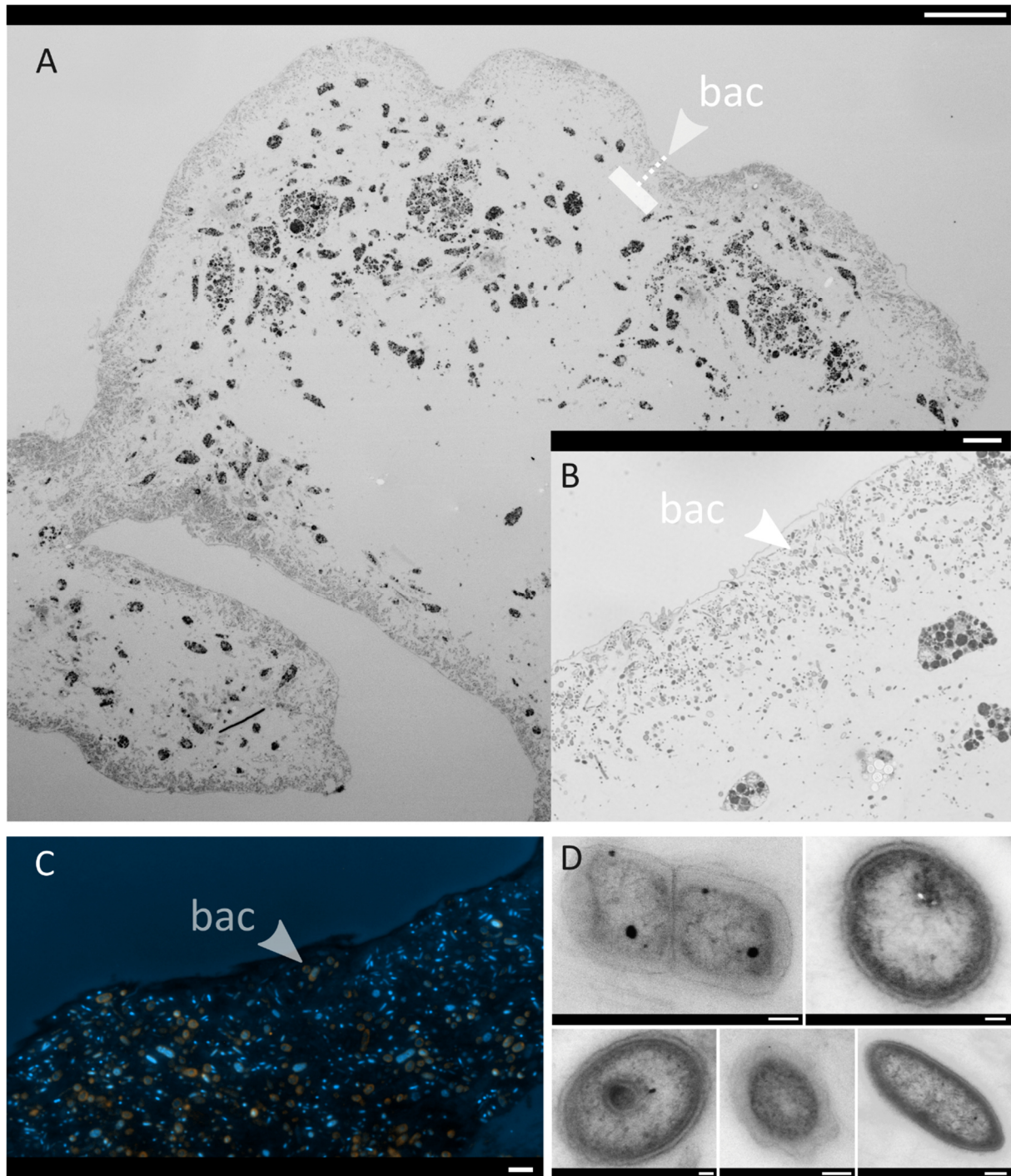


FIGURE 7 | Micrographs of one representative *C. infrequens* embryo ($n = 3$ inspected). Arrows point from the outside to the inside. bac = band of bacterial cells **(A)** A whole *C. infrequens* embryo stained with Richardson solution. Scale bar = 100 μm . The dashed line indicates the band with highest bacterial densities **(B)** Brightfield microscopy of the bacterial band, scale bar = 10 μm . **(C)** Fluorescence *in situ* hybridization of the bacterial band. Merged channels of SAR202 signals (orange) and the DAPI counter stain (blue) are shown, scale bar = 2 μm . Micrographs of isolated hybridization signals for the specific probes (*Anaerolineae*, *Caldilineae*, and SAR202) are included in Supplementary Figures S5A–D. **(D)** TEM close-up of microbial cells located within the bacterial band, scale bars = 200 nm.

2014; Moitinho-Silva et al.,2017b). We base this statement on our observation that both sponge species harbor diverse and dense microbial communities within their tissues, as ob-

served by TEM. This finding is further supported by the observation that the sponge microbiomes are dominated by *Chloroflexi* (SAR202), which are known indicator taxa of HMA sponges (Moitinho-Silva et al.,2017b).

The microbiomes of *C. zetlandica* and *C. infrequens* were host-species specific and distinctly different from seawater which is consistent with previous reports on sponge microbiomes (Easson and Thacker,2014; Thomas et al.,2016; Steinert et al.,2017). The prokaryotic community composition and alpha diversity were very similar between adults, embryos and recruits within the same sponge species, indicating that entire microbial consortia are very likely vertically transmitted. Vertical transmission has been previously reported for a number of shallow-water HMA sponges (i.e., Schmitt et al.,2007; Björk et al.,2019), but is, to the best of our knowledge, novel in the context of sponge species from sponge grounds. Our results indicate that young deep-sea sponges can already contain a high diversity and density of microorganisms. Despite the observation of very similar prokaryotic community compositions between all examined life stages per species in our study, the recruits of *C. zetlandica* fell somewhat distant from all other samples in the nMDS plots and were dominated by *Proteobacteria*. It is noteworthy that these samples were collected in an aquarium system and not in the field. Indeed, previous studies have shown that the maintenance of marine organisms in captivity can affect the relative abundances of some associated microbial taxa (Galand et al.,2018). In addition, microbial variability has been observed to be particularly high in recently settled recruits compared to adult individuals (Schmitt et al.,2007; Busch et al.,2020).

The most dominant microbial phylum in all *in situ* sampled sponges of both sponge species and across all developmental stages were *Chloroflexi*. Generally, the probability of vertical symbiont-transmission is assumed to be positively related with the dependence of the host on the microbial partner partners (Ewald,1987; Bull et al.,1991; Yamamura,1993; Douglas,1994; Thompson,1994; Herre et al.,1999; Wilkinson,2001; Sachs et al.,2011). Bacteria of the phylum *Chloroflexi* (in particular *SAR202*) are abundant and wide-spread in HMA sponges (Moitinho-Silva et al.,2017b; Bayer et al.,2018) and are suggested to play an

important function in the degradation of recalcitrant organic matter (Landry et al.,2017; Bayer et al.,2018). We propose that the 16 *Craniella* core ASVs detected within our study should be considered as sponge symbionts as they were found in high abundances across both species and all developmental stages, while being absent from seawater. Almost half of the 16 *Craniella* core ASVs belong to the *SAR202* clade, making it tempting to speculate about a potential crucial functional role in respect to (recalcitrant) organic matter cycling. Our FISH analyses revealed high abundances of *Chloroflexi* scattered within the bacterial band at the inner periphery of sponge embryos. We conclude that the transmission of *Chloroflexi* through the embryonal stages reflects their tight and evolutionarily long-standing integration in the sponge holobiont. Future research efforts should be geared toward understanding the specific functional roles of *Chloroflexi* symbionts for deep-sea sponges.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below:
<https://www.ncbi.nlm.nih.gov/>, PRJNA612549.

AUTHOR CONTRIBUTIONS

KBu and UH designed the study, wrote the manuscript, and were responsible for the microbial pipeline. KBu, HR, and EW conducted sampling. KBa and KBu performed fluorescence *in situ* hybridization. HR led sponge taxonomic analysis. AF was involved in sequencing. KBu performed the data analysis (bioinformatics and visualizations). EW and KBa reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was funded by the European Union's Horizon 2020 Research and Innovation Program under Grant Agreement No. 679849 (the SponGES project). This document reflects only the authors' view and the Executive Agency for Small and Medium-sized Enterprises (EASME) is not responsible for any use that may be made of the information it contains.

ACKNOWLEDGMENTS

We express our gratitude to HR, coordinator of the H2020-SponGES project on deep-sea sponge grounds of the North Atlantic, deep-sea sponge taxonomist, explorer, mentor, and friend. We thank the crews and scientific parties of RV G.O.Sars cruises "GS2017110" and

"GS2018108". On these cruises the ROV AEGir-team (University of Bergen) enabled sampling of precious samples. Ina Clefsen and Andrea Hethke provided valuable support in the lab back on land. We also thank Heikki Savolainen and Tone Ulvatn for help in aquaria facility management and technical support at the University of Bergen. Heidi Kristina Meyer's support in the acquisition of video material is much appreciated.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at:

<https://www.frontiersin.org/articles/10.3389/fmars.2020.00674/full#supplementary-material>.

REFERENCES

- Anbuezhian, R., and Li, Z. (2014). "Marine sponge *Craniella australiensis*- associated bacterial," in *Encyclopedia of Metagenomics*, ed. K. E. Nelson (Berlin: Springer-Verlag), 319–326. doi: 10.1007/978-1-4899-7475-4_604
- Apprill, A., Marlow, H. Q., Martindale, M. Q., and Rappé, M. S. (2009). The onset of microbial associations in the coral *Pocillopora meandrina*. *ISME J.* 3, 685–699. doi: 10.1038/ismej.2009.3
- Bayer, K., Jahn, M. T., Slaby, B. M., Moitinho-Silva, L., and Hentschel, U. (2018). Marine sponges as *Chloroflexi* hot spots: genomic insights and high-resolution visualization of an abundant and diverse symbiotic clade. *mSystems* 3:e00150-18.
- Björk, J. R., Díez-Vives, C., Astudillo-García, C., Archie, E. A., and Montoya, J. M. (2019). Vertical transmission of sponge microbiota is inconsistent and unfaithful. *Nat Ecol Evol.* 3, 1172–1183. doi: 10.1038/s41559-019-0935-x
- Blom, J. (2019). *Ternary Plot Maker*. Available online at: <https://www.ternaryplot.com/> (accessed March 1, 2020).
- Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C., Al Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable, and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* 37, 852–857.
- Bull, J. J., Molineux, I. J., and Rice, W. R. (1991). Selection of benevolence in a host-parasite system. *Evolution* 45, 875–882. doi: 10.1111/j.1558-5646.1991.tb04356.x
- Busch, K., Beazley, L., Kenchington, E., Whoriskey, F., Slaby, B., and Hentschel, U. (2020). Microbial diversity of the glass sponge *Vazella pourtalesii* in response to anthropogenic activities. *bioRxiv* [Preprint]. doi: 10.1101/2020.05.19.102806
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P. (2016). DADA2: high-resolution sample inference from Illumina amplicon data. *Nat. Methods* 13, 581–583. doi: 10.1038/nmeth.3869
- Caporaso, J. G., Lauber, C. L., Walters, W. A., Berg-Lyons, D., Knight, R., Lozupone, C. A., et al. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. U.S.A.* 108, 4516–4522. doi: 10.1073/pnas.1000080107
- Carrier, T. J., Dupont, S., and Reitzel, A. M. (2019). Geographic location and food availability offer differing levels of influence on the bacterial communities associated with larval sea urchins. *FEMS Microbiol. Ecol.* 95, 1–9.
- De Caralt, S., Uriz, M. J., and Wijffels, R. H. (2007). Vertical transmission and successive location of symbiotic bacteria during embryo development and larva formation in *Corticium candellabrum* (Porifera: Demospongiae). *J. Mar. Biol. Assoc. U. K.* 87, 1693–1699. doi: 10.1017/s0025315407056846
- Douglas, A. E. (1994). *Symbiotic Interactions*. Oxford: Oxford University Press.
- Easson, C. G., and Thacker, R. W. (2014). Phylogenetic signal in the community structure of host-specific microbiomes of tropical marine sponges. *Front. Microbiol.* 5:532. doi: 10.3389/fmicb.2014.00532
- Eberl, G. (2010). A new vision of immunity: homeostasis of the superorganism. *Mucosal. Immunol.* 3, 450–460. doi: 10.1038/mi.2010.20
- Endow, K., and Ohta, S. (1990). Occurrence of bacteria in the primary oocytes of vesicomid clam *Calyptogena soyoeae*. *Mar. Ecol. Prog. Ser.* 64, 309–311. doi: 10.3354/meps064309
- Ereskovsky, A. V., Gonobobleva, E., and Vishnyakov, A. (2005). Morphological evidence for vertical transmission of symbiotic bacteria in the viviparous sponge *Halisarca dujardini* Johnston (Porifera: Demospongiae, Halisarcida). *Mar. Biol.* 146, 869–875. doi: 10.1007/s00227-004-1489-1
- Ewald, P. W. (1987). Transmission modes and evolution of the parasitism-mutualism continuum. *Ann. N. Y. Acad. Sci.* 503, 295–306. doi: 10.1111/j.1749-6632.1987.tb040616.x
- Flórez, L. V., Biedermann, P. H. W., Engl, T., and Kaltenpoth, M. (2015). Defensive symbioses of animals with prokaryotic and eukaryotic microorganisms. *Nat. Prod. Rep.* 32, 904–936. doi: 10.1039/c5np00010f
- Galac, M. R., Bosch, I., and Janies, D. A. (2016). Bacterial communities of oceanic sea star (Asteroidea: Echinodermata) larvae. *Mar. Biol.* 163:162.
- Galand, P. E., Chapron, L., Meistertzheim, A.-L., Peru, E., and Lartaud, F. (2018). The effect of captivity on the dynamics of active bacterial communities differs between two deep-sea coral species. *Front. Microb.* 9:2565. doi: 10.3389/fmicb.2018.02565
- Gloeckner, V., Lindquist, N., Schmitt, S., and Hentschel, U. (2013). *Ectyoplasia ferox*, an experimentally tractable model for vertical microbial transmission in marine sponges. *Microb. Ecol.* 65, 462–474. doi: 10.1007/s00248-012-0142-7
- Gloeckner, V., Wehr, M., Moitinho-silva, L., Schupp, P., Pawlik, J. R., Lindquist, N. L., et al. (2014). The HMA-LMA dichotomy revisited?: an electron microscopical survey of 56 sponge species. *Biol. Bull.* 227, 78–88. doi: 10.1086/bblv227n1p78
- Harrington, B., and the Inkscape Development Team (2005). *Inkscape*. Available online at: <http://www.inkscape.org/> (accessed February 18, 2019).

- Hentschel, U., Hopke, J., Horn, M., Friedrich, A. B., Wagner, M., Hacker, J., et al. (2002). Molecular evidence for a uniform microbial community in sponges from different oceans. *Appl. Environ. Microbiol.* 68, 4431–4440. doi: 10.1128/AEM.68.9.4431-4440.2002
- Herre, E. A., Knowlton, N., Mueller, U. G., and Rehner, S. A. (1999). The evolution of mutualisms: exploring the paths between conflict and cooperation. *Trends Ecol. Evol.* 14, 49–53. doi: 10.1016/S0169-5347(98)01529-8
- Klitgaard, A. B., and Tendal, O. S. (2004). Distribution and species composition of mass occurrences of large-sized sponges in the northeast Atlantic. *Prog. Oceanogr.* 61, 57–98. doi: 10.1016/j.pocean.2004.06.002
- Koropatnick, T. A., Engle, J. T., Apicella, M. A., Stabb, E. V., Goldman, W. E., and McFall-ngai, M. J. (2004). Microbial factor-mediated development in a host-bacterial mutualism. *Science* 306, 1186–1189.
- Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., and Schloss, P. D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the Miseq Illumina sequencing platform. *Appl. Environ. Microbiol.* 79, 5112–5120. doi: 10.1128/aem.01043-13
- Landry, Z., Swa, B. K., Herndl, G. J., Stepanauskas, R., and Giovannoni, S. J. (2017). SAR202 genomes from the dark ocean predict pathways for the oxidation of recalcitrant dissolved organic matter. *mBio* 8, e00413-17.
- Love, G. D., Grosjean, E., Stalvies, C., Fike, D. A., Grotzinger, J. P., Bradley, A. S., et al. (2009). Fossil steroids record the appearance of Demospongiae during the Cryogenian period. *Nature* 457, 718–721. doi: 10.1038/nature07673
- Maldonado, M. (2007). Intergenerational transmission of symbiotic bacteria in oviparous and viviparous sponges, with emphasis on intracytoplasmically-compartmented bacterial types. *J. Mar. Biol. Assoc. U. K.* 87, 1701–1713. doi: 10.1017/s0025315407058080
- Maldonado, M., Aguilar, R., Bannister, R. J., Bell, J. J., Conway, K. W., Dayton, P. K., et al. (2016). “Sponge grounds as key marine habitats: a synthetic review of types, structure, functional roles, and conservation concerns,” in *Marine Animal Forests*, eds S. Rossi, L. Bramanti, A. Gori, and C. Orejas (Berlin: Springer-Verlag), 1–39. doi: 10.1007/978-3-319-17001-5_24-1
- McFall-Ngai, M., Hadfield, M. G., Bosch, T. C. G., Carey, H. V., Domazet-Lošo, T., and Douglas, A. E. (2013). Animals in a bacterial world, a new imperative for the life sciences. *Proc. Natl. Acad. Sci. U.S.A.* 110, 3229–3236.
- Moitinho-Silva, L., Nielsen, S., Amir, A., Gonzalez, A., Ackermann, G. L., Cerrano, C., et al. (2017a). The sponge microbiome project. *Gigascience* 6:gix077.
- Moitinho-Silva, L., Steinert, G., Nielsen, S., Hardoim, C. C. P., Wu, Y. C., McCormack, G. P., et al. (2017b). Predicting the HMA-LMA status in marine sponges by machine learning. *Front. Microbiol.* 8:752. doi: 10.3389/fmicb.2017.00752
- Montgomery, M. K., and McFall-Ngai, M. (1994). Bacterial symbionts induce host organ morphogenesis during early postembryonic development of the squid *Euprymna scolopes*. *Development* 120, 1719–1729.
- Mortzfeld, B. M., Urbanski, S., Reitzel, A. M., Künzel, S., Technau, U., and Fraune, S. (2016). Response of bacterial colonization in *Nematostella vectensis* to development, environment and biogeography. *Environ. Microbiol.* 18, 1764–1781. doi: 10.1111/1462-2920.12926
- Mulisch, M., and Welsch, U. (2015). *Romeis Mikroskopische Technik*. Berlin: Springer, 1–603.
- Murillo, F. J., Kenchington, E., Koen-Alonso, M., Guijarro, J., Kenchington, T. J., Sacau, M., et al. (2020). Mapping benthic ecological diversity and interactions with bottom-contact fishing on the Flemish Cap (northwest Atlantic). *Ecol. Indic.* 112:106135. doi: 10.1016/j.ecolind.2020.106135
- Muyzer, G., Waal, E. C. D. E., and Utierlinden, A. G. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl. Environ. Microbiol.* 59, 1–6.
- Nicholson, J. K., Holmes, E., Kinross, J., Burcelin, R., Gibson, G., Jia, W., et al. (2012). Metabolic Interactions. *Science* 306, 1262–1268.
- Nussbaumer, A. D., Fisher, C. R., and Bright, M. (2006). Horizontal endosymbiont transmission in hydrothermal vent tubeworms. *Nature* 441, 345–348. doi: 10.1038/nature04793
- Pham, C. K., Murillo, F. J., Lirette, C., Maldonado, M., Colaço, A., Ottaviani, D., et al. (2019). Removal of deep-sea sponges by bottom trawling in the Flemish Cap area: conservation, ecology and economic assessment. *Sci. Rep.* 9:15843.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res.* 41, 590–596.
- R Development Core Team (2008). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Rosenberg, E., and Zilber-Rosenberg, I. (2011). Symbiosis and development: the hologenome concept. *Birth Defects Res. Part C Embryo Today Rev.* 93, 56–66. doi: 10.1002/bdrc.20196
- Sachs, J. L., Skophammer, R. G., and Regus, J. U. (2011). Evolutionary transitions in bacterial symbiosis. *Proc. Natl. Acad. Sci. U.S.A.* 108, 10800–10807. doi: 10.1073/pnas.1100304108
- Schimak, M. P., Schoffelen, N. J., Gruhl, A., and Dubilier, N. (2015). *Transmission of Bacterial Symbionts in the Gutless Oligochaete Olavius Algarvensis*. Bremen: University of Bremen, 161–199.
- Schmitt, S., Angermeier, H., Schiller, R., Lindquist, N., and Hentschel, U. (2008). Molecular microbial diversity survey of sponge reproductive stages and mechanistic insights into vertical transmission of microbial symbionts. *Appl. Environ. Microbiol.* 74, 7694–7708. doi: 10.1128/aem.00878-08
- Schmitt, S., Weisz, J. B., Lindquist, N., and Hentschel, U. (2007). Vertical transmission of a phylogenetically complex microbial consortium in the viviparous sponge *Ircinia felix*. *Appl. Environ. Microbiol.* 73, 2067–2078. doi: 10.1128/aem.01944-06
- Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., et al. (2011). Metagenomic biomarker discovery and explanation. *Genome Biol.* 12:R60.
- Sharp, K. H., Eam, B., John Faulkner, D., and Haygood, M. G. (2007). Vertical transmission of diverse microbes in the tropical sponge *Corticium* sp. *Appl. Environ. Microbiol.* 73, 622–629. doi: 10.1128/aem.01493-06
- Steinert, G., Rohde, S., Janussen, D., Blaurock, C., and Schupp, P. J. (2017). Host-specific assembly of sponge-associated prokaryotes at high taxonomic ranks. *Sci. Rep.* 7, 1–9.
- Thomas, T., Moitinho-Silva, L., Lurgi, M., Björk, J. R., Easson, C., Astudillo-García, C., et al. (2016). Diversity, structure and convergent evolution of the global sponge microbiome. *Nat. Commun.* 7:11870.
- Thompson, J. N. (1994). *The Coevolutionary Process*. Chicago, IL: The University of Chicago Press, 1–383.
- Thompson, T., and Fuller, S. D. (2020). *Technical Measures and Environmental Risk Assessments for Deep-Sea Sponge Conservation*. Rome: FAO, 1–30.
- van Soest, R. W. M., Boury-Esnault, N., Hooper, J. N. A., Rützler, K., de Voogd, N. J., Alvarez, B., et al. (2020). *World Porifera Database*. Accessed at: <http://www.marinespecies.org/porifera> (accessed February 06, 2020).
- Wallner, G., Amann, R., and Beisker, W. (1993). Optimizing fluorescent *in situ* hybridization with rRNA-targeted oligonucleotide probes for flow cytometric identification of microorganisms. *Cytometry* 14, 136–143. doi: 10.1002/cyto.990140205
- Webster, N. S., Taylor, M. W., Behnam, F., Lückner, S., Rattei, T., Whalan, S., et al. (2010). Deep sequencing reveals exceptional diversity and modes of transmission for bacterial sponge symbionts. *Environ. Microbiol.* 12, 2070–2082.
- Weisburg, W. G., Barns, S. M., Pelletier, D. A., and Lane, D. J. (1991). 16S ribosomal DNA amplification for phylogenetic study. *J. Bacteriol.* 173, 697–703. doi: 10.1128/jb.173.2.697-703.1991
- Wentrup, C., Wendeborg, A., Schimak, M., Borowski, C., and Dubilier, N. (2014). Forever competent: deep-sea bivalves are colonized by their chemosynthetic symbionts throughout their lifetime. *Environ. Microbiol.* 16, 3699–3713. doi: 10.1111/1462-2920.12597
- Wickham, H. (2016). *ggplot2: Elegant graphics for Data Analysis*. New York, NY: Springer-Verlag, 1–213.
- Wilkinson, D. M. (2001). Horizontally acquired mutualisms, an unsolved problem in ecology?? *Oikos* 92, 377–384. doi: 10.1034/j.1600-0706.2001.920222.x

Xavier, J. R., Cárdenas, P., Cristobo, J., Van Soest, R., and Rapp, H. T. (2015). Systematics and biodiversity of deep-sea sponges of the Atlanto-Mediterranean region. *J. Mar. Biol. Assoc. U. K.* 95, 1285–1286.

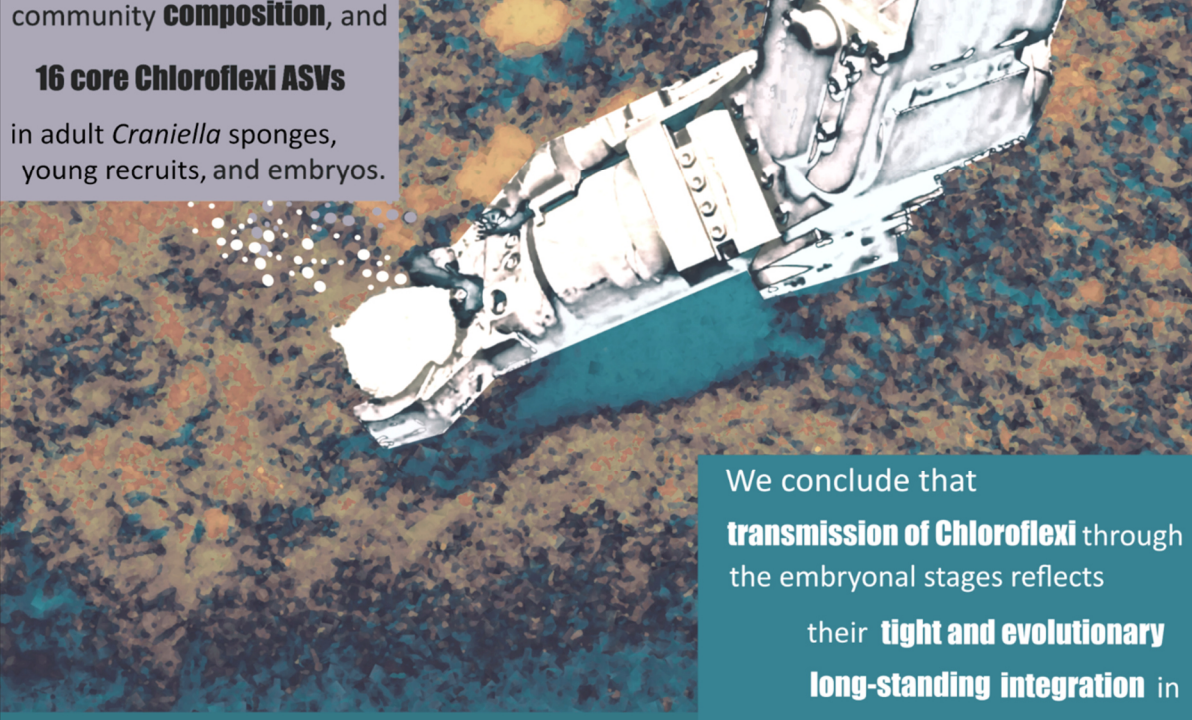
Yamamura, N. (1993). Vertical transmission and evolution of mutualism from parasitism. *Theor. Popul. Biol.* 44, 95–109. doi: 10.1006/tpbi.1993.1020

Zumberge, J. A., Love, G. D., Cárdenas, P., Sperling, E. A., Gunasekera, S., Rohrsen, M., et al. (2018). Demosponge steroid biomarker 26-methylstigmastane provides evidence for Neoproterozoic animals. *Nat. Ecol. Evol.* 2, 1709–1714. doi: 10.1038/s41559-018-0676-2

GRAPHICAL SUMMARY

Chloroflexi dominate the deep-sea golf ball sponges

Craniella zetlandica and *Craniella infrequens* throughout different life stages

Busch	Wurz	Rapp	Bayer	Franke	Hentschel	
illustrated by K. Busch We observed equal diversity metrics , a similar microbial community composition, and 16 core Chloroflexi ASVs in adult <i>Craniella</i> sponges, young recruits, and embryos.		BACKGROUND	The fitness of many organisms is influenced by the establishment of host-microbial colonisation during development. Beneficial symbionts are expected to be vertically transmitted .			
			We conclude that transmission of Chloroflexi through the embryonal stages reflects their tight and evolutionary long-standing integration in the sponge holobiont.			

Chapter 3

Microbial diversity of the glass sponge *Vazella pourtalesii* in response to anthropogenic activities

© 2020 Busch et al.

Original publication by *Conservation Genetics* (Springer Nature)

Reprinted under the terms of the Creative Commons Attribution License (CC-BY 4.0).

doi: 10.1007/s10592-020-01305-2.

Microbial diversity of the glass sponge *Vazella pourtalesii* in response to anthropogenic activities

Kathrin Busch^{1*}, Lindsay Beazley², Ellen Kenchington², Frederick Whoriskey³, Beate M. Slaby¹, Ute Hentschel^{1,4}

¹ GEOMAR Helmholtz Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany, ² Department of Fisheries and Oceans, Bedford Institute of Oceanography, Dartmouth, NS, Canada, ³ Ocean Tracking Network, Dalhousie University, Halifax, NS, Canada, ⁴ Christian-Albrechts University of Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany. *Correspondence author.

Establishment of adequate conservation areas represents a challenging but crucial task in the conservation of genetic diversity and biological variability. Anthropogenic pressures on marine ecosystems and organisms are steadily increasing. Whether and to what extent these pressures influence marine genetic biodiversity is only starting to be revealed. Using 16S rRNA gene amplicon sequencing, we analysed the microbial community structure of 33 individuals of the habitat-forming glass sponge *Vazella pourtalesii*, as well as reference seawater, sediment, and biofilm samples. We assessed how two anthropogenic impacts, i.e. habitat destruction by trawling and artificial substrate provision (moorings made of composite plastic), correspond with in situ *V. pourtalesii* microbiome variability. In addition, we evaluated the role of two bottom fishery closures in preserving sponge-associated microbial diversity on the Scotian Shelf, Canada. Our results illustrate that *V. pourtalesii* sponges collected from protected sites within fishery closures contained distinct and taxonomically largely novel microbial communities. At the trawled site we recorded significant quantitative differences in distinct microbial phyla, such as a reduction in Nitrospinae in the four sponges from this site and the environmental references. Individuals of *V. pourtalesii* growing on the mooring were significantly enriched in Bacteroidetes, Verrucomicrobia and Cyanobacteria in comparison to sponge individuals growing on the natural seabed. Due to a concomitant enrichment of these taxa in the mooring biofilm, we propose that biofilms on artificial substrates may 'prime' sponge-associated microbial communities when small sponges settle on such substrates. These observations likely have relevant management implications when considering the increase of artificial substrates in the marine environment, e.g., marine litter, off-shore wind parks, and petroleum platforms.

Keywords: *Vazella pourtalesii*, Sponge conservation areas (SCAs), Glass sponge grounds, Microbiome, Ocean tracking network (OTN), Anthropogenic impact, Marine litter, Trawling

INTRODUCTION

Glass sponges (Hexactinellida) are extraordinary animals with a skeleton made of silicon dioxide and a unique histology which is distinct from all other known sponge classes (Leys et al. 2007). The greatest taxonomic diversity of glass sponges is found between 300 and 600 m depth, with only a few populations occurring in shallow (euphotic) waters (Leys et al. 2007). Hexactinellids are among the most ancient

metazoans with an estimated origin of 800 million years ago as determined by molecular clocks (Leys et al. 2007), and therefore represent promising candidates for examining evolutionary ancient biological mechanisms and relationships such as sponge-microbe interactions (Pita et al. 2018). The microbial communities associated with sponges can be very diverse, with more than 63 phyla reported previously (Thomas et al. 2016; Moitinho-Silva et al.

2017). In order to gain a holistic understanding of multicellular organisms, also their internal symbiotic associations should be considered (McFall-Ngai et al. 2013). Along similar lines, the term ‘holobiont’ (syn. ‘metaorganism’ (Bosch and McFall-Ngai 2011)) has been defined to cover the host plus its associated microbiota (Bordenstein and Theis 2015; Rohwer et al. 2002). While the histology and trophic ecology of glass sponges has been a matter of several previous studies (Kahn and Leys 2017; Kahn et al. 2018), the microbiology of glass sponges still remains largely unexplored (but see Steinert et al. 2020; Tian et al. 2016; Savoca et al. 2019).

The glass sponge *Vazella pourtalesii* (Schmidt 1870) is distributed along the continental margin of eastern North America from the Florida Keys in the southeastern USA to the Scotian Shelf off Nova Scotia, Canada, where it forms pronounced monospecific aggregations with densities reaching up to 4 individuals per m² (Fig. 1; Fig. 2a, Kenchington et al. unpublished data). In 2013, Fisheries and Oceans Canada (DFO) established two Sponge Conservation Areas (the Sambro Bank and Emerald Basin Sponge Conservation Areas, referred to herein as SCAs) to protect two of the most significant concentrations of *V. pourtalesii* from bottom-fishing activities. However, these closures protect less than 2% of the total area covered by the *V. pourtalesii* sponge grounds (Kenchington et al. unpublished data), and bottom fishing activities continue to occur almost immediately adjacent to their borders (DFO 2017).

Areas situated in the direct vicinity of fisheries closures represent potentially lucrative areas for fishing, as ‘spillover’ of fish stocks from the protected sites may occur. Bottom trawling is one of the most destructive ways to catch fish (Kelleher 2005). Impacts include reduced fishing stocks, by-catch of non-target species and destruction of habitat for benthic invertebrates in the trawl path and neighboring vicinity. At trawled sites, the local hydrodynamics may be altered due to a removal of habitat-

forming benthic structures (living and non-living). Further, impacts at the population-, community-, and ecosystem-levels have resulted from the removals of unintentionally (by-caught) fished taxa (Ortuño Crespo and Dunn 2017). Benthic trawling also impacts the microbial community: For example Jackson et al. (2001) (and references therein) report a strong increase in microbial cell numbers in the water column associated with overfishing. For glass sponges such as *V. pourtalesii*, the effects of bottom trawling on the variability of microbial community compositions are currently unknown. In terms of sponge physiology, glass sponges were previously observed to arrest pumping in response to high sediment levels in the water column (Tompkins-Macdonald and Leys 2008), which may have an impact on the oxygenation status inside the sponge tissue and thus on the microbial community.

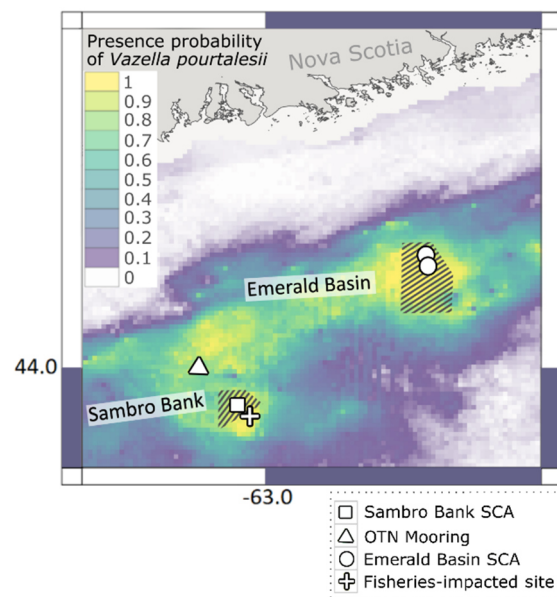


FIGURE 1 | Map of sampling region on the Scotian Shelf, Canada. Colours depict the probability of *V. pourtalesii* occurrence based on the data presented in Beazley et al. (2018). Yellow indicates areas where the probability of occurrence is highest. The Sambro Bank and Emerald Basin SCAs are depicted by stripes. Sampling locations of *V. pourtalesii* in the Emerald Basin SCA are indicated by white dots and the white squares at the Sambro Bank SCA. The white triangle depicts the position of the OTN mooring and the white cross represents samples from the fisheries-impacted site

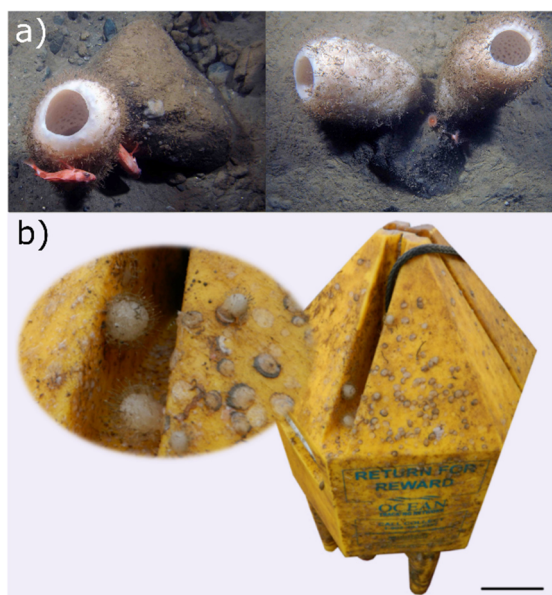


FIGURE 2 | **a** In situ photographs of *V. pourtalesii* with individuals being up to around 20 cm in height. **b** Close-up of small *V. pourtalesii* growing on a mooring float of the Ocean Tracking Network (photo credits: DFO). Average size of small *V. pourtalesii* growing on the mooring was approximately 1 cm. Scale bar at right bottom represents 10 cm and applies only to the whole mooring float in Fig. 2b

Several recent studies have explored the response of sponges and their microbiomes (hereafter referred to as ‘holobionts’) to anthropogenic stressors, such as ocean warming, acidification, eutrophication, sedimentation and pollution (reviewed in Pita et al. (2018); Slaby et al. (2019)). The increase in artificial substrates such as plastics in the oceans represents another anthropogenic-induced impact that could potentially adversely affect sponge holobionts. Especially in shelf areas, petroleum platforms, off-shore wind parks and increased loads of litter may represent new settlement opportunities for sponge larvae or gemmules. Artificial substrates have been observed to be exposed to microfouling and biofilm development (Balqadi et al. 2018). It is thus conceivable that biofilms on artificial substrates may influence the sponge holobiont settlement and development on such substrates. The recent discovery that *V. pourtalesii* uses oceanographic moorings (belonging to the

Halifax Line of the Ocean Tracking Network (OTN)) as settlement substrate enabled the testing of this hypothesis. The present study aims to describe and compare the microbial community composition of the glass sponge *V. pourtalesii* collected from protected environments (i.e., within the Sambro Bank and Emerald Basin SCAs), a fisheries-impacted site immediately adjacent to the border of the Sambro Bank SCA, as well as from an artificial substrate (plastic mooring float cover). This study contributes to a deeper understanding of how marine sponge holobionts respond to human pressures.

METHODS

Field work

Samples were collected during two oceanographic research missions to the *Vazella* sponge grounds on the Scotian Shelf, off Eastern Canada. The first mission took place in July/August 2016 onboard Canadian Coast Guard Ship (CCGS) *Hudson* (*Hudson2016-019*; Kenchington et al. 2017), where Oceaneering’s Spectrum remotely operated vehicle (ROV) and a mega-box corer were used to collect *V. pourtalesii* from the Emerald Basin SCA (protected site) between 184 and 206 m depth. The second mission was conducted in September 2017 onboard CCGS *Martha L. Black* (*MLB2017001*; Beazley et al. 2017), where sponge samples were collected during multiple deployments of the ROV ROPOS (Canadian Scientific Submersible Facility, Victoria, Canada) between 153 and 161 m depth at a single station near the centre of the Sambro Bank SCA (protected site). Furthermore, during this mission, sponge samples were collected outside the closure immediately adjacent to its southeast border at 198 m depth, in an area that was subjected to relatively recent bottom trawling activity (indicated by the presence of deep trawl marks in the general vicinity of the collection site). In total 33 sponges were

collected: eight and 13 individuals from the Emerald Basin and Sambro Bank SCAs (protected sites), and four from the fisheries-impacted site adjacent to the Sambro Bank SCA. Additionally, eight specimens were collected from an Ocean Tracking Network (OTN) mooring located approximately 10 km northwest of the Sambro Bank SCA (for map of sampling locations see **Fig. 1**). The mooring (for a picture of the OTN mooring see **Fig. 2b**) was anchored ~ 5 m above the seabed and was deployed for ~ 13 months (15th of August 2016–5th of September 2017) prior to its recovery. Sample metadata were deposited in the Pangaea database.

From each *V. pourtalesii* specimen, pieces of sponge biomass were subsampled and rinsed by transferring them three times through sterile filtered seawater. Excess liquid was then removed, the samples were flash frozen, kept continuously frozen during transportation on dry-ice, and stored at – 80 °C until DNA extraction. For the mooring samples, complete sponge individuals were frozen at – 80 °C due to their small size. Seawater samples (2 L) were collected (concomitantly with sponge sampling) at every location in quadruplicates and each filtered onto PVDF filter membranes (Merck Millipore) with a pore size of 0.22 µm and a diameter of 47 mm. Filters were stored at – 80 °C until DNA extraction. Sediments were collected using ROV push corers or box corer deployments at the same locations where sponges and seawater were collected contemporaneously. The top of each core was sliced off, the upper 2 cm of the sediment was collected and stored at – 80 °C until DNA extraction. Biofilms were scraped off the mooring float (where no visible encrusting fauna was present) using cotton swabs (Carl Roth, Karlsruhe, Germany), placed into sterile tubes and frozen at – 80 °C. All sponge, seawater, sediment and biofilm samples covered in this study originated from the same water mass (discussed in Beazley et al. 2018). Throughout this study we use the term ‘environmental references’ for seawater and sediment microbial

communities. For assessing the impact of fisheries and artificial substrate provision on *V. pourtalesii* microbial communities, we use Sambro Bank protected sponges as the baseline for comparison.

Amplicon sequencing DNA extraction was conducted on ~ 0.25 g of sponge tissue, ~ 0.25 g of sediment, half a seawater filter or the complete cotton woolen part of a swab using the DNeasy Power Soil Kit (Qiagen, Venlo, The Netherlands). Extracted DNA was quantified using Qubit fluorometer measurements and their quality assessed by a polymerase chain reaction (PCR) with the universal 16S rRNA gene primers 27F + 1492R and subsequent gel electrophoresis on 1% agarose. The V3 and V4 variable regions of the 16S rRNA gene were amplified in a one-step PCR using the primer pair 341F-806R in a dual-barcoding approach (Kozich et al. 2013). The nucleotide sequences of the primers are as follows: 5'-CCTACGG-GAGGCAGCAG-3' (Muyzer et al. 1993) and 5'-GGACTA CHVGGGTWTCTAAT-3' (Caporaso et al. 2011). PCR-products were verified by gel electrophoresis, normalised and pooled. Sequencing was performed on a MiSeq platform (MiSeqFGx, Illumina, San Diego, USA) using v3 chemistry (producing 2 × 300 bp). Demultiplexing after sequencing was based on 0 mismatches in the barcode sequences. Raw reads were archived in the NCBI Sequence Read Archive.

Bioinformatics and statistics For analysis of sequencing data, raw sequences were first quality-filtered using BBDOUK (BBMAP version 37.75; Bushnell 2017). Quality trimming was conducted on both ends of the reads (removal of first and last 13 nt) with Q20 and a minimum length of 250 nt. Quality of sequences was evaluated with FastQC (version 0.10.1; Andrews 2010) and output aggregated with MultiQC (version 0.9; Ewels et al. 2016). The post-filtered sequences were processed with QIIME2 (versions 2018.6 and 2018.8;

Bolyen et al. 2019). The DADA2 algorithm (Callahan et al. 2016) was used, which retains Amplicon Sequence Variants (ASVs). Reads were denoised and consensus removal of chimera was conducted. One million reads were used to train the error model. Chloroplast and mitochondrial sequences were removed from further analyses. Assignment of taxonomy was conducted using a Naive Bayes classifier (Bokulich et al. 2018) trained on the Silva 132 99% OTUs 16S database (Quast et al. 2013). We used the term ‘unclassified’ which is defined as ‘no hit’ in this taxonomic classification. This information was used to identify potentially novel microbial diversity in *V. pourtalesii*. Ambiguous classifications (such as ‘unidentified bacteria’) are included in the ‘known sequences’ fraction as those sequence have been recorded and deposited before.

FastTree2 (Price et al. 2010) was used to generate a phylogeny based on ASVs. Subsequently, weighted UniFrac distances (phylogeny-based β -diversity metrics) were calculated (Lozupone et al. 2005). Non-metric multidimensional scaling (NMDS) was performed on weighted UniFrac distances to visually assess sample separation in ordination space. A permutational multivariate analysis of variance (PERMANOVA) was conducted on weighted UniFrac distances to determine whether groups of samples were significantly different from one another. Pairwise tests were performed between all pairs of groups (applied significant level $\alpha = 0.05$). Multivariate analyses were run inside QIIME2 and R.

In addition to phylogeny-based metrics, quantitative measures of community richness were also calculated (e.g. Shannon’s Diversity Index). The Linear Discriminant Analysis (LDA) Effect Size (LEfSe) algorithm (Segata et al. 2011) was applied to determine if microbial phyla differed significantly among groups and rank them according to estimated effect sizes of the significant variations. To accomplish this, first a factorial Kruskal–Wallis sum-rank test was performed on feature tables to

detect features (i.e. microbial taxa) with significant differential abundance between groups (applied significant level $\alpha = 0.05$). Wilcoxon rank-sum tests were then run to perform pairwise tests among subgroups (applied significant level $\alpha = 0.05$). Finally, a LDA estimated the effect size of each differentially abundant feature (applied threshold on the logarithmic LDA score for discriminative features = 2.0).

The majority of plots were produced using R statistical software version 3.0.2 (R Development Core Team 2008) and arranged using Inkscape (version 0.92.3; Harrington and Team 2005). The *V. pourtalesii* presence probability raster (Fig. 1) was visualised with QGIS (version 2.18.4; QGIS Development Team 2017).

RESULTS

Microbial diversity of *V. pourtalesii*

In the present study, we assessed the microbial community structure of 33 *V. pourtalesii* individuals as well as environmental reference samples by 16S rRNA gene amplicon sequencing. The microbiomes of the *V. pourtalesii* sponges clustered together, and were distinct from seawater and sediment microbiomes (Fig. 3). Overall, the *V. pourtalesii* microbiomes showed a higher variability of their α -diversity than the environmental reference microbiomes (seawater, biofilm, seawater). Rarefaction curves (α -diversity) revealed further that microbial richness (average number of ASVs per sample) was highest in the biofilm samples collected from the OTN mooring, second highest in seawater, and lowest in *V. pourtalesii* samples (Online Resource 1: S1). Also, in terms of diversity indices (Shannon index), alpha diversity was lowest in the *V. pourtalesii* microbiomes. While Proteobacteria were the most dominant phylum in all sample types (sponge: 78.8% of total; seawater: 47.4%; sediment: 54.7%; and biofilm: 56.6%), Patescibacteria were particularly abundant members of the

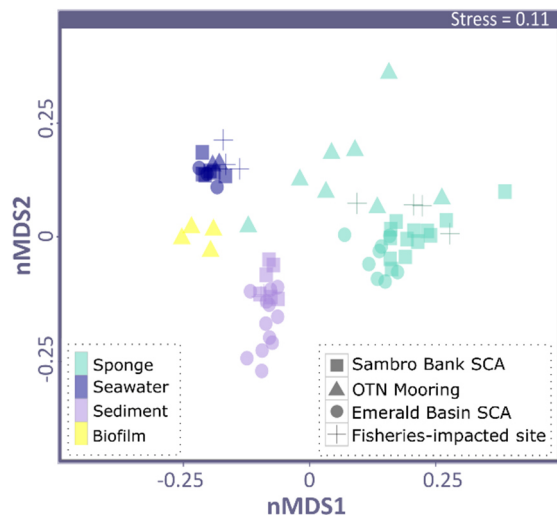


FIGURE 3 | Non-metric Multidimensional Scaling (NMDS) plot on weighted UniFrac distances. Each marker represents a single microbial community. Symbols represent sampling location: squares show samples from Sambro Bank, circles represent samples from Emerald Basin, triangles are mooring samples and crosses are samples from the fisheries impacted site. Colors depict sample type: green marks *V. pourtalesii* samples, blue marks seawater samples, yellow is for mooring biofilm and red represents sediment samples

V. pourtalesii microbiome (4.6%), followed by Bacteroidetes (3.4%), Spirochaetes (2.9%), and Planctomycetes (2.8%), (**Fig. 4**). Despite a distinct microbial fingerprint on different taxonomic levels (e.g., at ASV-level, as well as phylum-level), *V. pourtalesii* microbiomes showed a higher variability in terms of their α -diversity than environmental microbiomes. We further observed a higher proportion of previously unrecorded microbial taxa at every hierarchical taxonomic level (microbial phylum down to microbial order level) in the microbiomes of *V. pourtalesii* than in the seawater samples (**Table 1**). At the genus level, 40.2% of the amplicon sequences of *V. pourtalesii* could not be taxonomically assigned (compared to 16.2% in seawater). The largest increase of unclassified reads was observed from class to order level (from 0.4% to 31.6% respectively). In comparison to environmental microbiomes (seawater,

sediment, and biofilm) *V. pourtalesii* microbiomes were particularly enriched in Proteobacteria and Patescibacteria (**Fig. 5a**).

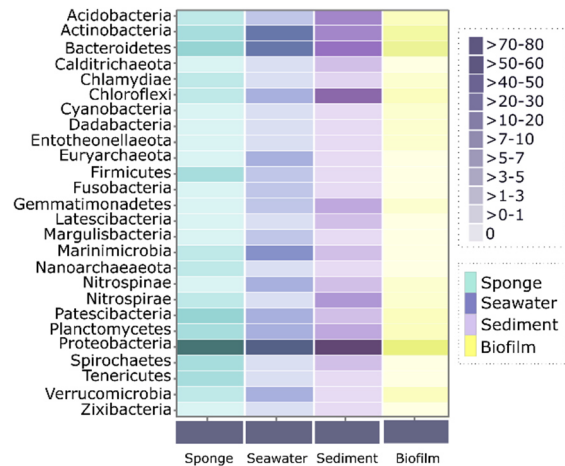


FIGURE 4 | Heatmap showing the relative abundances [%] of microbial phyla per sample type: sponge, seawater, sediment and mooring biofilm

Table 1 | Average fractions of unclassified 16S rRNA gene sequences in *V. pourtalesii* and in seawater at the different hierarchical taxonomic levels indicating a high level of novelty in particularly sponge associated microbial communities

Taxonomic level	Unclassified sequences [%] in <i>Vazella pourtalesii</i>	Unclassified sequences [%] in seawater
Phylum	0.3	0.1
Class	0.4	0.3
Order	31.6	1.5
Family	36.1	10.9
Genus	40.2	16.2

Response of the *V. pourtalesii* microbiome to anthropogenic activities

We then assessed the changes of the sponge microbiome upon human activity, including also implemented protection efforts in form of sponge conservation areas. A significant difference in *V. pourtalesii* microbiomes between the Sambro Bank and Emerald Basin SCAs was observed (**Table 2**), with several microbial phyla significantly enriched in one site over the other (**Fig. 5b**). For example, Patescibacteria were significantly enriched in Sambro Bank

sponges, while Proteobacteria were significantly enriched in Emerald Basin sponges. We further compared inter-location variability (i.e., *V. poutalesii* individuals between the Emerald

Basin SCA and Sambro Bank SCA) against intra-location variability (i.e., individuals collected from within Emerald Basin SCA only). While the inter-location variability was significant, the

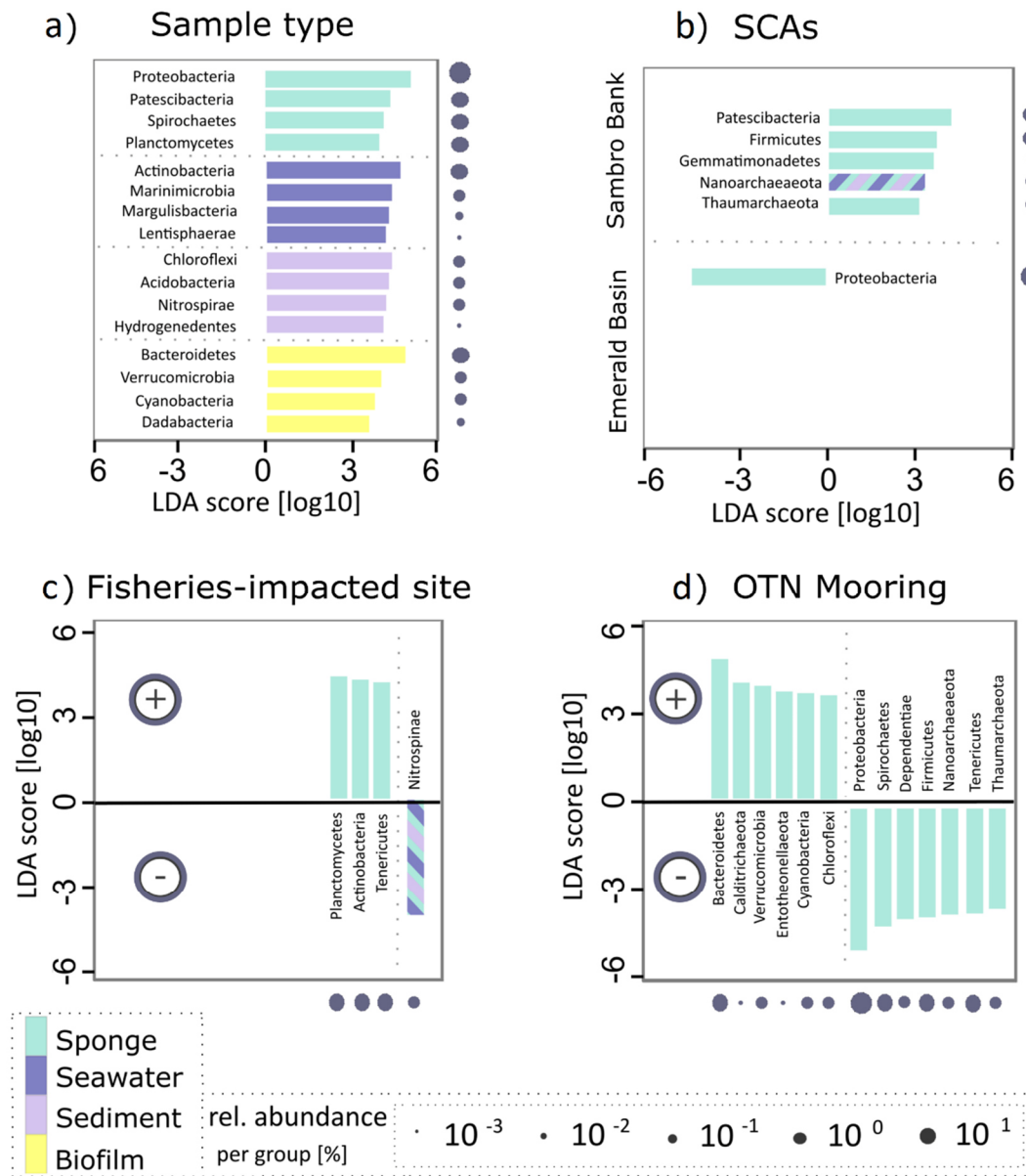


FIGURE 5 | Linear discriminant analysis (LDA) Effect Size (LEfSe) plots. Analysis was performed on microbial phylum level, the rank in the plot is given according to effect size. a Shows data for different sample types (sponge, seawater, sediment and mooring biofilm samples). Here, only the four most significantly different phyla are shown per group. For b-d, where only sponge samples were inspected, all phyla are shown. b Shows differences between sponge conservation areas (SCAs), Sambro Bank and Emerald Basin. c-d) Show microbial taxa increased or decreased upon anthropogenic pressures. c Shows protected vs. fisheries-impacted sites and d shows growth of sponges in their natural habitat in sponge grounds vs moorings). For all plots, filled bars show microbial taxa with significant effects only in sponge samples; striped bars additionally indicate that these taxa have a significant effect for all three sample types (sponge, seawater and sediment). Dots next to or below plots indicate relative abundance of taxa within the *V. poutalesii* microbial community. The respective scale is indicated at the lower part of the illustration

TABLE 2 | Pairwise comparisons of beta-diversity are shown for the microbial communities across different sample types, sampling locations, and anthropogenic impacts. Technically, pseudo-F and p-values of pairwise PERMANOVAs on weighted UniFrac matrices are shown

Group1	Group2	Pseudo-F	p-value
SambroBankSCA_Sponge	SambroBankSCA_Seawater	47.65	<0.01
SambroBankSCA_Sponge	SambroBankSCA_Sediment	29.76	<0.01
SambroBankSCA_Seawater	SambroBankSCA_Sediment	31.35	<0.01
EmeraldBasinSCA_Sponge	EmeraldBasinSCA_Seawater	35.62	0.01
EmeraldBasinSCA_Sponge	EmeraldBasinSCA_Sediment	45.33	<0.01
EmeraldBasinSCA_Seawater	EmeraldBasinSCA_Sediment	26.87	<0.01
SambroBankSCA_Sponge	OTNMooring_Sponge	7.86	<0.01
Sponge	OTNMooring_Biofilm	26.37	<0.01
Seawater	OTNMooring_Biofilm	18.16	<0.01
Sediment	OTNMooring_Biofilm	14.14	<0.01
SambroBankSCA_Sponge	EmeraldBasinSCA_Sponge	6.81	<0.01
SambroBankSCA_Seawater	EmeraldBasinSCA_Seawater	1.57	0.11
SambroBankSCA_Sediment	EmeraldBasinSCA_Sediment	5.88	<0.01
SambroBankSCA_Sponge	FisheriesImpactedSite_Sponge	1.67	0.12
SambroBankSCA_Seawater	FisheriesImpactedSite_Seawater	2.75	0.02
SambroBankSCA_Sediment	FisheriesImpactedSite_Sediment	1.73	0.07
FisheriesImpactedSite_Sponge	FisheriesImpactedSite_Seawater	11.76	0.03
FisheriesImpactedSite_Sponge	FisheriesImpactedSite_Sediment	10.57	0.08
FisheriesImpactedSite_Seawater	FisheriesImpactedSite_Sediment	9.83	0.09

intra-location variability was not (PERMANOVA, Online Resource 1: Table S1, mean weighted UniFrac distance = 0.29). Further, a smaller mean weighted UniFrac distance was observed between sponges of the same location (Emerald Basin SCA), compared to sponges at different sampling locations (Emerald Basin SCA vs. Sambro Bank SCA). In terms of trawling impact, certain microbial taxa were depleted or enriched in *V. pourtalesii* individuals originating from the trawled site compared to the protected sites (Fig. 5c). For example, Nitrospinae were significantly depleted in *V. pourtalesii* individuals originating from the trawled site adjacent to the Sambro Bank SCA (in both sponges as well as environmental references). Nitrospinae are nitrite-oxidising bacteria and have previously been found in deep-sea glass sponges (Tian et al. 2016). With respect to sponge growth on artificial substrates

(moorings), their microbial community composition differed significantly from sponges growing in their natural habitat (Table 2). Sponges from the artificial mooring substrate were significantly enriched in Bacteroidetes, Verrucomicrobia and Cyanobacteria which were also enriched in the mooring biofilm (Fig. 5d).

DISCUSSION

This study is the first to report on the microbiome composition of the glass sponge *Vazella pourtalesii*, that forms pronounced monospecific aggregations with high densities on the Scotian Shelf, Canada (Fuller 2011; Beazley et al. 2018) that are not found elsewhere in the world's oceans. Our results suggest that *V. pourtalesii* contains its own distinct microbial community that is different from seawater and sediment samples. The observed pattern of lower alpha diversity in *V. pourtalesii* compared to environmental reference samples is in line with previous observations showing that sponge-associated microbial communities are

usually less complex than those of seawater or sediments (Thomas et al. 2016). We propose that *V. pourtalesii* represents a rich reservoir of novel microbial taxa, even at high taxonomic ranks (e.g. class-level: 31.6% unclassified). The high abundance of Patescibacteria in *V. pourtalesii* is particularly striking when compared to the microbiomes of other glass sponges (Tian et al. 2016) and also when compared to the environmental reference samples in our study. This microbial phylum is generally assumed to depend on symbiotic animal hosts to cover their basic metabolic requirements (Castelle et al. 2018).

***V. pourtalesii* microbial community composition varies between the Sambro Bank and Emerald Basin Sponge Conservation Areas**

Compared to the conservation efforts towards the protection of animal and plant biodiversity, comparably little is currently known about the needs to protect microbial biodiversity. This is striking as Webster et al. (2018) and others point out very clearly that microorganisms underpin ecosystem health and hence should deserve committed conservation and management endeavors. In our study, we explored the variation in microbial community composition residing in sponges from two SCAs. We observed that the variations in sponge-associated microbiomes were smaller within sites than between sites. We cannot rule out whether a temporal effect may have led to the observed differences between sites, as all Emerald Basin samples were collected in 2016, while all other samples were collected in 2017. However, since no significant differences in seawater microbial communities between both years were observed (**Table 2**), we deduce that temporal effects were not detectable in the water surrounding the sponges. Our results of variable microbial community compositions between Emerald Basin SCA *V. pourtalesii*s and Sambro Bank SCA *V. pourtalesii*s illustrate that fisheries closures are important for the conservation of not only macro-, but also microbial biodiversity.

Deviations of the *V. pourtalesii* microbial community composition at a trawled site

We observed significantly different relative abundances of distinct microbial phyla from bottom-trawled areas, such as a reduction of Nitrospinae in sponges and reference samples. On the other hand, Planctomycetes, Actinobacteria as well as Tenericutes were significantly enriched in *V. pourtalesii* individuals originating from the fisheries-impacted site, but not in the seawater or other samples taken from these sites. Although we observed significant enrichments or depletions of individual microbial phyla, the overall *V. pourtalesii* microbial community composition did not differ significantly between the protected and fisheries-impacted site in a permanova analyses based on weighted UniFrac distances. As outcomes of permanova analyses strongly depend on sample sizes, we cannot exclude a technical issue in this regard (as $n = 4$). However, these results are in line with Luter et al. (2012) who observed no shifts in the microbial community of the Great Barrier reef sponge *Ianthella basta* associated with increased sedimentation, when considering the total microbial community rather than individual microbial taxa. In another study on five shallow water sponge species, again no significant effect of increased sediment loads was observed, while minor effects included an increased reliance on phototrophic feeding under high suspended sediment loads (Pineda et al. 2016). Future studies with targeted sampling, higher replication, and wet-lab experimentation are needed to fully assess the impact of trawling on deep-sea sponges and their microbiomes.

While seawater and sediment microbiomes differed significantly from each other at the protected sites in our study, microbial community composition was similar for the two sample types at the fisheries-impacted site. Seawater microbial community composition was significantly different at the fisheries-impacted site in comparison to the protected site, while the sediment microbial community was similar at both sites. From these two

observations we conclude that trawling might have an effect on the seawater microbial community in a way that the seawater community becomes more similar to that of sediments. We posit that this may be due to direct and indirect effects of trawling-induced sediment suspension. While a direct translocation of microbial community members may occur from the sediment to the seawater, sediment resuspension may also lead to an increased eutrophication, thus promoting a rise of opportunistic bacterial clades in response to the release of biogeochemical elements (such as e.g., phosphorous) from the sediment into the water column.

Deviations of the *V. pourtalesii* microbial community composition growing on an artificial substrate

Artificial structures have the potential to act as stepping stones, enabling species to broaden their distributions in the marine environment (Adams et al. 2014). Our results suggest that the biofilms on artificial substrates may 'prime' the sponge-associated microbial communities when small sponges settle on such substrates. Our finding of a different microbial community in mooring *V. pourtalesii* individuals compared to protected ones raises the possibility of these structures changing the microbial diversity of adjacent and newly established benthic populations seeded from such sources. Individuals of *V. pourtalesii* growing on the artificial substrate showed a higher degree of microbiome variation compared to those collected from their natural habitat (Online Resource 1: S2). This variation may be due to the fact that sponges growing on the moorings were still small and had presumably settled within a maximum of 13 months following mooring deployment, and thus their microbiomes may not yet have been fully established.

Conclusions

In the present study two anthropogenic impacts, exposure to trawling activity and growth on plastic moorings, on the *V. pourtalesii* microbiomes were explored. In addition, we evaluated the role of two fisheries closures in

preserving sponge-associated microbial diversity on the Scotian Shelf, Canada. We conclude first and foremost, that the Emerald Basin and Sambro Bank Sponge Conservation Areas are ecologically important, as they contain distinct and largely unclassified benthic microbial communities found within sponges or sediments. Variability of the microbial community was comparably high between *V. pourtalesii* individuals, pointing towards the need to minimise sponge loss through anthropogenic pressures (e.g., direct trawling) in order to protect the full range of microbial biodiversity. We further observed significant quantitative differences (in terms of relative abundances) in distinct microbial phyla, such as a reduction in Nitrospinae in sponges and samples from the surrounding environment in areas subjected to bottom trawling. With respect to sponges growing on plastic moorings, significant differences in their microbiome were observed compared to sponges growing on natural substrates. *V. pourtalesii* individuals growing on a mooring surface were significantly enriched in Bacteroidetes, Verrucomicrobia and Cyanobacteria compared to those growing on natural substrates. This could be explained by either the plastic settlement substrate or the younger age of sponges growing on the moorings. Importantly, *V. pourtalesii* is a rich and unique reservoir of microbial biodiversity that deserves conservation effort.

ACKNOWLEDGEMENTS

Open Access funding provided by Projekt DEAL. This manuscript is written in memory of Hans Tore Rapp, whose research efforts were dedicated to the understanding and preservation of deep-sea sponge grounds in the North Atlantic Ocean. We thank the crews and scientific parties of the two research cruises *Hudson2016-019* and *MLB2017001*. Ulrike Hanz and Furu Mienis are acknowledged for support during field work. We further thank Andrea Hethke, Ina Clefsen, and the CRC1182 Z3 team (Katja Cloppenborg-Schmidt, Malte Rühlemann, John Baines) for assistance with the amplicon pipeline.

AUTHOR CONTRIBUTIONS

KB, UH and BS designed the study. EK, LB, FW organised the cruises and provided samples. KB performed material preparations and data analyses (bioinformatics and visualisation). The first draft of the manuscript was written by KB and UH. All authors commented on previous versions of the manuscript and approved the final version.

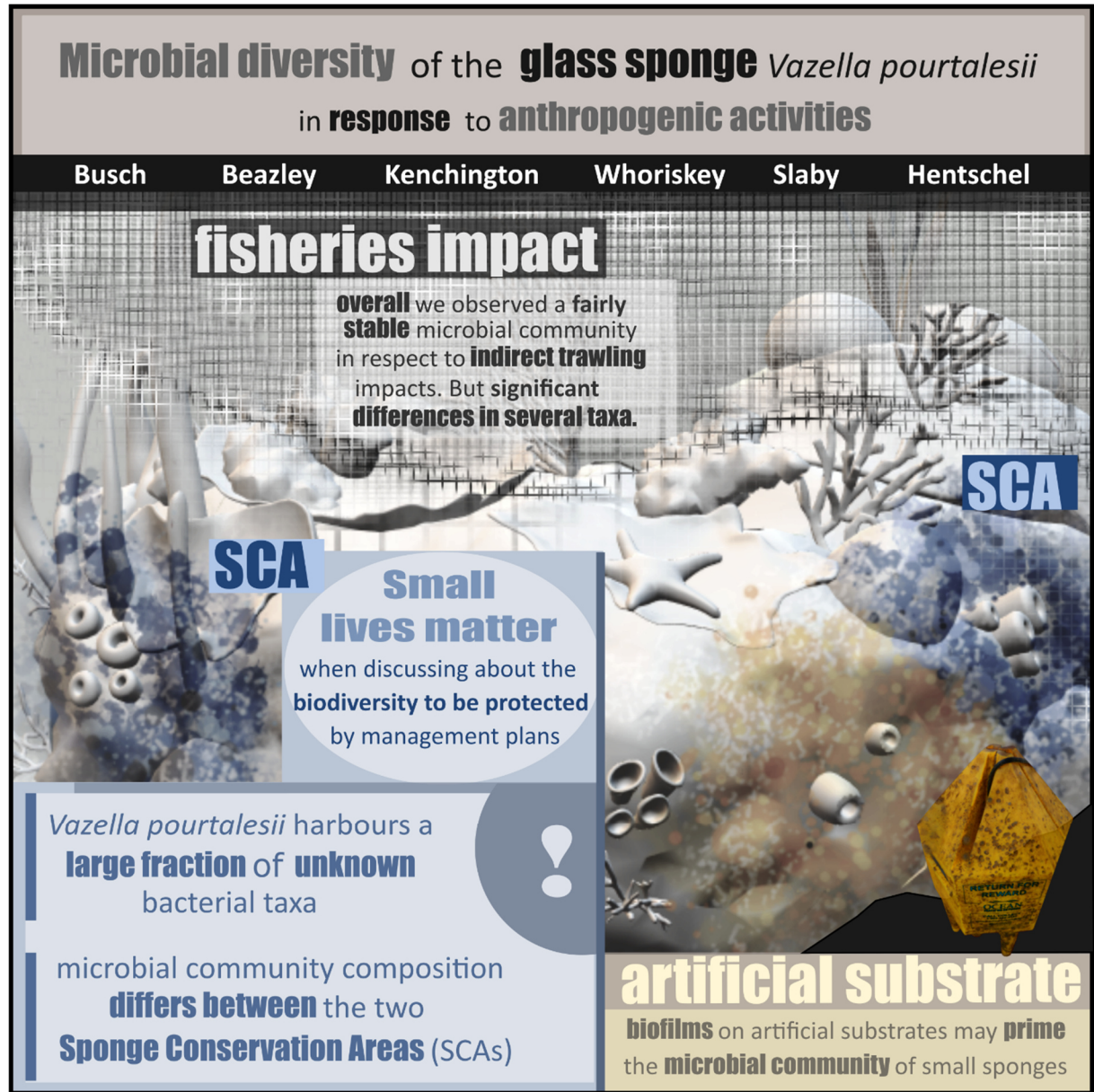
FUNDING

This study was funded by the European Union's Horizon 2020 research and innovation program under Grant Agreement No. 679849 ('SponGES'). This document reflects only the authors' view and the Executive Agency for Small and Medium-sized Enterprises (EASME) is not responsible for any use that may be made of the information it contains. Canadian cruises and contributions were funded by Fisheries

REFERENCES

- Adams TP, Miller RG, Aleynik D, Burrows MT (2014) Offshore marine renewable energy devices as stepping stones across biogeographical boundaries. *J Appl Ecol* 51(2):330–338
- Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. Available online at: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Balqadi AA, Salama AJ, Sathesh S (2018) Microfouling development on artificial substrates deployed in the central Red Sea. *Oceanologia* 60:219–231
- Beazley L, Pham C, Murillo J, Kenchington E (2017) Cruise report for the DFO/SponGES CCGS Martha L. Black oceanographic mission (MLB2017001), August 31 to September 7, 2017. Canadian Technical Report of Fisheries and Aquatic Sciences 3242: vi+42.
- Beazley L, Wang Z, Kenchington E, Yashayaev I, Rapp HT, Xavier JR et al (2018) Predicted distribution of the glass sponge *Vazella pourtalesi* on the Scotian Shelf and its persistence in the face of climatic variability. *PLoS One* 13:e0205505
- Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R et al (2018) Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* 6:1–17
- Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA et al (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 37:852–857. <https://doi.org/10.1038/s41587-019-0209-9>
- Bordenstein SR, Theis KR (2015) Host biology in light of the microbiome: ten principles of holobionts and hologenomes. *PLoS Biol* 13:e1002226
- Bosch TCG, McFall-Ngai MJ (2011) Metaorganisms as the new frontier. *Zoology* 114:185–190
- Bushnell B (2017) BBMap short read aligner, and other bioinformatic tools. Available online at: <https://sourceforge.net/projects/bbmap/>.
- Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* 13:581–583
- Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Knight R, Lozupone CA et al (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci* 108:4516–4522
- Castelle CJ, Brown CT, Anantharaman K, Probst AJ, Huang RH, Banfield JF (2018) Biosynthetic capacity, metabolic variety and unusual biology in the CPR and DPANN radiations. *Nat Rev Microbiol* 16:629–645
- DFO (2017) Delineation of significant areas of coldwater corals and sponge-dominated communities in Canada's Atlantic and Eastern Arctic marine waters and their overlap with fishing activity. DFO Can Sci Advis Sec Sci Advis Rep 2017:007
- Ewels P, Magnusson M, Lundin S, Käller M (2016) MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* 32:3047–3048
- Fuller S (2011) Diversity of marine sponges in the Northwest Atlantic. PhD Thesis, Dalhousie Univ 1–229.
- Harrington B and Team, and the D. (2005) Inkscape. Available online at: <https://www.inkscape.org/>.
- Jackson JBC, Kirby MX, Berger WH, Bjorndal KA, Botsford LW, Bourque BJ et al (2001) Historical overfishing and the recent collapse of coastal ecosystems. *Science* (80–) 629:1–17
- Kahn AS, Chu JWF, Leys SP (2018) Trophic ecology of glass sponge reefs in the Strait of Georgia, British Columbia. *Sci Rep* 8:1–11
- Kahn AS, Leys SP (2017) Spicule and flagellated chamber formation in a growth zone of *Aphrocallystes vastus* (Porifera, Hexactinellida). *Invertebr Biol* 136:22–30
- Kelleher K (2005) Discards in the world's marine fisheries. An update. FAO Fisheries Technical Paper. 470.
- Kenchington E, Beazley L, Yashayaev I (2017) Hudson 2016–2019 International deep sea science expedition cruise report. Canadian Data Report of Fisheries and Aquatic Sciences 1277: v + 55p.
- Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD (2013) Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. *Appl Environ Microbiol* 79:5112–5120
- Leys SP, Mackie GO, Reiswig HM (2007) The biology of glass sponges. *Adv Mar Biol* 52:1–145
- Lozupone C, Knight R, Lozupone C, Knight R (2005) UniFrac: a new phylogenetic method for comparing microbial communities. *Appl Environ Microbiol* 71:8228–8235
- Luter HM, Whalan S, Webster NS (2012) Thermal and sedimentation stress are unlikely causes of brown spot syndrome in the coral reef sponge, *Ianthella basta*. *PLoS One* 7:1–9

- McFall-Ngai M, Hadfield MG, Bosch TCG, Carey HV, Domazet-Lošo T, Douglas AE et al (2013) Animals in a bacterial world, a new imperative for the life sciences. *Proc Natl Acad Sci USA* 110:3229–3236
- Moitinho-Silva L, Nielsen S, Amir A, Gonzalez A, Ackermann GL, Cerrano C, et al (2017) The sponge microbiome project. *Gigascience* 6(10):gix077
- Muyzer G, Waal ECDE, Uitterlinden AG (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* 59:1–6
- Ortuño Crespo G, Dunn DC (2017) A review of the impacts of fisheries on open-ocean ecosystems. *ICES J Mar Sci* 74:2283–2297
- Pineda M-C, Strehlow B, Duckworth A, Doyle J, Jones R, Webster NS (2016) Effects of light attenuation on the sponge holobiont-implications for dredging management. *Sci Rep* 6:39038
- Pita L, Rix L, Slaby BM, Franke A, Hentschel U (2018) The sponge holobiont in a changing ocean: from microbes to ecosystems. *Microbiome* 6:46
- Price MN, Dehal PS, Arkin AP (2010) FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490
- QGIS Development Team (2017) Geographic information system. Open source geospatial foundation project. Available online at: <https://qgis.osgeo.org>.
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P et al (2013) The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:590–596
- R Development Core Team (2008) R: a language and environment for statistical computing.
- Rohwer F, Seguritan V, Azam F, Knowlton N (2002) Diversity and distribution of coral-associated bacteria. *Mar Ecol Prog Ser* 243:1–10
- Savoca S, Lo Giudice A, Papale M, Mangano S, Caruso C, Spanò N et al (2019) Antarctic sponges from the Terra Nova Bay (Ross Sea) host a diversified bacterial community. *Sci Rep* 9:1–15
- Schmidt O (1870) Grundzüge einer Spongien-Fauna des atlantischen Gebietes. Wilhelm Engelmann, Leipzig, pp 14–15
- Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C (2011) Metagenomic biomarker discovery and explanation. *Genome Biol* 12:R60
- Slaby B, Franke A, Rix L, Pita L, Bayer K, Jahn M, Hentschel U (2019) Marine sponge holobionts in health and disease. In: Li Z (ed) *Symbiotic microbiomes of coral reefs sponges and corals*. Springer, Netherlands, Dordrecht, pp 81–104
- Steinert G, Busch K, Bayer K, Khodami S, Arbizu PM, Kelly M, et al (2020) Compositional and quantitative insights into bacterial and archaeal communities of South Pacific deep-sea sponges (Demospongiae and Hexactinellida). *Front Microbiol* 11:716
- Thomas T, Moitinho-Silva L, Lurgi M, Björk JR, Easson C, Astudillo-García C et al (2016) Diversity, structure and convergent evolution of the global sponge microbiome. *Nat Commun* 7:11870
- Tian RM, Sun J, Cai L, Zhang WP, Zhou GW, Qiu JW, Qian PY (2016) The deep-sea glass sponge *Lophophysema eversa* harbours potential symbionts responsible for the nutrient conversions of carbon, nitrogen and sulfur. *Environ Microbiol* 18:2481–2494
- Tompkins-Macdonald GJ, Leys SP (2008) Glass sponges arrest pumping in response to sediment: implications for the physiology of the hexactinellid conduction system. *Mar Biol* 154:973–984
- Webster NS, Wagner M, Negri AP (2018) Microbial conservation in the Anthropocene. *Environ Microbiol* 20:1925–1928



Chapter 4

Microbial strategies for survival in the glass sponge *Vazella pourtalesii*

© 2020 Bayer, Busch et al.

Original publication by *mSystems* (American Society for Microbiology)

Reprinted under the terms of the Creative Commons Attribution License (CC-BY 4.0).

doi: [10.1128/mSystems.00473-20](https://doi.org/10.1128/mSystems.00473-20).

Microbial strategies for survival in the glass sponge *Vazella pourtalesii*

Kristina Bayer,^a Kathrin Busch,^a Ellen Kenchington,^b Lindsay Beazley,^b Sören Franzenburg,^c Jan Michels,^d Ute Hentschel,^{a,e} Beate M. Slaby^a

^aGEOMAR Helmholtz Centre for Ocean Research Kiel, RD3 Marine Symbioses, Kiel, Germany, ^bDepartment of Fisheries and Oceans, Bedford Institute of Oceanography, Dartmouth, Nova Scotia, Canada, ^cInstitute for Clinical Molecular Biology, Kiel University, Kiel, Germany, ^dFunctional Morphology and Biomechanics, Institute of Zoology, Kiel University, Kiel, Germany, ^eKiel University, Kiel, Germany.

Kristina Bayer and Kathrin Busch are co-first authors and contributed equally to the manuscript. As the main focus lies on genomic interpretation, Kristina Bayer is named first.

Few studies have explored the microbiomes of glass sponges (Hexactinellida). The present study seeks to elucidate the composition of the microbiota associated with the glass sponge *Vazella pourtalesii* and the functional strategies of the main symbionts. We combined microscopic approaches with metagenome-guided microbial genome reconstruction and amplicon community profiling toward this goal. Microscopic imaging revealed that the host and microbial cells appeared within dense biomass patches that are presumably syncytial tissue aggregates. Based on abundances in amplicon libraries and metagenomic data, SAR324 bacteria, *Crenarchaeota*, *Patescibacteria*, and *Nanoarchaeota* were identified as abundant members of the *V. pourtalesii* microbiome; thus, their genomic potentials were analyzed in detail. A general pattern emerged in that the *V. pourtalesii* symbionts had very small genome sizes, in the range of 0.5 to 2.2 Mb, and low GC contents, even below those of seawater relatives. Based on functional analyses of metagenome-assembled genomes (MAGs), we propose two major microbial strategies: the “givers,” namely, *Crenarchaeota* and SAR324, heterotrophs and facultative anaerobes, produce and partly secrete all required amino acids and vitamins. The “takers,” *Nanoarchaeota* and *Patescibacteria*, are anaerobes with reduced genomes that tap into the microbial community for resources, e.g., lipids and DNA, likely using pilus-like structures. We posit that the existence of microbial cells in sponge syncytia together with the low-oxygen conditions in the seawater environment are factors that shape the unique compositional and functional properties of the microbial community associated with *V. pourtalesii*.

IMPORTANCE

We investigated the microbial community of *V. pourtalesii* that forms globally unique, monospecific sponge grounds under low-oxygen conditions on the Scotian Shelf, where it plays a key role in its vulnerable ecosystem. The microbial community was found to be concentrated within biomass patches and is dominated by small cells (<1 μm). MAG analyses showed consistently small genome sizes and low GC contents, which is unusual compared to known sponge symbionts. These properties, as well as the (facultatively) anaerobic metabolism and a high degree of interdependence between the dominant symbionts regarding amino acid and vitamin synthesis, are likely adaptations to the unique conditions within the syncytial tissue of their hexactinellid host and the low-oxygen environment.

Keywords: glass sponge, Porifera, Hexactinellida, symbiosis, microbiome, microbial metabolism, metagenomic binning, SAR324, *Crenarchaeota*, *Patescibacteria*, *Nanoarchaeota*, glass sponge

INTRODUCTION

The fossil record shows that sponges (Porifera) have been essential members of reef communities in various phases of Earth's history and even built biohermal reefs in the mid-Jurassic to early-Cretaceous (1). Today, extensive sponge aggregations, also known as sponge grounds, are found throughout the World's oceans, from temperate to arctic regions along shelves, on ridges, and on seamounts (2). They can be mono- to multispecific with a single or various sponge species dominating the benthic community, respectively. In sponge ground ecosystems, these basal animals play a crucial role in the provision of habitat, adding structural complexity to the environment and thereby attracting other organisms, ultimately causing an enhancement of local biodiversity (3–5).

Studies on demosponges have shown that they harbor distinct and diverse microbial communities that interact with each other, their host, and the environment in various ways (6, 7). The microbial consortia of sponges are represented by diverse bacterial and archaeal communities, with ≥ 63 prokaryotic phyla having been found in sponges so far (6, 8). These sponge microbiomes display host species-specific patterns that are distinctly different from those of seawater in terms of richness, diversity, and community composition. Microbial symbionts contribute to holobiont metabolism (e.g., via nitrogen cycling and vitamin production) and defense (e.g., via secondary metabolite production) (reviewed in reference 7). Sponges and their associated microbial communities (here termed holobionts) further contribute to fundamental biogeochemical cycles like nitrogen, phosphorous, and dissolved organic matter in the ecosystem, but the relative contribution of microbial symbionts remains mostly unresolved (1, 7, 9, 10).

Sponges of the class Hexactinellida (glass sponges) are largely present and abundant in the mesopelagic realm below 400

feet. They can form extensive reefs of biohermal character and can dominate the sponge ground ecosystems (1, 11). Glass sponges are characterized by a skeleton of siliceous spicules that is six-rayed symmetrical with square axial proteinaceous filament (12). Much of the body is composed of syncytial tissue, which represents extensive and continuous regions of multinucleated cytoplasm (12, 13). Nutrients also are transported via the cytoplasmic streams of these trabecular syncytia (12). Some discrete cell types exist, including choanocytes and the pluripotent archaeocytes that are likely nonmotile and, thus, not involved in nutrient transport in Hexactinellida (12). While the microbial symbiont diversity and functions are well studied in Demospongiae, much less is known about the presence and function of microbes in glass sponges. In fact, microorganisms have rarely been seen in glass sponges (12). However, a recent study of South Pacific sponge microbial communities has shown that general patterns seen previously in shallow-water sponge microbiomes, such as host specificity and low microbial abundance-high microbial abundance dichotomy, are generally applicable for these deep-sea sponge microbiomes as well, including those of glass sponges (14). Another study underlined the importance of ammonia-oxidizing archaea (family *Nitrosopumilaceae*, phylum *Thaumarchaeota*) in the deep-sea hexactinellid *Lophophysema eversa* using metagenomic data (15).

Here, we investigate the microbial community of *Vazella pourtalesii* (16), a glass sponge (class Hexactinellida) that forms globally unique, monospecific sponge grounds on the Scotian Shelf off Nova Scotia, Canada. This ecosystem is characterized by relatively warm and nutrient-rich water with low oxygen concentrations (1, 17, 18). While there have been a number of studies on the distribution, biomarkers, and possible functional roles of *V. pourtalesii* in the ecosystem (1, 17, 19), little has been published to date on its associated microbiota (20). According to phylogenetic and

fossil studies, sponges (including glass sponges) originate from Neoproterozoic times when oxygen was limited (21). Moreover, laboratory experiments have shown that sponges can cope with low oxygen levels for extended periods of time (22–24). Due to the observed low-oxygen conditions at the sampling location, we

explored whether the *V. pourtalesii* microbiome contains compositional as well as functional adaptations to the low-oxygen environment. Microscopy, metagenome-guided microbial genome reconstruction, and amplicon community profiling were employed toward this goal.

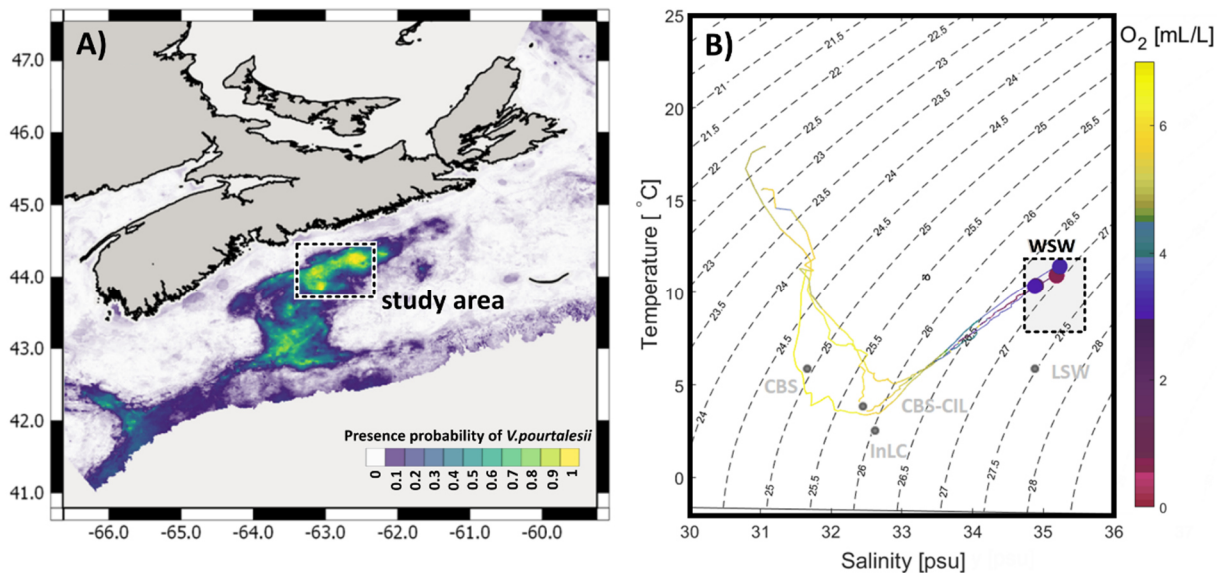


FIGURE 1 | Map of sampling region on the Canadian shelf (A) and TS diagram (B). (A) Colors depict presence probability of *Vazella pourtalesii* based on data presented in Beazley et al. (17), with yellow indicating areas of highest occurrence probability. (B) Coloring corresponds to oxygen concentrations measured during representative CTD casts at the study area. Water masses (light gray dots, labels, and square) were added according to Dever et al. (101) and Fratantoni et al. (102): CBS, Cabot Strait Subsurface Water; InLC, Inshore Labrador Current; CBS-CIL, Cold Intermediate Layer of Cabot Strait Subsurface Water; LSW, Labrador Slope Water; WSW, Warm Slope Water.

RESULTS

Site description.

On the Scotian Shelf off Nova Scotia, eastern Canada, the highest densities of *V. pourtalesii* were observed and/or predicted in the Emerald Basin and the Sambro Bank areas (Fig. 1A). The water column of this region has a characteristic vertical structure, with water masses of different temperatures and salinities gradually mixing and creating a distinct temperature-salinity profile (Fig. 1B). Main water masses influencing the sampling sites are, from surface to deep sea, the following: Cabot Strait Subsurface Water (CBS), Inshore Labrador Current (InLC), Cold Intermediate Layer of

Cabot Strait Subsurface Water (CBS-CIL), Labrador Slope Water (LSW), and Warm Slope Water (WSW). All *V. pourtalesii* samples of this study originate from a relatively warm (>10°C) and nutrient-rich water mass, called Warm Slope Water, which originates from the Gulf Stream (18). Relatively low oxygen concentrations (<4 mL/L) were measured at the sampling locations and depths that lay in the range of a mild hypoxia (25).

Microscopic analyses of *V. pourtalesii*.

In contrast to other sponges, Hexactinellidae mainly consist of a single syncytium, a fusion of eukaryotic cells forming multinucleate tissue that permeates the whole sponge (12). By

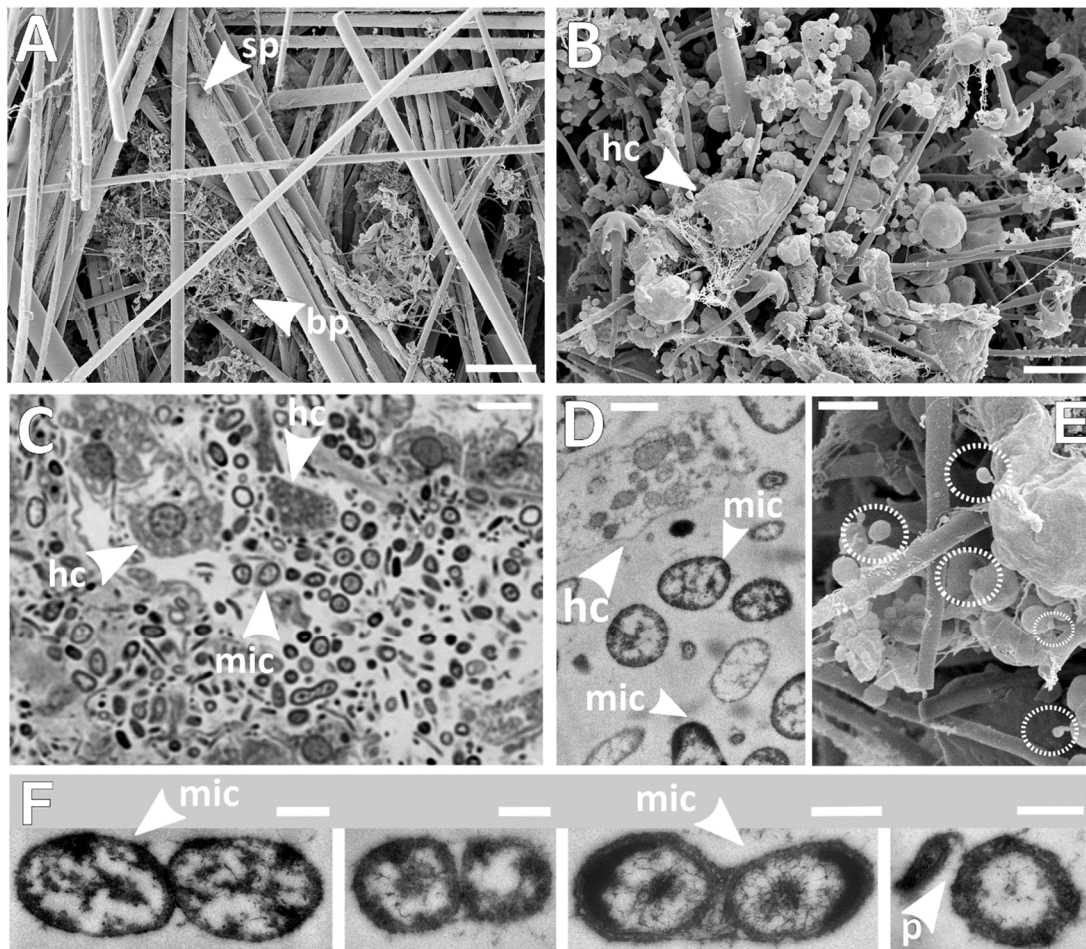


FIGURE 2 | Microscopy of *Vazella pourtalesii* tissue. **(A)** Scanning electron microscopy overview of spicule scaffolds (scale bar, 75 μm). **(B)** SEM closeup image of a biomass patch (scale bar, 3 μm). **(C and D)** Light microscopy image (scale bar, 5 μm) **(C)**, and TEM image of the same biomass patch (scale bar, 1 μm) **(D)**. **(E)** SEM closeup presumably showing smaller microbes attached to larger ones by stalk- or filament-like structures (scale bar, 1 μm). **(F)** TEM images of adjacent microbial cells (scale bars, 500 nm). Acronyms: sp, spicule; bp, biomass patch; hc, host cell; mic, microbes; p, potential pilus.

scanning electron microscopy (SEM), we observed that the overall amount of sponge biomass in *V. pourtalesii* was low, and its distribution within the spicule scaffolds was patchy (**Fig. 2A and B**). Closer inspection of such biomass patches by light microscopy and by transmission electron microscopy (TEM) (**Fig. 2C and D**) revealed numerous host cells with their characteristic nuclei, as well as high densities of microbial cells of various morphologies and with a dominance of comparatively small cell sizes (<1 μm). In addition, we frequently observed microbial cells that were attached to each other with pilus-like structures (**Fig. 2E and F**).

MAG selection.

In total, 137 metagenome-assembled genomes (MAGs) of >50% estimated completeness and <10% redundancy were retrieved (see Table S1 in the supplemental material) from seven sponges and five seawater controls that were sampled from natural *Vazella* grounds and a mooring float (Table S2). Proteobacteria followed by Patescibacteria were the dominant bacterial phyla according to amplicon analyses. In addition, linear discriminant analysis (LDA) scores were obtained for MAGs based on their read abundance in the different metagenomic sample types (**Fig. S1**): (i) *V. pourtalesii* metagenomes versus seawater metagenomes and (ii) pristine *V. pourtalesii* versus mooring *V. pourtalesii* versus seawater metagenomes.

TABLE 1 | MAGs of bacterial candidate phyla Patescibacteria and SAR324 and archaeal phyla *Crenarchaeota* and *Nanoarchaeota*^a

Phylum-level affiliation and MAG	No. of contigs	Estimated genome size (Mb)	GC (%)	<i>N</i> ₅₀	Cov (%)	Red (%)
<i>Patescibacteria</i>						
Patesci_129	109	0.71	31.8	5,408	67.37	0
Patesci_30	95	0.46	31.7	5,089	66.19	1.18
Patesci_136	206	0.95	31.6	4,840	61.21	0.47
Patesci_48	163	0.86	32.6	8,405	71.52	0
Patesci_98	15	0.71	28.4	66,386	79.44	0
<i>SAR324 (Deltaproteobacteria)</i>						
SAR324_126	562	2.16	33.0	8,540	75.63	5.39
SAR324_140	257	1.49	32.2	4,131	54.53	0.75
SAR324_8	270	1.47	32.4	3,229	54.81	2.45
<i>Crenarchaeota</i>						
CrenArch_143	507	2.04	38.2	3,080	63.44	1.46
CrenArch_74	366	1.41	40.1	3,432	53.2	8.16
CrenArch_90	115	1.11	40.4	8,261	59.77	2.59
CrenArch_101	336	1.74	31.8	3,376	60.75	3.40
<i>Nanoarchaeota</i>						
NanoArch_78	39	0.67	24.3	39,559	76.63	0

^aGenome properties were determined by QUASt, and completeness and contamination estimations were performed by CheckM implemented in the metaWRAP pipeline. Acronyms: Cov, genome coverage; Red, redundancy.

Based on these assessments, we selected 13 representative *V. pourtalesii*-enriched MAGs of the four dominant phyla for detailed analyses (Table 1). Five MAGs belonged to the candidate phylum Patescibacteria, three to the candidate phylum SAR324, four to the phylum *Crenarchaeota*, and one to the phylum *Nanoarchaeota*. The selected MAGs are evidently not redundant representations of the same microbial genomes, which is visualized by the comparatively long branches in the phylogenomic tree (Fig. 3). Additionally, the MAGs showed a maximum of 86.8%, 93.2%, and 88.3% similarity to each other in the average nucleotide identity (ANI) analysis for SAR324, *Crenarchaeota*, and Patescibacteria, respectively (Table S4).

Phylogenetic placement of the major players.

Three of the selected MAGs belong to the candidate phylum SAR324. The SAR324 clade was recently moved to the level of candidate phylum along with the publication of the whole-genome-based classification of microbial genomes (26). The three SAR324 MAGs clustered outside known orders with relatively long branches, showing that they are genomically distinct from published genomes of their closest relatives (Fig. 3). In the amplicon analysis,

this taxon was placed within the class *Deltaproteobacteria*, in which it was the most abundant order (Fig. S2).

Three crenarchaeal MAGs clustered with *Cenarchaeum symbiosum* A (GCA000200715.1), associated with the sponge *Axinella mexicana* (27). One MAG (CrenArch_101) was closely related to *Nitrosopumilus* species isolated from Arctic marine sediment (28) and the deep-sea sponge *Neamphius huxleyi* (29). *Crenarchaeota* are not represented adequately in our amplicon data due to the sequencing primer bias toward bacteria.

Three of the five *V. pourtalesii*-enriched Patescibacteria MAGs clustered with genome GCA002747955.1 from the oral metagenome of a dolphin (30). This clade is a sister group to other families of the order UBA9983 in the class *Paceibacteria*. The other two patescibacterial MAGs belonged to the family *Kaiserbacteriaceae*. They were placed separate from each other and outside known genera, where they cluster with two different groundwater bacteria (GCA_000998045.1 and GCA_002773335.1). Patescibacteria were the second most abundant phylum in the amplicon data, with a dominance of the class *Parcubacteria* that showed high abundances in *V. pourtalesii* compared to seawater and sediment

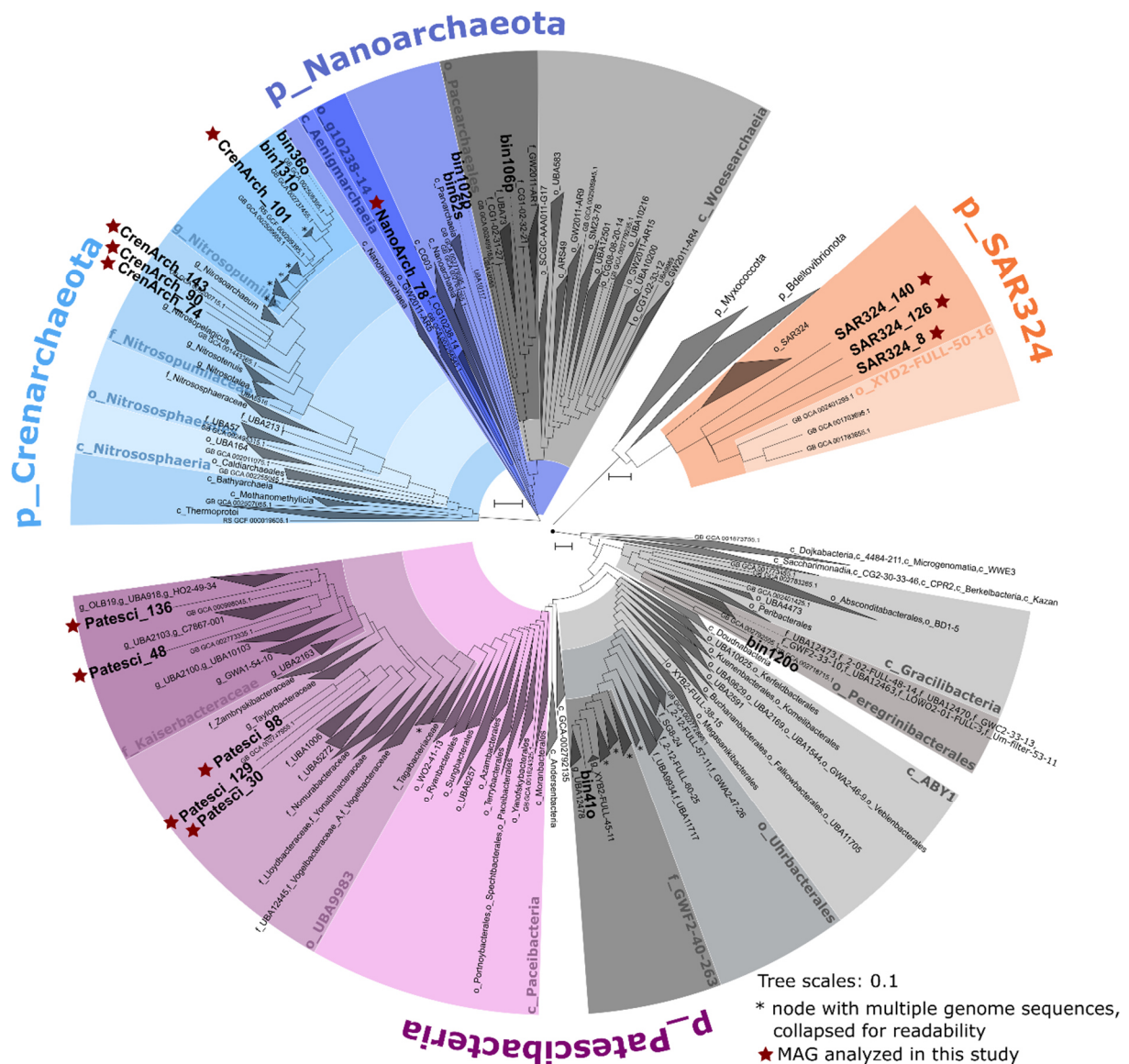


FIGURE 3 | Subtrees of the GTDB-Tk phylogenetic tree showing the setting of the MAGs of this study (shown in boldface) within the microbial phyla selected for detailed analysis. Class names are indicated by a leading “c_,” order names by “o_,” family names by “f_,” and genus names by “g_.” *Vazella*-enriched MAGs are marked with a red star.

controls (Fig. S3). The majority of the Parcubacteria (73%) remain unclassified, while 27% are classified as order *Kaiserbacteria*. As this phylum has not been noticed as particularly abundant in sponge microbiomes before, we tested whether they are sponge specific by comparison with the reference database of the Sponge Microbiome Project (SMP) (8 and data not shown). The 899 patescibacterial *V. pourtalesii* ASVs matched to 42 SMP operational taxonomic units (OTUs) (mostly listed as “unclassified Bacteria” due to an older SILVA version). We identified three SMP OTUs matching *V. pourtalesii* *Kaiserbacteria* ASVs, namely,

OTU0005080, OTU0007201, and OTU0159142. These OTUs occurred in 45, 26, and 7 sponge species, respectively, in the SMP reference database showing a global distribution.

The one *Nanoarchaeota* MAG, NanoArch_78, belongs phylogenomically to the class Aenigmarchaeia, where it is placed together with MAG GCA_002254545.1 from a deep-sea hydrothermal vent sediment metagenome and outside known families. Despite the primer bias toward bacteria in the amplicon analysis, the phylum *Nanoarchaeota* was among the

most abundant microbial phyla in this analysis (Fig. S2).

Genome sizes and GC contents.

With respect to genome size, the MAGs range from 0.46 Mb (Patesci_30) to 2.16 Mb (SAR324_126), including completeness values into the genome size estimations, with N50 values between 3,080 (CrenArch_143) and 66,386 (Patesci_98) (Table 1). According to CheckM, between 53.2% and 79.4% of the genomes are covered and redundancies range from 0% to 8.16%. GC contents range from 24.3% (NanoArch_78) to 40.38% (CrenArch_90). We compared the MAGs to genomes of symbionts from other sponge species and of seawater-derived microbes of each respective phylum regarding their estimated genome sizes and GC contents (Fig. 4 and Table S3). Due to the lack of published genomes of

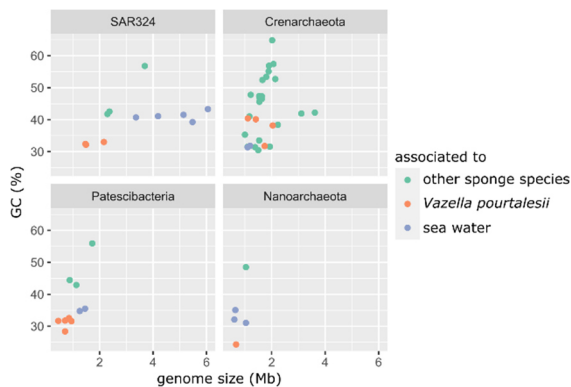


FIGURE 4 | Comparison of the MAGs retrieved in this study to published MAGs from sponge and seawater metagenomes. From this study, only the MAGs enriched in either *Vazella pourtalesii* or in water were considered; neutral ones were excluded from this analysis.

SAR324 and *Nanoarchaeota* sponge symbionts, we included genome size and GC content data of unpublished symbionts of *Phakellia* species and *Stryphnus fortis* in this analysis. This comparison revealed that the genomes of *V. pourtalesii*-enriched microbes are exceptionally small, with very low GC percentages. This trend is especially striking for Patescibacteria, SAR324, and *Nanoarchaeota*, as their values are not only low for sponge symbionts but even

below the levels of the respective related seawater microbes. While this is not the case for *Crenarchaeota*, the *V. pourtalesii* MAGs are, nevertheless, in the lower ranges regarding size and GC content compared to other sponge symbionts.

Predicted lifestyle of the major players.

(i) SAR324.

Metagenomic analysis of the three SAR324 MAGs from *V. pourtalesii* (Fig. 5A) revealed the presence of a nearly complete glycolysis pathway up to pyruvate (Pyr) along with the genes for the tricarboxylic acid (TCA) cycle and for conversions of the pentose phosphate pathway (PPP) (see Text S1 for in-depth analysis of more complex SAR324 and crenarchaeal MAGs). Pyruvate is converted aerobically by the pyruvate dehydrogenase enzyme complex into acetyl-coenzyme A (CoA), which fuels the completely annotated (thiamine-dependent) TCA cycle. While SAR324 have the genes for a nearly complete respiratory chain, their lifestyle appears to be facultatively anaerobic. We detected enzymes of the glyoxylate bypass (orange arrows within the TCA cycle in Fig. 5A), which is required by bacteria to grow anaerobically on fatty acids and acetate (31). This is supported by the presence of a potential AMP-dependent acetyl-CoA synthetase to utilize acetate, whereas enzymes for fatty acid degradation were not found. SAR324 might gain additional energy by a cation-driven p-type ATPase and possibly also anaerobic respiration (fumarate and nitrite/sulfide respiration) (also see Text S1). There is evidence for assimilatory sulfate reduction, but the pathway was not fully resolved.

V. pourtalesii-associated SAR324 organisms are able to take up di- and tricarboxylates using TRAP and TTT transporters, respectively. The imported substances can feed the TCA cycle under aerobic conditions or serve as an energy source through fumarate respiration under anaerobic conditions (32, 33). The presence of lactate dehydrogenase (LDH),

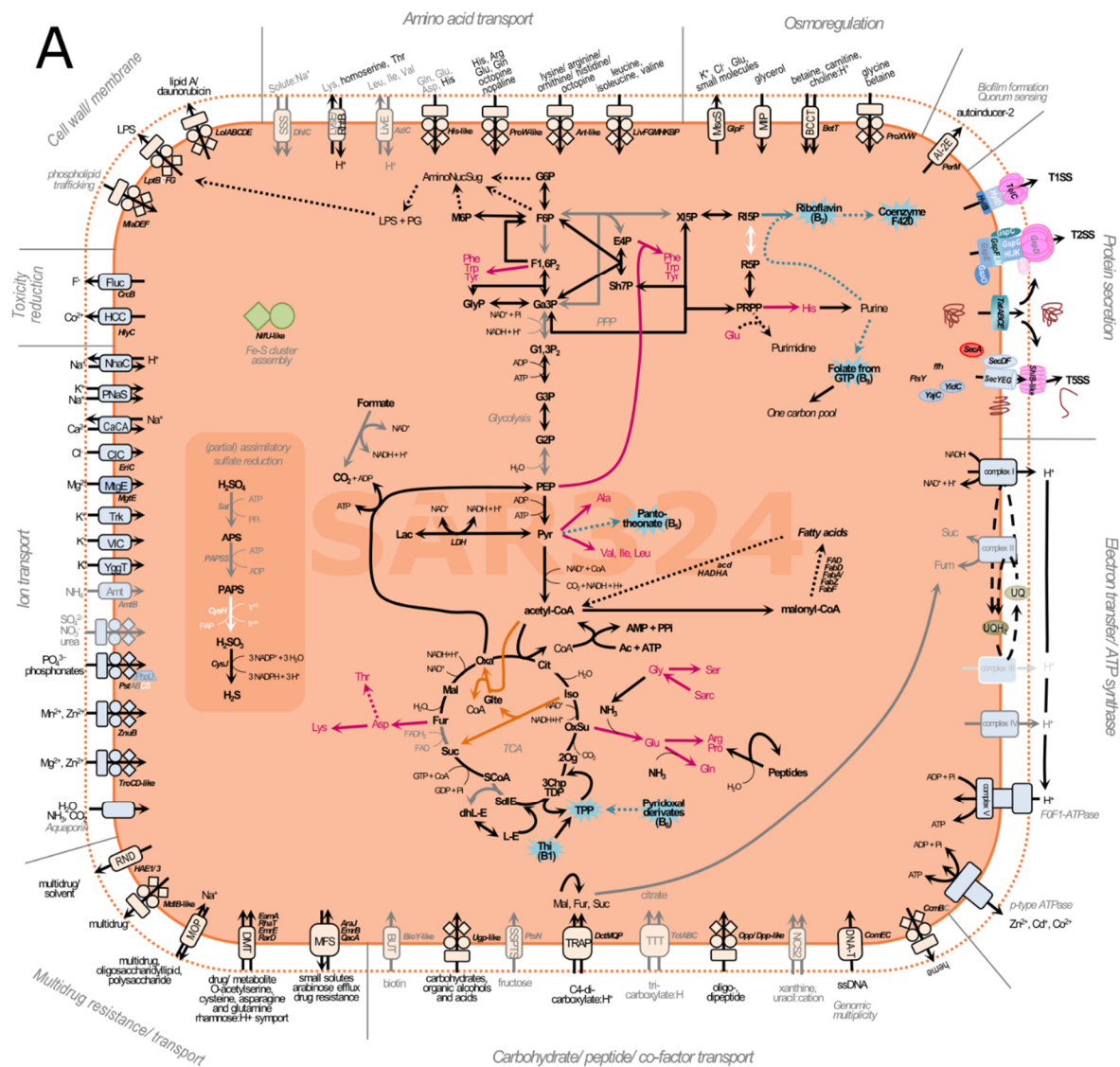


FIGURE 5 | Reconstruction of metabolic features found in the genomes of SAR324 (A), *Crenarchaeota* (B), *Patescibacteria* (C), and *Nanoarchaeota* (D). Solid lines indicate that genes/enzymes, or >50% of a given pathway, were found, and dashed lines indicate that <50% of a pathway was found. Gray arrows, writing, and lining indicate that the genes/enzymes were found in <50% of genomes of the respective phylum. White arrows and writing indicate missing genes/enzymes. Cofactor synthesis is indicated by turquoise color and amino acid production by magenta color. Symport, antiport, uniport, and direction are indicated by number and direction of arrows.

involved in fermentation, further supports a facultatively anaerobic lifestyle (31).

SAR324 symbionts are capable of synthesizing diverse amino acids and B vitamins (riboflavin, coenzyme F420, folate, pantothenate and thiamine from pyridoxal) using precursors from glycolysis, PPP, and the TCA cycle (summarized in Fig. 5A; see Text S1 for more details). Additionally, the genomes are well equipped with several transporters enabling the import and export of diverse substances (e.g., sugars, amino acids, peptides, and ions)

(Text S1). These transporters are involved in, e.g., osmoregulation and/or toxic ion reduction (cobalt and fluoride) and multidrug resistance/import. A p-type ATPase was annotated that may aid in the export of cations or may use an electrochemical gradient for ATP synthesis. Proteins can be secreted by *tat* and *sec* transport as well as type 1 (T1SS) and type 2 (T2SS) secretion systems, probably involved in excretion of symbiosis-relevant molecules. We further identified autoinducer-2 (AI-2) and DNA-T family transporter (Text S1).

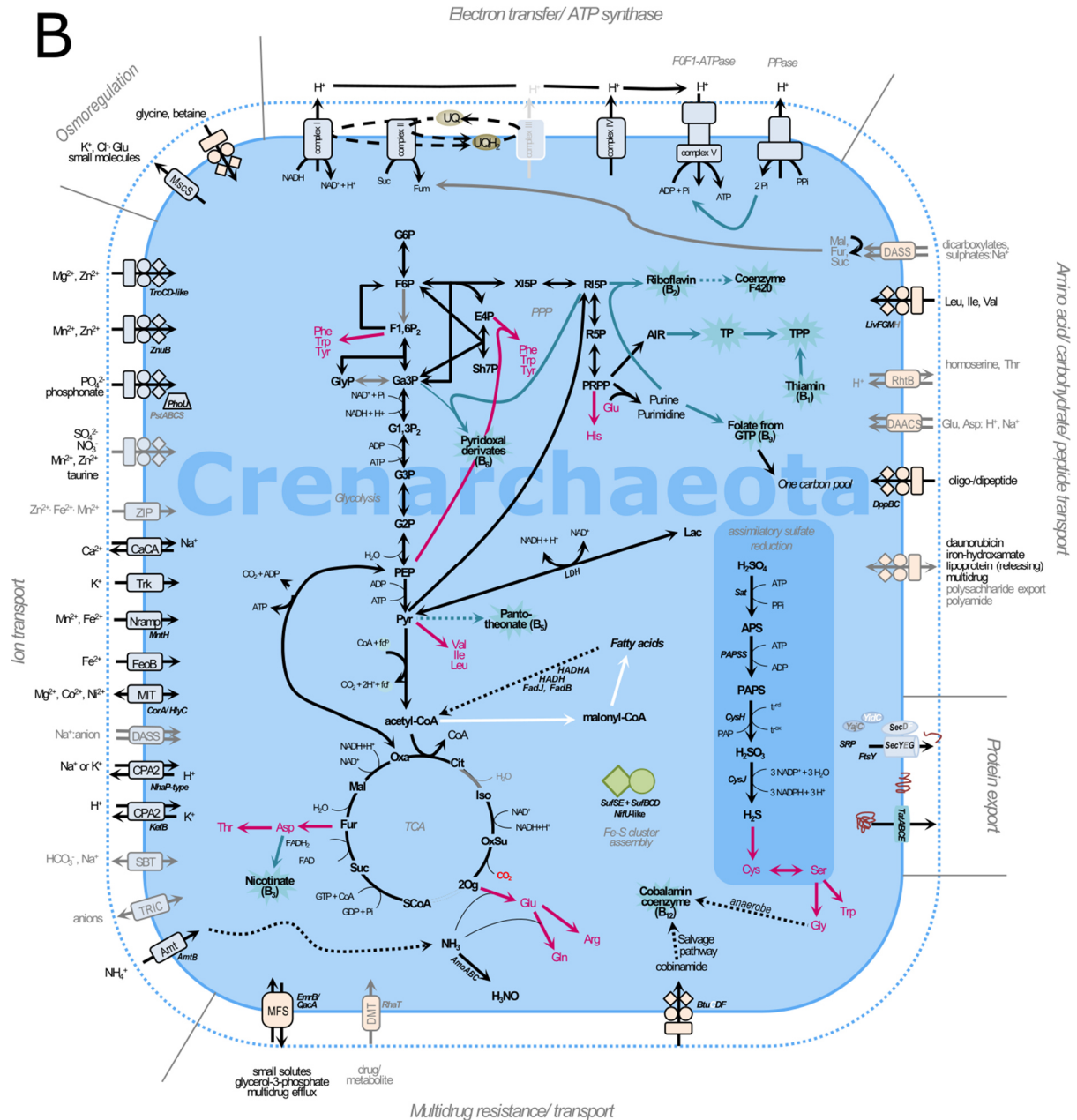


FIGURE 5 | (Continued)

(ii) Crenarchaeota.

The *Crenarchaeota* of *V. pourtalesii* (Fig. 5B) are capable of glycolysis from glucose-6-phosphate (G6P) to Pyr, which is likely anaerobically converted into acetyl-CoA using the enzyme pyruvate-ferredoxin oxidoreductase, which suggests facultatively anaerobic metabolism. The TCA cycle was almost completely annotated. Genomic evidence for aerobic and anaerobic respiration (fumarate and nitrite/sulfide respiration) was detected (Text S1). Genes for autotrophic CO₂ fixation in *V.*

pourtalesii-associated microbes were lacking, but assimilatory sulfate reduction was annotated completely. A PPase was annotated that might deliver phosphates to feed the ATP synthase, and in one MAG, a transporter to import dicarboxylates (fumarate, malate, and succinate) was annotated, a feature we found in SAR324 as well. These substances could feed the TCA cycle under aerobic conditions. Additionally supporting the hypothesis of a facultatively anaerobic lifestyle is the presence of the enzyme LDH, which

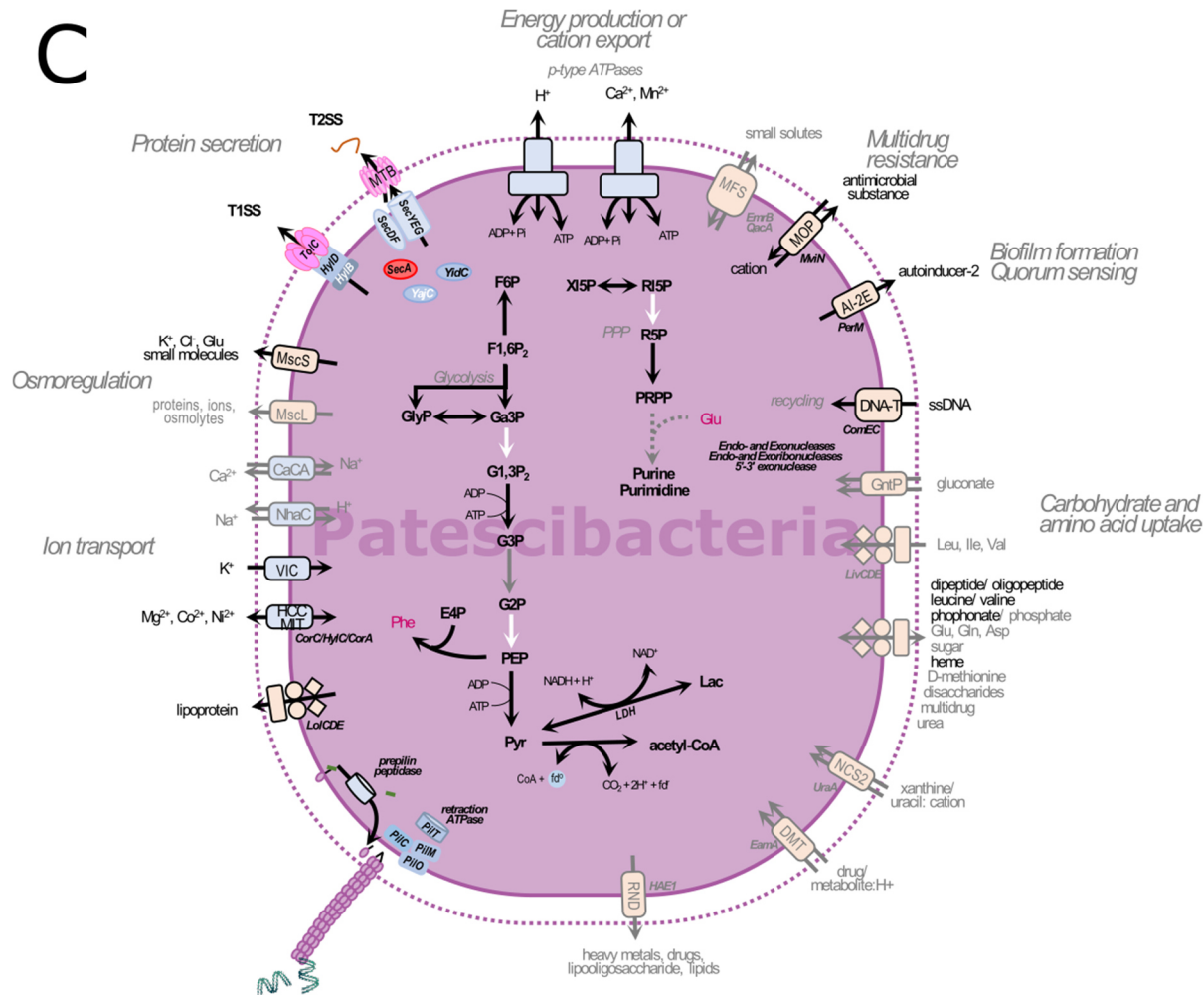


FIGURE 5 | (Continued)

is involved in fermentation (31). The crenarchaeal genomes encode the synthesis of an even greater number of B vitamins than the SAR324 genomes, including riboflavin, coenzyme F420, folate, pantothenate, pyridoxal, nicotinate, and cobalamin (anaerobically), using precursors of central metabolism. Interestingly, they can synthesize thiamine from 5-phosphoribosyl diphosphate (PRPP), while SAR324 partially encode thiamine synthesis from pyridoxal, which would need to be imported from an external source (e.g., from other community members).

The crenarchaeal genomes are further well equipped with transporters that facilitate import and export of diverse substances, such as sugars, amino acids, peptides, and ions, among others (Text S1). These transporters are

involved in, e.g., osmoregulation, reflecting the adaptation to a saline environment, and in multidrug resistance/import. A p-type ATPase was annotated in the genomes that may be involved in the export of cations (forced by ATP utilization). Protein secretion can be realized by *tat* and *sec* transport, which might be involved in the transport of proteins, such as those necessary for membrane formation and maintenance (Text S1).

(iii) Patescibacteria.

The *V. pourtalesii*-associated Patescibacteria (Fig. 5C) showed metabolic capacity similar to that of published Patescibacteria from other environments. While we found several enzymes involved in glycolysis, we could not resolve the pathway completely. The genomes

lack enzymes involved in oxidative phosphorylation (respiration) and the TCA cycle. The synthesis pathways of the important precursor PRPP and subsequent synthesis of purines and pyrimidines were only partially encoded. The biosynthesis of phenylalanine (Phe) from phosphoenolpyruvate (PEP) and erythrose-4-phosphate (E4P) is encoded, but biosynthesis pathways for other amino acids, cofactors, or vitamins are missing. We found two p-type ATPases that might export ions or provide energy (ATP) using cation and/or proton gradients present in the environment (holobiont). An anaerobic lifestyle is likely due to the presence of an LDH, an enzyme involved in lactate fermentation and anaerobic acetyl-CoA synthesis using the enzyme pyruvate-ferredoxin oxidoreductase. Patescibacterial MAGs also encode some transporters, albeit in lower

numbers than the above-described SAR324 and crenarchaeal genomes. These transporters may be involved in osmoregulation, multi-drug influx and efflux, and sugar, amino acid, and ion uptake. Regarding further symbiosis-relevant features, we detected the autoinducer transporter AI-2E. Additionally, we detected *ComEC/Rec2* and related proteins that are involved in the uptake of single-stranded DNA (34). This is supported by the presence of *PilT*, the motor protein that is thought to drive pilin retraction prior to DNA uptake, and the pilus assembly proteins *PilM*, *PilO*, and *PilC*. Even though the machinery was not annotated completely, Patescibacteria may be able to take up foreign DNA via retraction of type IV-pilin-like structures into the periplasm and via *ComEC* through the inner membrane.

D

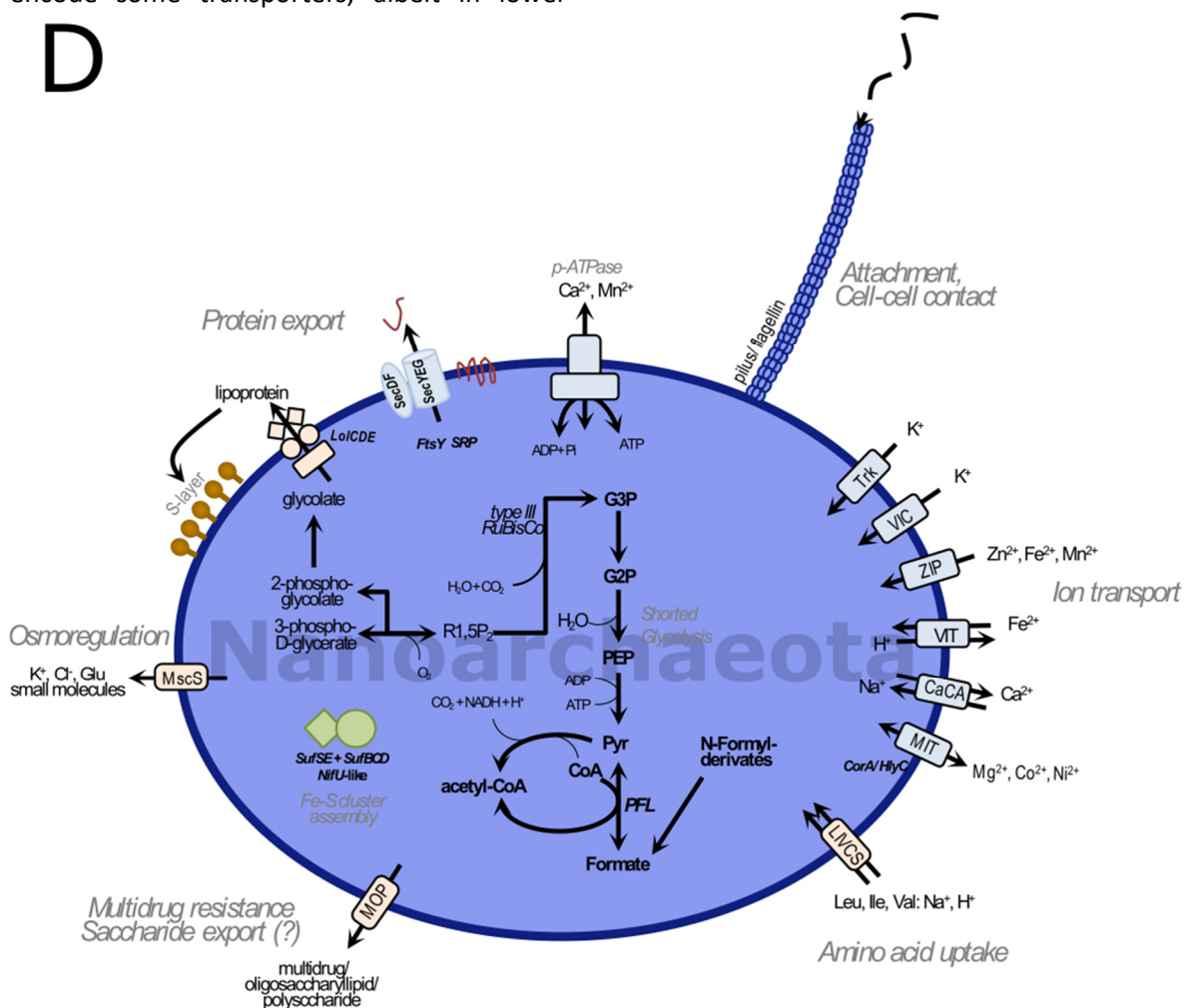


FIGURE 5 | (Continued)

(iv) Nanoarchaeota.

The nanoarchaeal MAG (Fig. 5D) shows the genomic potential to convert glycerate-3-phosphate (G3P) to Pyr, which represents a shortened glycolysis pathway and results in reduced potential for energy production (ATP synthesis). It could use an anaerobic pyruvate-formate lyase (PFL) for acetyl-CoA production. Interestingly, an archaeal type III RuBisCo is encoded that catalyzes light-independent CO₂ fixation using ribulose-1,5-bisphosphate and CO₂ as substrates to synthesize G3P (35) to fill the only partially encoded glycolysis. All enzymes needed for respiration were absent, supporting an anaerobic lifestyle. Energy might be gained using a cation-driven p-type ATPase. Like its published relatives (36, 37), the *V. pourtalesii*-associated nanoarchaeon lacks almost all known genes required for the de novo biosynthesis of amino acids, vitamins, nucleotides, cofactors, and lipids. The uptake of some amino acids and ions may be possible, as few transporters were detected. Typical archaeal S-layer membrane proteins are encoded that may be exported by an ABC-transporter (*LoiCDE*) or by sec transport. Other transporters are involved in osmoregulation and in multidrug resistance and/or transport. We found a prepilin type IV leader peptidase encoded in the genome that is synthesized as a precursor before flagellin/pilin is incorporated into a filament (38).

DISCUSSION

Microbial associations in *V. pourtalesii* syncytia.

The glass sponge *V. pourtalesii* consists of a scaffold of spicules with cellular biomass concentrated in biomass patches that contain sponge as well as microbial cells (Fig. 2). We assume that the biomass patches were probably formed by dehydration of syncytial tissues during fixation, resulting in higher biomass densities than in the *in vivo* situation (12). Surprisingly high numbers of microbial cells were found within the observed biomass patches,

considering that microbes have rarely been noticed in glass sponges previously (12). In *V. pourtalesii*, the microbes appeared in various morphotypes, indicating a taxonomically diverse microbial community. Microbial cells of strikingly small sizes (<1 μm) compared to those of shallow water demosponges (39) constituted a large fraction of the microbial community. Microbial cells were frequently seen in close association and even physically attached to each other (Fig. 2E and F). These associations were observed between equally sized cells but also between cells of distinctly different sizes, where the smaller microbes were attached to larger ones by stalk- or filament-like structures (Fig. 2E).

Main players in the *V. pourtalesii* microbial community with small, low-GC genomes.

While previously published sponge metagenomes and MAGs tended toward high-GC contents (typically around 65 to 70%), the *V. pourtalesii* MAGs show lower GC levels in the range of 24 to 40% that are more similar to those of seawater metagenomes (40–42) (Fig. 4; see also Table S3 in the supplemental material). Genome sizes are also on the smaller side compared to previously published sponge MAGs (40, 42). The large genome sizes of demosponge symbionts may be attributed to the specific genomic toolbox they require to utilize the mesohyl matrix, such as carbohydrate-active enzymes (CAZy) and arylsulfatases. These genes are frequently found enriched in the sponge symbiont genomes compared to free-living relatives (42, 43). This is, however, not the case in *V. pourtalesii* MAGs. On the contrary, here we see GC contents and genome sizes similar to and even below the ones of free-living marine microbes of the same respective phyla (Fig. 4). Trophic specialization and avoidance of DNA replication cost have been proposed as hypotheses for genome reduction in free-living marine bacteria of, e.g., the genera *Idiomarina* (44) and *Pelagibacter* (45). For the *V. pourtalesii*-associated microbial community, the Black Queen hypothesis may

best explain the apparent genomic streamlining: if some members carry out tasks that are beneficial to the whole microbial community, most other members will lose the ability to carry out these (often costly) tasks (46). Thus, the small sizes and low GC contents of the *V. pourtalesii* MAGs could be a sign of adaptation to generally nutrient-limiting environmental conditions (reviewed in reference 47) and specialization on nutrient sources that are available within the sponge host environment, such as ammonia.

The “givers” and “takers” hypothesis.

Formerly placed within the *Thaumarchaeota*, the *Crenarchaeota* are well known and widespread sponge symbionts (48–51). Different genera have recently been observed in South Pacific Hexactinellida and Demospongiae (14). *Nanoarchaeota* and Patescibacteria also have recently been noticed as members of sponge microbial communities, including glass sponges, but no sponge-associated nanoarchaeal genome has been studied to date, and, likely due to their low abundance in other sponge species, patescibacterial genomes have not been studied in detail (14, 52, 53). No sponge-derived genomes are available for the phylum SAR324 so far. The genomes of *V. pourtalesii* microbes lack a number of properties that we know from typical shallow-water sponge symbionts (7), such as the potential for the production of secondary metabolites and arylsulfatases, and they are not enriched in CAZy genes. This underlines the above-stated hypothesis that these sponge symbionts do not need a diverse toolbox of genes to make use of a complex mesohyl, like symbionts of Demospongiae. On the contrary, they seem to possess streamlined genomes to save resources and likely rely on each other for essential substances, such as certain amino acids and vitamins.

Based on the functional genetic content of the four microbial phyla that we analyzed in greater detail, we propose two major

strategies: the givers, namely, SAR324 and *Crenarchaeota*, with comparatively larger, more complex genomes and likely bigger in cell size, and the takers, Patescibacteria and *Nanoarchaeota*, with reduced genomes and likely smaller cell sizes. We posit that the givers, being genomically well equipped, could be producing and partly secreting all required amino acids and vitamins, drawing energy from various aerobic as well as anaerobic processes (Fig. 6). Regarding their metabolic repertoire, the *Crenarchaeota* described here are rather similar to the SAR324 bacteria, namely, in their facultatively anaerobic lifestyle, the reactions of the central metabolism, and their ability for amino acid and vitamin biosynthesis. At the same time, while published sponge-associated or free-living *Crenarchaeota* have the genomic repertoire to fix carbon (37, 50, 54–56), such pathways appear to be absent from *V. pourtalesii*-associated *Crenarchaeota*. These findings indicate that the microbes are specifically adapted to the conditions within their respective host sponge and the surrounding environment, e.g., low-oxygen conditions in this case. Supporting our hypothesis of genome streamlining in the sense of the Black Queen hypothesis, the two givers also seem to depend on each other metabolically: *Crenarchaeota* can produce several B vitamins, which might be used by members of SAR324. One example is pyridoxal, provided by *Crenarchaeota* to SAR324, which would be able to produce thiamine. Their genomic similarity, their difference from close relatives from other environments, and their metabolic interdependence reinforces our hypothesis that they are, in fact, symbionts specifically adapted to life within their *V. pourtalesii* host. Beyond the scope of the microbial community, microbial vitamin production may also have an important role in animal host metabolism, such as the respiratory chain, the synthesis of coenzyme A, proteins, fatty acids, nucleic acids, cofactor synthesis, and carbohydrate metabolism (Fig. 6). As previously hypothesized for Demospongiae symbionts (7), the capacity for vitamin

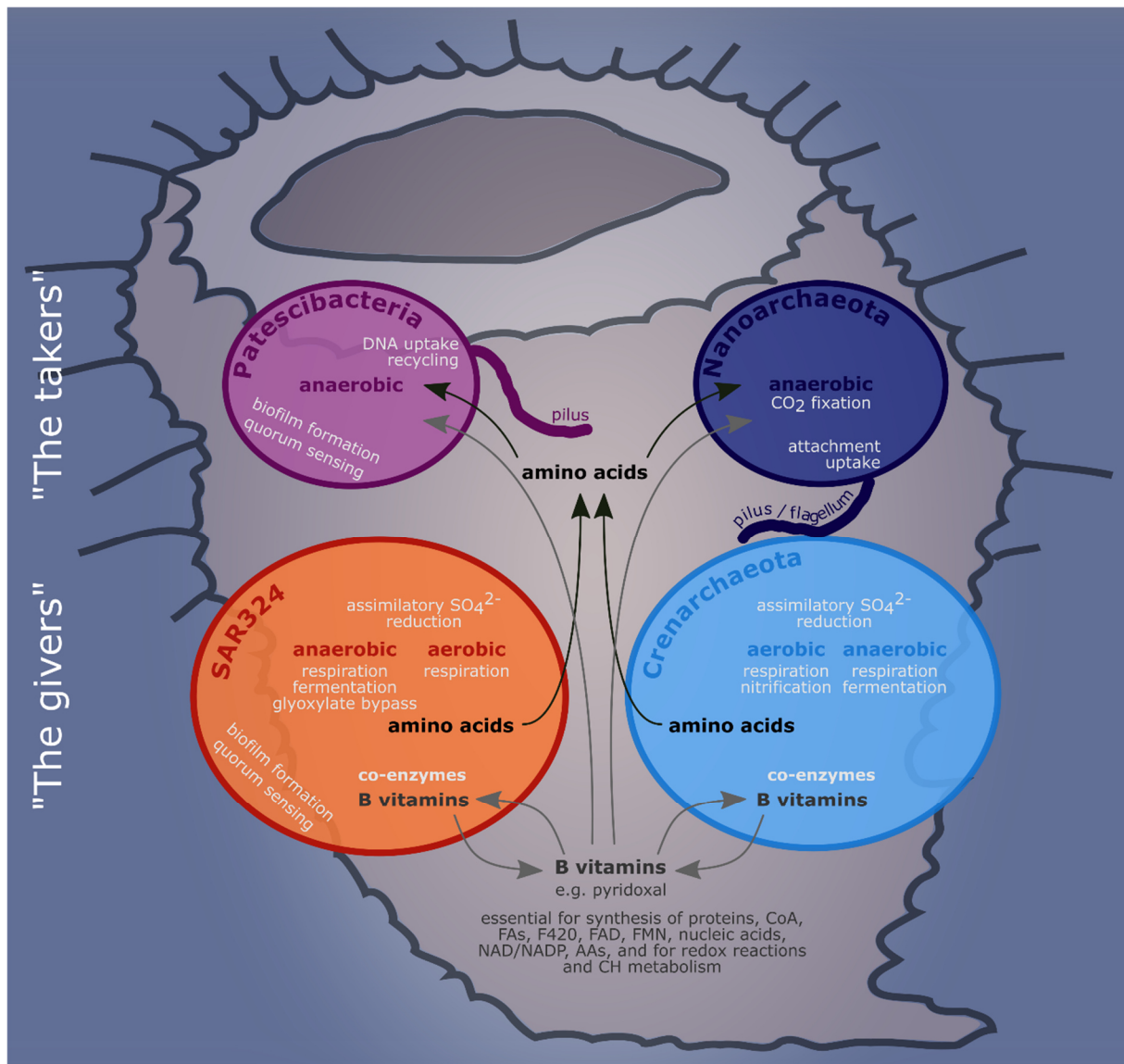


FIGURE 6 | Summary model of the main metabolic interactions between the four microbial taxa studied here. Black arrows indicate putative amino acid transport, gray arrows indicate B vitamin transport. CoA, acetyl-CoA; FAs, fatty acids; F420, coenzyme F420; FAD, flavin adenine dinucleotide; FMN, flavin mononucleotide; AAs, amino acids; CH, carbohydrate. The shape in the background depicts the host sponge *V. pourtalesii*.

synthesis by microbes associated with *V. pourtalesii* might be an important factor in maintaining the symbiosis with the animal host.

It is tempting to speculate that the takers, reduced in size and functional potential, would scavenge from their neighbors. Marine Patescibacteria are known (and named) for their reduced genomes and metabolic capacities (57–59), and *Nanoarchaeota* are known for their dependence on a crenarchaeal host, although in very different marine environments, such as hydrothermal vents (60, 61). Regarding the exchange of substances between microbes,

we propose that the nanoarchaeal symbionts “ride” on the crenarchaeal symbionts, analogous to what is described for *Nanoarchaeum equitans* and *Ignicoccus hospitalis* (37, 62), which use pili for attachment and possibly also for metabolite uptake. Jahn et al. (62) showed that *Nanoarchaeum equitans* may get large amounts of lipids (and possibly other substances) from its associated *Ignicoccus hospitalis*. We hypothesize that the nanoarchaeum in *V. pourtalesii* is likewise directly associated with the abundant *Crenarchaeota* and receives, e.g., lipids and DNA via cell attachment and using pilus-like structures to maintain cell-

cell contact (see, e.g., small cells attached to larger cells in **Fig. 2E**) (37, 60, 63). In turn, they might provide carbon to the microbial consortium by anaerobic CO₂ fixation. Patescibacteria could take up required nutrients from their microenvironment, also utilizing pili equipped with a *pilT* motor protein and *comEC* enabling the uptake of DNA for the recycling of nucleotides (64) that they cannot build themselves.

Interestingly, we detected copies of the *luxS* gene, the proposed AI-2 synthetase, in patescibacterial and SAR324 genomes. There is strong evidence that AI-2E family homologues function as an AI-2 exporter in *E. coli* cells to control biofilm formation. AI-2 is a proposed signaling molecule for interspecies communication in bacteria (reviewed in reference 65). Autoinducer production plays a crucial role in *Vibrio fischeri* colonization of (and maintenance in) the light organ of the host squid *Euprymna scolopes* (66) and was recently detected in sponge-associated *Vibrio* species (67). Microbial quorum-sensing processes (such as biofilm formation, bioluminescence, motility, virulence factor secretion, antibiotic production, sporulation, and competence for DNA uptake) (68) may display symbiosis-relevant features additional or alternative to the ones described before for sponge-associated microbes (e.g., arylsulfatases, TPRs, and CRISPR-Cas).

Conclusions.

The present study aimed to characterize the diversity and function of microbes residing in the glass sponge *Vazella pourtalesii*. A general pattern emerged in that the *V. pourtalesii* symbionts displayed smaller genome sizes and lower GC contents than bacterial relatives from seawater or from demosponge symbionts. Genomic analysis revealed two putative functional strategies: the givers (SAR324 and *Crenarchaeota*) producing and most likely providing required amino acids and vitamins to the microbial community and the takers (Patescibacteria and *Nanoarchaeota*) depending on the provision of compounds like lipids

and DNA that they likely take up via pilus-like structures. Their localization within biomass patches together with the environmental low-oxygen conditions could serve as an explanation for the unique compositional and functional properties of the microbial community of *V. pourtalesii*.

MATERIALS AND METHODS

Sampling and assessment of microbial community composition.

Sampling was performed on a cruise to the Scotian Shelf off Nova Scotia, eastern Canada, in August-September 2017 onboard CCGS *Martha L. Black* (MLB2017001). Here, we selected a subset of all samples received during this cruise (see Table S2 in the supplemental material) to study the lifestyle strategies of the dominant members of the microbial community. For details on sampling of sponges as well as seawater and sediment controls, DNA extraction, and amplicon sequencing, see reference 20, where we cover the complete data set to study the microbial diversity inside *V. pourtalesii* in response to anthropogenic activities. Briefly, sponge individuals were collected for this study from pristine areas by the remotely operated vehicle (ROV) ROPOS (Canadian Scientific Submersible Facility, Victoria, Canada), and tissue subsamples were taken, rinsed in sterile filtered seawater, and frozen at – 80°C. Samples were collected at an average sampling depth of 168 m (minimum, 161 m; maximum, 183 m), which coincides with the base of the euphotic zone. Oceanographic data, such as temperature, salinity, and oxygen, were collected using conductivity, temperature, and depth (CTD) rosette casts (sensors by Sea-Bird Electronics SBE 25). Water samples were taken during CTD casts and using Niskin bottles of the ROPOS ROV. Sediment samples were collected at the same location as sponge and seawater samples by ROV push corers (Table S2). Additionally, sponge samples were collected from an Ocean Tracking Network (OTN) acoustic mooring located approximately 10 km northwest of the Sambro Bank Sponge Conservation

Area on the Scotian Shelf (20). The mooring was anchored ~5 m above the seabed and was deployed for ~13 months (15 August 2016 to 5 September 2017) prior to its recovery.

DNA was extracted using the DNeasy power soil kit (Qiagen). After quantification and quality assessment by NanoDrop spectrophotometer and by PCR, the V3V4 variable regions of the 16S rRNA gene were amplified in a dual-barcoding approach (69) using a one-step PCR with the primers 5'-CCTACGG-GAGGCAGCAG-3' (70) and 5'-GGAC-TACHVGGGTWTCTAAT-3' (71). Samples were sequenced on a MiSeq platform (MiSeqFGx; Illumina) with v3 chemistry. The raw sequences were quality filtered using BBDUK (BBMAP version 37.75 [72]) with a Q20 and a minimum length of 250 nucleotides. Sequences were processed in QIIME2 (versions 2018.6 and 2018.8 [73]), implementing the DADA2 algorithm to determine amplicon sequence variants (ASVs). Sequences were denoised, and chimeras, chloroplasts, and mitochondrial sequences were removed. Taxonomy was assigned using a naïve Bayes classifier (75) trained on the SILVA 132 99% OTUs 16S database (76). An ASV-based phylogeny was generated using the FastTree2 plugin (77). The plots were produced with R (version 3.0.2 [78]), Inkscape (version 0.92.3 [79]), QGIS (version 2.18.4 [80]), and MATLAB (version R2016b, including Gibbs Seawater toolbox [81]).

Scanning electron microscopy.

Tissue subsamples of three sponge individuals (Table S2) were fixed for SEM onboard ship in 6.25% glutaraldehyde in phosphate-buffered saline (PBS) (Fisher Scientific) in two technical replicates each. Samples were then washed three times for 15 min each time in PBS, postfixed for 2h in 2% osmium tetroxide (Carl Roth, Germany), and washed again three times for 15 min each time in PBS. Samples were dehydrated in an ascending ethanol (EtOH) series (ROTIPURAN; Carl Roth, Germany): 1× 15 min 30% EtOH, 2× 15 min 50% EtOH, 2× 15 min 70% EtOH, 2× 15 min 80% EtOH, 2× 15 min 90%

EtOH, and 1× 15 min 100% EtOH. Subsequent dehydration was continued with carbon dioxide in a critical point dryer (BalzersCPD 030). After critical point drying, the samples were manually fractionated and sputter coated for 3 min at 25 mA with gold/palladium (Balzers SCD 004). The preparations were visualized using a Hitachi S-4800 field emission scanning electron microscope (Hitachi High-Technologies Corporation, Tokyo, Japan) with a combination of upper and lower detectors at an acceleration voltage of 3 kV and an emission current of 10 mA.

Transmission electron microscopy and light microscopy.

Tissue samples of three sponge individuals (Table S2) were fixed onboard ship in 2.5% glutaraldehyde in 0.1 M natriumcacodylate buffer (pH 7.4; Science Services GmbH) for TEM and light microscopy in two technical replicates each. Samples were then rinsed 3× with buffer at 4°C, postfixed for 2h in 2% osmium tetroxide (Carl Roth), and washed with buffer three times for 15 min each time at 4°C. Samples were partially dehydrated with an ascending ethanol (ROTIPURAN; Carl Roth) series (2× 15 min 30% EtOH, 1× 15 min 50% EtOH, up to 70% ethanol). Samples were stored at 4°C overnight before desilicification with 4% suprapure hydrofluoric acid (incubation of approximately 5 h; Merck). Afterwards, samples were washed eight times for 15 min each time in 70% EtOH (with an overnight storage at 4°C in between washings). Dehydration was continued with a graded ethanol series (1× 15 min 90% EtOH and 2× 15 min 100% EtOH) followed by gradual infiltration with LR-White resin (Agar Scientific) at room temperature (1× 1 h 2:1 ethanol:LR-White, 1× 1 h 1:1 ethanol:LR-White, 1× 1 h 1:2 ethanol:LR-White; 2× 2 h pure LR-White). Samples were incubated in pure LR-White resin at 4°C overnight before being transferred into fresh resin and polymerized in embedding capsules at 57°C for 2 days.

Semithin sections (0.5 µm) were cut (in technical replicates) using an ultramicrotome

(Ultracut E; Reichert- Jung) equipped with a diamond knife (Diatome, Switzerland) and were stained with Richardson solution (ingredients from Carl Roth; prepared as described in reference 82). Semithin sections were then mounted on SuperFrost Ultra Plus microscopy slices (using Biomount medium produced by Plano; Carl Roth) and visualized with an Axio Observer.Z1 microscope (Zeiss, Germany). Ultrathin sections (70 nm) were cut (in technical replicates) with the same ultramicrotome, mounted on pioloform coated copper grids (75 mesh; Plano), and contrasted with uranyl acetate (Science Services; 20-min incubation followed by washing steps with MilliQ water) and Reynold's lead citrate (ingredients from Carl Roth; 3-min incubation followed by washing steps with MilliQ water). The ultrathin preparations were visualized at an acceleration voltage of 80 kV on a Tecnai G2 Spirit Bio Twin transmission electron microscope (FEI Company).

Microbial functional repertoire.

For metagenomic sequencing, DNA was extracted from the seven sponge samples (four from natural *Vazella* grounds and three from the mooring to optimize for differential coverage binning) and five seawater controls (Table S2) with the Qiagen AllPrep DNA/RNA minikit. Two washing steps with buffer AW2 were employed during DNA extraction. DNase and protease-free RNase A (Thermo Scientific) were used to remove remnant RNA from the DNA extracts. For seawater controls, DNA was extracted from half of a polyvinylidene difluoride membrane filter (seawater filter; described above), while the other half of the filter was used for amplicon analyses. The DNA was concentrated by precipitation with 100% ethanol and sodium acetate buffer and reeuted in 50 µl water. For all extracts, DNA quantity and quality were assessed by NanoDrop measurements and Qubit assays, and 30 µl (diluted in water, if necessary) was sent for metagenomic Illumina Nextera sequencing (HiSeq 4000, 2× 150-bp paired ends) at the Institute of

Clinical Molecular Biology (IKMB) of Kiel University. The sequence quality of all read files was assessed with FastQC (83).

The raw reads were trimmed with Trimmomatic v0.36 (ILLUMINACLIP:NexteraPE-Pe.fa:2:30:10 LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:36) and coassembled with megahit v1.1.3 (84). The metaWRAP v1.0.2 pipeline was implemented for binning as follows (85). Initial binning was performed with metabat, metabat2, and maxbin2 (86–88) within metaWRAP. The bins were refined with the metaWRAP bin_refinement module and further improved, where possible, with the reassemble_bins module. This module uses the genome assembler SPAdes v3.12.0 (89) on two sets of reads mapped to the original bin with strict and more permissive settings and then compares the original bin with the two newly assembled genomes. Which of the three versions of the MAG was the best in each respective case and, thus, was used for further analyses is indicated by the trailing letter in the names (o, original; p, permissive; s, strict) in Table S1. The MAGs that were further analyzed in detail were renamed, indicating their phylum-level affiliation and their bin number.

MAG taxonomy was determined by GTDB-Tk based on whole-genome information and by following the recently published, revised microbial taxonomy by Parks and colleagues (26, 90, 91). The phylogenomic trees produced by GTDB-Tk (77, 92, 93) were visualized on the Interactive Tree Of Life (iTOL) platform v4.3 (94). MAG abundance in the different metagenomic data sets was quantified with the metaWRAP quant_bins module and used to determine which MAGs were enriched in which sample type by calculating LDA scores with LEfSe v1.0 (95) in two ways: (i) "*V. pourtalesii*" versus "water" and (ii) "pristine *V. pourtalesii*" versus "mooring *V. pourtalesii*" versus "water." We identified MAGs belonging to the bacterial candidate phyla Patescibacteria and SAR324 and the archaeal phyla *Crenarchaeota* and *Nanoarchaeota* that were enriched in *V. pour-*

talesii over seawater or in one of the *V. pourtalesii* subsets. The MAGs were compared to each other within their taxonomic groups using ANI of the pangenomic workflow of *anvi'o* v5.2 (96, 97), and they were compared to seawater and other host sponge-derived reference genomes (Table S3) regarding their genome sizes and GC contents. For functional annotations, *Interproscan* v5.30-69.0, including GO term and pathway annotations, was used (98, 99). The resulting EC numbers were converted to K terms with an in-house R script using the KEGG Orthology (KO) reference hierarchy to apply the online tool Reconstruction Pathway in KEGG mapper (<https://www.genome.jp/kegg/mapper.html>). Additionally, manual search in the annotation tables (<https://doi.org/10.6084/m9.figshare.12280313>) allowed the identification of several enzymes completing some pathways. Potential transporters were identified in the above-described annotation and using the online tool TransportDB 2.0 (100).

Data deposition.

Detailed sample metadata was deposited in the PANGAEA database (<https://doi.pangaea.de/10.1594/PANGAEA.917599>). Amplicon and metagenomic raw read data were deposited in the NCBI database under BioProject PRJNA613976. Individual accession numbers for assembled MAGs are listed in Table S1. Interpro annotation output is available on figshare at <https://doi.org/10.6084/m9.figshare.1228031>.

REFERENCES

- Maldonado M, Aguilar R, Bannister RJ, Bell JJ, Conway KW, Dayton PK, Diaz C, Gutt J, Kelly M, Kenchington ELR, Leys SP, Pomponi SA, Rapp HT, Rützler K, Tendal OS, Vacelet J, Young CM. 2015. Sponge grounds as key marine habitats: a synthetic review of types, structure, functional roles, and conservation concerns, p 1–39. *In* Rossi S, Bramanti L, Gori A, Orejas C (ed), *Marine animal forests*. Springer International Publishing, Cham, Switzerland.
- Howell K-L, Piechaud N, Downie A-L, Kenny A. 2016. The distribution of deep-sea sponge aggregations in the North Atlantic and implications for their effective spatial management. *Deep Sea Res Part I Oceanogr Res Pap* 115:309–320. <https://doi.org/10.1016/j.dsr.2016.07.005>.
- Beazley L, Kenchington E, Yashayaev I, Murillo FJ. 2015. Drivers of epibenthic megafaunal composition in the sponge grounds of the Sackville Spur, northwest Atlantic. *Deep Res Part I Oceanogr Res Pap* 98:102–114. <https://doi.org/10.1016/j.dsr.2014.11.016>.
- Hawkes N, Korabik M, Beazley L, Rapp HT, Xavier JR, Kenchington E. 2019. Glass sponge grounds on the Scotian Shelf and their associated biodiversity. *Mar Ecol Prog Ser* 614:91–109. <https://doi.org/10.3354/meps12903>.
- Murillo FJ, Kenchington E, Koen-Alonso M, Guijarro J, Kenchington TJ, Sacau M, Beazley L, Rapp HT. 2020. Mapping benthic ecological diversity and interactions with bottom-contact fishing on the Flemish Cap (northwest Atlantic). *Ecol Indic* 112:106135. <https://doi.org/10.1016/j.ecolind.2020.106135>.
- Thomas T, Moitinho-Silva L, Lurgi M, Björk JR, Easson C, Astudillo-García C, Olson JB, Erwin PM, López-Legentil S, Luter H, Chaves-Fonnegra A, Costa R, Schupp PJ, Steindler L, Erpenbeck D, Gilbert J, Knight R, Ackermann G, Lopez JV, Taylor MW, Thacker RW, Montoya JM, Hentschel U, Webster NS. 2016. Diversity, structure

SUPPLEMENTAL MATERIAL

Supplemental material is available online only.

ACKNOWLEDGMENTS

The project “SponGES” received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement no. 679849. We thank our colleague Hans Tore Rapp for leading this great consortium for the last 4 years; you are dearly missed. Ship time and Canadian participation were enabled by Fisheries and Oceans Canada’s (DFO) International Governance Strategy Science Program through project “Marine Biological Diversity Beyond Areas of National Jurisdiction (BBNJ): 3-Tiers of Diversity (Genes-Species-Communities),” led by E.K. (2017 to 2019). We acknowledge financial support by Land Schleswig-Holstein within the funding programme Open Access Publikations-fonds.

We appreciated the onboard support of the crew and scientific party of the expedition *MLB2017001*. Fred Whoriskey provided the OTN mooring samples. Andrea Hethke, Ina Clefsen, and the CRC1182 Z3 team (Katja Cloppenburg-Schmidt, Malte Rühlemann, and John Baines) provided valuable support with the amplicon pipeline. Further, we thank the following people for microscopy-related support: Yu-Chen Wu, Marie Sieberns, Anke Bleyer, Cay Kruse, and Julia-Vanessa Böge. Martin Jahn provided advice for metabolic analyses.

- and convergent evolution of the global sponge microbiome. *Nat Commun* 7:11870. <https://doi.org/10.1038/ncomms11870>.
7. Pita L, Rix L, Slaby BM, Franke A, Hentschel U. 2018. The sponge holobiont in a changing ocean: from microbes to ecosystems. *Microbiome* 6:46. <https://doi.org/10.1186/s40168-018-0428-1>.
 8. Moitinho-Silva L, Nielsen S, Amir A, Gonzalez A, Ackermann GL, Cerrano C, Astudillo-García C, Easson C, Sipkema D, Liu F, Steinert G, Kotoulas G, McCormack GP, Feng G, Bell JJ, Vicente J, Björk JR, Montoya JM, Olson JB, Reveillaud J, Steindler L, Pineda M-C, Marra MV, Ilan M, Taylor MW, Polymenakou P, Erwin PM, Schupp PJ, Simister RL, Knight R, Thacker RW, Costa R, Hill RT, Lopez-Legentil S, Dailianis T, Ravasi T, Hentschel U, Li Z, Webster NS, Thomas T. 2017. The sponge microbiome project. *Gigascience* 6:1–7. <https://doi.org/10.1093/gigascience/gix077>.
 9. Maldonado M, Ribes M, van Duyl FC. 2012. Nutrient fluxes through sponges: biology, budgets, and ecological implications, p 113–182. In Becerro M, Uriz M, Maldonado M, Turon X (ed), *Advances in marine biology*, vol 62. Elsevier Ltd. Academic Press, Amsterdam, The Netherlands.
 10. de Goeij JM, van Oevelen D, Vermeij MJA, Osinga R, Middelburg JJ, de Goeij AFPM, Admiraal W. 2013. Surviving in a marine desert: the sponge loop retains resources within coral reefs. *Science* 342:108–110. <https://doi.org/10.1126/science.1241981>.
 11. Krautter M, Conway KW, Barrie JV, Neuweiler M. 2001. Discovery of a “living dinosaur”: globally unique modern hexactinellid sponge reefs off British Columbia, Canada. *Facies* 44:265–282. <https://doi.org/10.1007/BF02668178>.
 12. Leys SP, Mackie GO, Reiswig HM. 2007. The biology of glass sponges, p 1–145. In Sims D (ed), *Advances in marine biology*, vol 52. Elsevier Ltd., Academic Press, Amsterdam, The Netherlands.
 13. van Soest RWM, Boury-Esnault N, Vacelet J, Dohrmann M, Erpenbeck D, de Voogd NJ, Santodomingo N, Vanhoorne B, Kelly M, Hooper JNA. 2012. Global diversity of sponges (Porifera). *PLoS One* 7:e35105. <https://doi.org/10.1371/journal.pone.0035105>.
 14. Steinert G, Busch K, Bayer K, Kodami S, Arbizu PM, Kelly M, Mills S, Erpenbeck D, Dohrmann M, Wörheide G, Hentschel U, Schupp PJ. 2020. Compositional and quantitative insights into bacterial and archaeal communities of South Pacific deep-sea sponges (Demospongiae and Hexactinellida). *Front Microbiol* 11:716. <https://doi.org/10.3389/fmicb.2020.00716>.
 15. Tian R-M, Sun J, Cai L, Zhang W-P, Zhou G-W, Qiu J-W, Qian P-Y. 2016. The deep-sea glass sponge *Lophophysema eversa* harbours potential symbionts responsible for the nutrient conversions of carbon, nitrogen and sulfur. *Environ Microbiol* 18:2481–2494. <https://doi.org/10.1111/1462-2920.13161>.
 16. Schmidt O. 1870. Grundzüge einer Spongien-Fauna des atlantischen Gebietes. Wilhelm Engelmann, Leipzig, Germany.
 17. Beazley L, Wang Z, Kenchington E, Yashayaev I, Rapp HT, Xavier JR, Murillo FJ, Fenton D, Fuller S. 2018. Predicted distribution of the glass sponge *Vazella pourtalesi* on the Scotian Shelf and its persistence in the face of climatic variability. *PLoS One* 13:e0205505. <https://doi.org/10.1371/journal.pone.0205505>.
 18. Townsend DW, Pettigrew NR, Thomas MA, Neary MG, McGillicuddy DJ, O'Donnell J. 2015. Water masses and nutrient sources to the Gulf of Maine. *J Mar Res* 73:93–122. <https://doi.org/10.1357/002224015815848811>.
 19. Hendry KR, Cassarino L, Bates SL, Culwick T, Frost M, Goodwin C, Howell KL. 2019. Silicon isotopic systematics of deep-sea sponge grounds in the North Atlantic. *Quat Sci Rev* 210:1–14. <https://doi.org/10.1016/j.quascirev.2019.02.017>.
 20. Busch K, Beazley L, Kenchington E, Whoriskey F, Slaby B, Hentschel U. 2020. Microbial diversity of the glass sponge *Vazella pourtalesii* in response to anthropogenic activities. *bioRxiv* <https://doi.org/10.1101/2020.05.19.102806>.
 21. Dohrmann M, Wörheide G. 2017. Dating early animal evolution using phylogenomic data. *Sci Rep* 7:3599. <https://doi.org/10.1038/s41598-017-03791-w>.
 22. Mentel M, Röttger M, Leys S, Tielsens AGM, Martin WF. 2014. Of early animals, anaerobic mitochondria, and a modern sponge. *Bioessays* 36:924–932. <https://doi.org/10.1002/bies.201400060>.
 23. Mills DB, Francis WR, Vargas S, Larsen M, Elemans CP, Canfield DE, Wörheide G. 2018. The last common ancestor of animals lacked the HIF pathway and respired in low-oxygen environments. *Elife* 7:e31176. <https://doi.org/10.7554/eLife.31176>.
 24. Leys SP, Kahn AS. 2018. Oxygen and the energetic requirements of the first multicellular animals. *Integr Comp Biol* 58:666–676. <https://doi.org/10.1093/icb/icy051>.
 25. Hofmann AF, Peltzer ET, Walz PM, Brewer PG. 2011. Hypoxia by degrees: establishing definitions for a changing ocean. *Deep Sea Res Part I Oceanogr Res Pap* 58:1212–1226. <https://doi.org/10.1016/j.dsr.2011.09.004>.
 26. Parks DH, Chuvochina M, Waite DW, Rinke C, Skarshewski A, Chaumeil P-A, Hugenholtz P. 2018. A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol* 36:996–1004. <https://doi.org/10.1038/nbt.4229>.
 27. Hallam SJ, Konstantinidis KT, Putnam N, Schleper C, Watanabe Y, Sugahara J, Preston C, de la Torre J, Richardson PM, DeLong EF. 2006. Genomic analysis of the uncultivated marine crenarchaeote *Cenarchaeum symbiosum*. *Proc Natl Acad Sci USA* 103:18296–18301. <https://doi.org/10.1073/pnas.0608549103>.
 28. Park S-J, Kim J-G, Jung M-Y, Kim S-J, Cha I-T, Ghai R, Martín-Cuadrado A-B, Rodríguez-Valera F, Rhee S-K. 2012. Draft genome sequence of an ammonia-oxidizing archaeon, “*Candidatus Nitrosopumilus sediminis*” AR2, from Svalbard in the Arctic Circle. *J Bacteriol* 194:6948–6949. <https://doi.org/10.1128/JB.01869-12>.
 29. Parks DH, Rinke C, Chuvochina M, Chaumeil P-A, Woodcroft BJ, Evans PN, Hugenholtz P, Tyson GW. 2017. Recovery of nearly 8,000 metagenome-assembled genomes substantially expands the tree of life. *Nat Microbiol* 2:1533–1542. <https://doi.org/10.1038/s41564-017-0012-7>.
 30. Dudek NK, Sun CL, Burstein D, Kantor RS, Aliaga Goltsman DS, Bik EM, Thomas BC, Banfield JF, Relman DA. 2017. Novel microbial diversity and functional potential in the marine mammal oral microbiome. *Curr Biol* 27:3752–3762. <https://doi.org/10.1016/j.cub.2017.10.040>.
 31. White D, Drummond J, Fuqua C. 2012. The physiology and biochemistry of prokaryotes. Oxford University Press, Oxford, United Kingdom.
 32. Uden G, Strecker A, Kleefeld A, Bin Kim O. 2016. C4-dicarboxylate utilization in aerobic and anaerobic growth. *Ecosal Plus* 7:1–33. <https://doi.org/10.1128/ecosalplus.ESP-0021-2015>.
 33. Rosa LT, Bianconi ME, Thomas GH, Kelly DJ. 2018. Tripartite ATP-independent periplasmic (TRAP) transporters and tripartite tricarboxylate transporters (TTT): from uptake to pathogenicity. *Front Cell Infect Microbiol* 8:33. <https://doi.org/10.3389/fcimb.2018.00033>.
 34. Draskovic I, Dubnau D. 2005. Biogenesis of a putative channel protein, ComEC, required for DNA uptake: membrane topology, oligomerization and formation of disulfide bonds. *Mol Microbiol* 55:881–896. <https://doi.org/10.1111/j.1365-2958.2004.04430.x>.
 35. Sato T, Atomi H, Imanaka T. 2007. Archaeal type III Ru-BisCOs function in a pathway for AMP metabolism. *Science* 315:1003–1006. <https://doi.org/10.1126/science.1135999>.
 36. Waters E, Hohn MJ, Ahel I, Graham DE, Adams MD, Barnstead M, Beeson KY, Bibbs L, Bolanos R, Keller M, Kretz K, Lin X, Mathur E, Ni J, Podar M, Richardson T, Sutton GG, Simon M, Söll D, Stetter KO, Short JM, Noordeewier M. 2003. The genome of *Nanoarchaeum equitans*: insights into early archaeal evolution and derived parasitism. *Proc Natl Acad Sci U S A* 100:12984–12988. <https://doi.org/10.1073/pnas.1735403100>.
 37. Podar M, Anderson I, Makarova KS, Elkins JG, Ivanova N, Wall MA, Lykidis A, Mavromatis K, Sun H, Hudson ME, Chen W, Deciu C, Hutchison D, Eads JR, Anderson A, Fernandes F, Szeto E, Lapidus A, Kyrpides NC, Saier MH, Richardson PM, Rachel R, Huber H, Eisen JA, Koonin EV, Keller M, Stetter KO. 2008. A genomic analysis of the archaeal system *Ignicoccus hospitis*-*Nanoarchaeum equitans*. *Genome Biol* 9:R158.

- <https://doi.org/10.1186/gb-2008-9-11-r158>.
38. Bardy SL, Jarrell KF. 2003. Cleavage of preflagellins by an aspartic acid signal peptidase is essential for flagellation in the archaeon *Methanococcus voltae*. *Mol Microbiol* 50:1339–1347. <https://doi.org/10.1046/j.1365-2958.2003.03758.x>.
 39. Gloeckner V, Wehrl M, Moitinho-Silva L, Gernert C, Schupp P, Pawlik JR, Lindquist NL, Erpenbeck D, Wörheide G, Hentschel U. 2014. The HMA-LMA dichotomy revisited: an electron microscopical survey of 56 sponge species. *Biol Bull* 227:78–88. <https://doi.org/10.1086/BBLv227n1p78>.
 40. Moreno-Pino M, Cristi A, Gillooly JF, Trefault N. 2020. Characterizing the microbiomes of Antarctic sponges: a functional metagenomic approach. *Sci Rep* 10:645. <https://doi.org/10.1038/s41598-020-57464-2>.
 41. Horn H, Slaby BM, Jahn MT, Bayer K, Moitinho-Silva L, Förster F, Abdelmohsen UR, Hentschel U. 2016. An enrichment of CRISPR and other defense-related features in marine sponge-associated microbial metagenomes. *Front Microbiol* 7:1751. <https://doi.org/10.3389/fmicb.2016.01751>.
 42. Slaby BM, Hackl T, Horn H, Bayer K, Hentschel U. 2017. Metagenomic binning of a marine sponge microbiome reveals unity in defense but metabolic specialization. *ISME J* 11:2465–2478. <https://doi.org/10.1038/ismej.2017.101>.
 43. Kamke J, Sczyrba A, Ivanova N, Schwientek P, Rinke C, Mavromatis K, Woyke T, Hentschel U. 2013. Single-cell genomics reveals complex carbohydrate degradation patterns in poribacterial symbionts of marine sponges. *ISME J* 7:2287–2300. <https://doi.org/10.1038/ismej.2013.111>.
 44. Qin Q-L, Li Y, Sun L-L, Wang Z-B, Wang S, Chen X-L, Oren A, Zhang Y-Z. 2019. Trophic specialization results in genomic reduction in free-living marine *Idiomarina* bacteria. *mBio* 10:e02545-18. <https://doi.org/10.1128/mBio.02545-18>.
 45. Giovannoni SJ, Tripp HJ, Givan S, Podar M, Vergin KL, Baptista D, Bibbs L, Eads J, Richardson TH, Noordewier M, Rappé MS, Short JM, Cirrington JC, Mathur EJ. 2005. Genome streamlining in a cosmopolitan oceanic bacterium. *Science* 309:1242–1245. <https://doi.org/10.1126/science.1114057>.
 46. Morris JJ, Lenski RE, Zinser ER. 2012. The Black Queen hypothesis: evolution of dependencies through adaptive gene loss. *mBio* 3:e00036-12. <https://doi.org/10.1128/mBio.00036-12>.
 47. Dutta C, Paul S. 2012. Microbial lifestyle and genome signatures. *Curr Genomics* 13:153–162. <https://doi.org/10.2174/138920212799860698>.
 48. Bayer K, Schmitt S, Hentschel U. 2008. Physiology, phylogeny and *in situ* evidence for bacterial and archaeal nitrifiers in the marine sponge *Aplysina aerophoba*. *Environ Microbiol* 10:2942–2955. <https://doi.org/10.1111/j.1462-2920.2008.01582.x>.
 49. Moeller FU, Webster NS, Herbold CW, Behnam F, Doman D, Albertsen M, Mooshammer M, Markert S, Turaev D, Becher D, Rattei T, Schweder T, Richter A, Watzka M, Nielsen PH, Wagner M. 2019. Characterization of a thaumarchaeal symbiont that drives incomplete nitrification in the tropical sponge *Ianthella basta*. *Environ Microbiol* 21:3831–3854. <https://doi.org/10.1111/1462-2920.14732>.
 50. Moitinho-Silva L, Diez-Vives C, Batani G, Esteves AIS, Jahn MT, Thomas T. 2017. Integrated metabolism in sponge–microbe symbiosis revealed by genome-centered metatranscriptomics. *ISME J* 11:1651–1616. <https://doi.org/10.1038/ismej.2017.25>.
 51. Zhang S, Song W, Wemheuer B, Reveillaud J, Webster N, Thomas T. 2019. Comparative genomics reveals ecological and evolutionary insights into sponge-associated *Thaumarchaeota*. *mSystems* 4:e00288-19. <https://doi.org/10.1128/mSystems.00288-19>.
 52. Engelberts JP, Robbins SJ, de Goeij JM, Aranda M, Bell SC, Webster NS. 2020. Characterization of a sponge microbiome using an integrative genome-centric approach. *ISME J* 14:1100–1110. <https://doi.org/10.1038/s41396-020-0591-9>.
 53. Turon M, Uriz MJ. 2020. New insights into the archaeal consortium of tropical sponges. *Front Mar Sci* 6:789. <https://doi.org/10.3389/fmars.2019.00789>.
 54. Berg IA. 2011. Ecological aspects of the distribution of different autotrophic CO₂ fixation pathways. *Appl Environ Microbiol* 77:1925–1936. <https://doi.org/10.1128/AEM.02473-10>.
 55. Bayer B, Vojvoda J, Offire P, Alves RJE, Elisabeth NH, Garcia JAL, Volland J-M, Srivastava A, Schleper C, Herndl GJ. 2016. Physiological and genomic characterization of two novel marine thaumarchaeal strains indicates niche differentiation. *ISME J* 10:1051–1063. <https://doi.org/10.1038/ismej.2015.200>.
 56. Hügler M, Sievert SM. 2011. Beyond the Calvin cycle: autotrophic carbon fixation in the ocean. *Annu Rev Mar Sci* 3:261–289. <https://doi.org/10.1146/annurev-marine-120709-142712>.
 57. Rinke C, Schwientek P, Sczyrba A, Ivanova NN, Anderson IJ, Cheng J-F, Darling AE, Malfatti S, Swan BK, Gies EA, Dodsworth JA, Hedlund BP, Tsiamis G, Sievert SM, Liu W-T, Eisen JA, Hallam SJ, Kyrpides NC, Stepanauskas R, Rubin EM, Hugenholtz P, Woyke T. 2013. Insights into the phylogeny and coding potential of microbial dark matter. *Nature* 499:431–437. <https://doi.org/10.1038/nature12352>.
 58. Woyke T, Doud DFR, Eloë-Fadrosch EA. 2019. Genomes from uncultivated microorganisms, p 437–442. In Schmidt TM (ed), *Encyclopedia of microbiology*, 4th ed. Elsevier Ltd., Academic Press, Amsterdam, the Netherlands.
 59. Castelle CJ, Brown CT, Anantharaman K, Probst AJ, Huang RH, Banfield JF. 2018. Biosynthetic capacity, metabolic variety and unusual biology in the CPR and DPANN radiations. *Nat Rev Microbiol* 16:629–645. <https://doi.org/10.1038/s41579-018-0076-2>.
 60. Forterre P, Gribaldo S, Brochier-Armanet C. 2009. Happy together: genomic insights into the unique *Nanoarchaeum/Ignicoccus* association. *J Biol* 8:7. <https://doi.org/10.1186/jbiol110>.
 61. Nicks T, Rahn-Lee L. 2017. Inside out: archaeal ectosymbionts suggest a second model of reduced-genome evolution. *Front Microbiol* 8:384. <https://doi.org/10.3389/fmicb.2017.00384>.
 62. Jahn U, Summons R, Sturt H, Grosjean E, Huber H. 2004. Composition of the lipids of *Nanoarchaeum equitans* and their origin from its host *Ignicoccus* sp. strain KIN4/I. *Arch Microbiol* 182:404–413. <https://doi.org/10.1007/s00203-004-0725-x>.
 63. Albers S-V, Meyer BH. 2011. The archaeal cell envelope. *Nat Rev Microbiol* 9:414–426. <https://doi.org/10.1038/nrmicro2576>.
 64. Mell JC, Redfield RJ. 2014. Natural competence and the evolution of DNA uptake specificity. *J Bacteriol* 196:1471–1483. <https://doi.org/10.1128/JB.01293-13>.
 65. Miller MB, Bassler BL. 2001. Quorum sensing in bacteria. *Annu Rev Microbiol* 55:165–199. <https://doi.org/10.1146/annurev.micro.55.1.165>.
 66. Chun CK, Troll JV, Koroleva I, Brown B, Manzella L, Snir E, Almabrazi H, Scheetz TE, de Fatima Bonaldo M, Casavant TL, Soares MB, Ruby EG, McFall-Ngai MJ. 2008. Effects of colonization, luminescence, and auto-inducer on host transcription during development of the squid-*Vibrio* association. *Proc Natl Acad Sci USA* 105:11323–11328. <https://doi.org/10.1073/pnas.0802369105>.
 67. Zan J, Fuqua C, Hill RT. 2011. Diversity and functional analysis of *luxS* genes in vibrios from marine sponges *Mycale laxissima* and *Ircinia strobilina*. *ISME J* 5:1505–1516. <https://doi.org/10.1038/ismej.2011.31>.
 68. Ng W-L, Bassler BL. 2009. Bacterial quorum-sensing network architectures. *Annu Rev Genet* 43:197–222. <https://doi.org/10.1146/annurev-genet-102108-134304>.
 69. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. 2013. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. *Appl Environ Microbiol* 79:5112–5120. <https://doi.org/10.1128/AEM.01043-13>.
 70. Muyzer G, de Waal EC, Uitterlinden AG. 1993. Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* 59:695–700. <https://doi.org/10.1128/AEM.59.3.695-700.1993>.
 71. Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D,

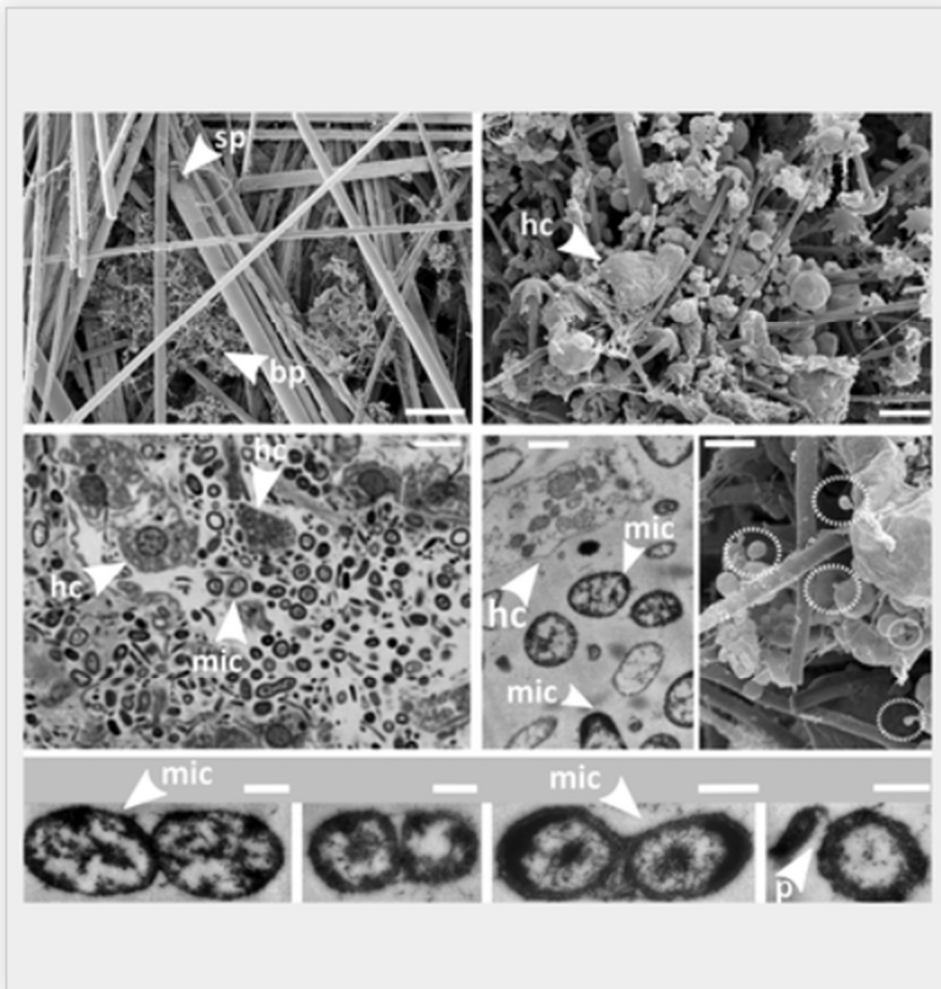
- Lozupone CA, Tumbaugh PJ, Fierer N, Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci U S A* 108:4516–4522. <https://doi.org/10.1073/pnas.1000080107>.
72. Bushnell B. 2017. BBMap short read aligner, and other bioinformatic tools. <https://sourceforge.net/projects/bbmap/>.
 73. Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Da Silva R, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kachler BD, Kang KB, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolk T, Kreps J, Langille MGI, Lee J, Ley R, Liu Y-X, Lofffield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik AV, Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Tumbaugh PJ, Ul-Hasan S, van der Hoof JJJ, Vargas F, Vázquez-Baeza Y, Vogtmann E, von Hippel M, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG. 2019. Reproducible, interactive, scalable, and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 37: 852–857. <https://doi.org/10.1038/s41587-019-0209-9>.
 74. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods* 13:581–583. <https://doi.org/10.1038/nmeth.3869>.
 75. Bokulich NA, Kachler BD, Rideout JR, Dillon M, Bolyen E, Knight R, Huttley GA, Caporaso JG. 2018. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome* 6:90. <https://doi.org/10.1186/s40168-018-0470-z>.
 76. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* 41:590–596. <https://doi.org/10.1093/nar/gks1219>.
 77. Price MN, Dehal PS, Arkin AP. 2010. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. <https://doi.org/10.1371/journal.pone.0009490>.
 78. R Core Team. 2008. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
 79. The Inkscape Project. 2005. Inkscape. <http://www.inkscape.org/>.
 80. QGIS Development Team. 2017. Geographic information system. Open Source Geospatial Foundation Project. <http://qgis.osgeo.org>.
 81. McDougall TJ, Barker PM. 2011. Getting started with TEOS-10 and the Gibbs Seawater (GSW) oceanographic toolbox. SCOR/IAPSO Working Group, Newark, DE.
 82. Aeschl E, Büchl-Zimmermann S, Burmester A, Dänhardt-Pfeiffer S, Desel C, Hamers C, Jach G, Kässens M, Makovitzky J, Mulisch M, Nixdorf-Bergweiler B, Pütz D, Riedelsheimer B, van den Boom F, Wegerhoff R, Welsch U. 2010. Romeis mikroskopische technik. Spektrum Akademischer Verlag, Heidelberg, Germany.
 83. Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
 84. Li D, Luo R, Liu C-M, Leung C-M, Ting H-F, Sadakane K, Yamashita H, Lam T-W. 2016. MEGAHIT v1.0: a fast and scalable metagenome assembler driven by advanced methodologies and community practices. *Methods* 102:3–11. <https://doi.org/10.1016/j.ymeth.2016.02.020>.
 85. Uritskiy GV, DiRuggiero J, Taylor J. 2018. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. *Microbiome* 6:158. <https://doi.org/10.1186/s40168-018-0541-1>.
 86. Kang DD, Froula J, Egan R, Wang Z. 2015. MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* 3:e1165. <https://doi.org/10.7717/peerj.1165>.
 87. Kang DD, Li F, Kirton E, Thomas A, Egan R, An H, Wang Z. 2019. MetaBAT 2: an adaptive binning algorithm for robust and efficient genome reconstruction from metagenome assemblies. *PeerJ* 7:e7359. <https://doi.org/10.7717/peerj.7359>.
 88. Wu Y-W, Simmons BA, Singer SW. 2016. MaxBin 2.0: an automated binning algorithm to recover genomes from multiple metagenomic datasets. *Bioinformatics* 32:605–607. <https://doi.org/10.1093/bioinformatics/btv638>.
 89. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <https://doi.org/10.1089/cmb.2012.0021>.
 90. Hyatt D, Chen G-L, LoCascio PF, Land ML, Larimer FW, Hauser LJ. 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. <https://doi.org/10.1186/1471-2105-11-119>.
 91. Eddy SR. 2011. Accelerated profile HMM searches. *PLoS Comput Biol* 7:e1002195. <https://doi.org/10.1371/journal.pcbi.1002195>.
 92. Matsen FA, Kodner RB, Armbrust EV. 2010. pplacer: linear time maximum-likelihood and Bayesian phylogenetic placement of sequences onto a fixed reference tree. *BMC Bioinformatics* 11:538. <https://doi.org/10.1186/1471-2105-11-538>.
 93. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. 2018. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun* 9:5114. <https://doi.org/10.1038/s41467-018-07641-9>.
 94. Letunic I, Bork P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res* 44:W242–W245. <https://doi.org/10.1093/nar/gkw290>.
 95. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C. 2011. Metagenomic biomarker discovery and explanation. *Genome Biol* 12:R60. <https://doi.org/10.1186/gb-2011-12-6-r60>.
 96. Delmont TO, Eren AM. 2018. Linking pangenomes and metagenomes: the *Prochlorococcus* metapangenome. *PeerJ* 6:e4320. <https://doi.org/10.7717/peerj.4320>.
 97. Eren AM, Esen ÖC, Quince C, Vineis JH, Morrison HG, Sogin ML, Delmont TO. 2015. Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3:e1319. <https://doi.org/10.7717/peerj.1319>.
 98. Jones P, Binns D, Chang H-Y, Fraser M, Li W, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, Pesseat S, Quinn AF, Sangrador-Vegas A, Scheremetjew M, Yong S-Y, Lopez R, Hunter S. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30: 1236–1240. <https://doi.org/10.1093/bioinformatics/btu031>.
 99. Sangrador-Vegas A, Mitchell AL, Chang H-Y, Yong S-Y, Finn RD. 2016. GO annotation in InterPro: why stability does not indicate accuracy in a sea of changing annotations. *Database* 2016:baw027. <https://doi.org/10.1093/database/baw027>.
 100. Elbourne LDH, Tetu SG, Hassan KA, Paulsen IT. 2017. TransportDB 2.0: a database for exploring membrane transporters in sequenced genomes from all domains of life. *Nucleic Acids Res* 45:D320–D324. <https://doi.org/10.1093/nar/gkw1068>.
 101. Dever M, Hebert D, Greenan BJW, Sheng J, Smith PC. 2016. Hydrography and coastal circulation along the Halifax line and the connections with the Gulf of St. Lawrence. *Atmos-Ocean* 54:199–217. <https://doi.org/10.1080/07055900.2016.1189397>.
 102. Fratantoni PS, Pickart RS. 2007. The Western North Atlantic Shelfbreak Current System in summer. *J Phys Oceanogr* 37:2509–2533. <https://doi.org/10.1175/JPO3123.1>.



AMERICAN
SOCIETY FOR
MICROBIOLOGY

mSystems
AN OPEN ACCESS JOURNAL PUBLISHED BY
THE AMERICAN SOCIETY FOR MICROBIOLOGY

Cover image July/August 2020; volume 5, issue 4



Chapter 5

Population connectivity of fan-shaped sponges and their associated microbiomes in the deep Cantabrian Sea

© 2020 Busch et al.

Original publication by *Deep Sea Research Part I: Oceanographic Research Papers* (Elsevier)

Reprinted under the terms of the Creative Commons Attribution License (CC-BY 4.0).

doi: 10.1016/j.dsr.2020.10342.

Population connectivity of fan-shaped sponge holobionts in the deep Cantabrian Sea

Kathrin Busch¹, Sergi Taboada^{2,3,4}, Ana Riesgo^{3,4}, Vasiliki Koutsouveli³, Pilar Ríos^{4,5}, Javier Cristobo^{4,6}, Andre Franke⁷, Klaus Getzlaff¹, Christina Schmidt¹, Arne Biastoch^{1,8}, Ute Hentschel^{1,8*}

¹ GEOMAR Helmholtz Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany, ² Universidad Autónoma de Madrid, Departamento de Biología (Zoología), Cantoblanco 28049, Madrid, Spain, ³ The Natural History Museum of London, Cromwell Road, London SW7 5BD, United Kingdom, ⁴ Universidad de Alcalá, Departamento de Ciencias de la Vida, Apdo. 20, 28805, Alcalá de Henares, Spain, ⁵ Instituto Español de Oceanografía. Centro Oceanográfico de Santander C/ Promontorio San Martín s/n, 39004 Santander, Spain, ⁶ Instituto Español de Oceanografía. Centro Oceanográfico de Gijón, C/ Príncipe de Asturias 70 bis, 33212 Gijón, Asturias, Spain, ⁷ Christian-Albrechts University of Kiel, Institute of Clinical Molecular Biology (IKMB), Rosalind-Franklin-Straße 12, 24105 Kiel, Germany, ⁸ Christian-Albrechts University of Kiel, Christian-Albrechts-Platz 4, 24118 Kiel, Germany. *Correspondence author.

Connectivity is a fundamental process driving the persistence of marine populations and their adaptation potential in response to environmental change. In this study, we analysed the population genetics of two morphologically highly similar deep-sea sponge clades (*Phakellia hironellei* and the ‘Topsentia-and-Petromica’ clade, (hereafter referred to as ‘TaP clade’)) at three locations in the Cantabrian Sea and simultaneously assessed the corresponding host microbiome by 16S rRNA gene sequencing. A virtual particle tracking approach (Lagrangian modelling) was applied to assess oceanographic connectivity in the study area. We observed overall genetic uniformity for both sponge clades. Notably, subtle genetic differences were observed for sponges of the TaP clade and also their microbiomes between a canyon and bank location, < 100 km apart and with the same depth range. The Lagrangian model output suggests a strong retention of larvae in the study area with variable inter-annual connectivity via currents between the three sampling regions. We conclude that geologic features (canyons) and the prevailing ocean currents may dictate sponge holobiont connectivity and that differentiation can emerge even on small spatial scales.

Keywords: Porifera, population genetics, single-nucleotide polymorphisms (SNPs), amplicon sequencing, Lagrangian modelling, ocean sensing

INTRODUCTION

The ocean is the largest interconnected habitat on planet Earth (Ramirez-Llodra et al., 2011). At the same time, it is a highly dynamic system showing natural fluctuations, but also discrete responses to human impacts. Connectivity studies are important for our understanding regarding the resilience of ecosystems to changing oceanic conditions, as well as for an evaluation of population vulnerability to environmental change and anthropogenic stressors (Fox et al., 2016; Kenchington et al., 2019). Understanding population- and ecosystem-connectivity in the ocean is crucial for the design of

appropriate conservation measures, in particular for the design of marine protected area (MPA) networks (White et al., 2014; Gallego et al., 2017; Andrello et al., 2017; Kenchington et al., 2019). Many marine benthic organisms have pelagic larvae that ensure the maintenance of genetic diversity and allow for a colonization of new habitats (e.g. review by Levin, 2006; Cowen and Sponaugle, 2009). Larval dispersal is difficult to observe directly in the field and in particular in the deep-sea. Therefore, indirect methods such as molecular markers and virtual particle tracking are commonly applied

to analyse genetic connectivity between organisms (e.g. review by Cowen and Sponaugle, 2009; Baltazar-Soares et al., 2014; Breusing et al., 2016)). To conduct virtual particle tracking, biophysical models based on a Lagrangian approach (i.e. individual particle tracking in space and time; Cowen and Sponaugle, 2009) are powerful tools. As the term 'biophysical' implies, in these approaches biological parameters (e.g. time point of larval release and pelagic larval duration) are integrated into a physical framework (i.e. ocean current models) to acquire large ensembles of passive drift trajectories, which represent larval dispersal (Breusing et al., 2016). Information about the biology of the studied organism, including its reproductive cycle, is crucial to fine-tune respective models, but this kind of information is generally not available for the deep ocean. Along these lines, the number of population genetic studies addressing connectivity is still very small (Taylor and Roterman, 2017). Previous studies suggest that the offspring of sessile benthic invertebrates is typically transported 10-100 km distance away from the parents (Pallumbi, 2004). However, identification and exploration of biophysical barriers, which may inhibit gene flow, is fundamental to understanding population connectivity. While depth is commonly known to be a main factor related to species and population structure (e.g. Van Soest and de Voogd, 2015; Taylor and Roterman, 2017; Indraningrat et al., 2019), the barriers separating regions within a similar depth range are less understood.

The geology of the seafloor is relevant for the formation and maintenance of biophysical barriers on different spatial scales. On large spatial

scales, massive features such as shelf breaks, trenches and ridges direct major ocean currents. On regional scales, the seafloor may lead to up- or downwelling, or circulatory water flows. Seamounts and banks are prominent elevated geologic features on the regional scale. These settings are often habitats for diverse macroorganisms (Morato et al., 2010) and microorganisms (Busch et al., 2020a), which frequently are endemic to that location (De Forges et al., 2000). Submarine canyons are prominent plunging geologic structures which represent crucial connections between productive shelf waters and the deep open ocean (Martín et al., 2006) but may also potentially lead to retention of organic matter (Hickey et al., 1986).

Around the northern shores of Spain, the continental shelf is comparably narrow (Gómez-Ballesteros et al., 2014) (**Figure 1**) and streaked with numerous submarine canyons. The Avilés Canyon System, which is located in front of the Asturian coast, is one of the largest submarine canyons in Europe (Rumín-Caparrós et al., 2016). Overall, the seafloor on the Cantabrian shelf has a high structural complexity, with elevated topography (e.g. banks) occurring in close proximity to plunging geologic features (e.g. canyons). Along these lines, the Le Danois Bank (Spanish name: El Cachucho) is found within a few kilometers distance next to the Avilés Canyon System (**Figure 1**). Le Danois Bank is a designated Marine Protected Area (MPA) and the Avilés Canyon System is a designated Special Area of Conservation (SAC), as both areas harbor a great diversity of benthic organisms (Sánchez et al., 2008; Sánchez et al., 2014).

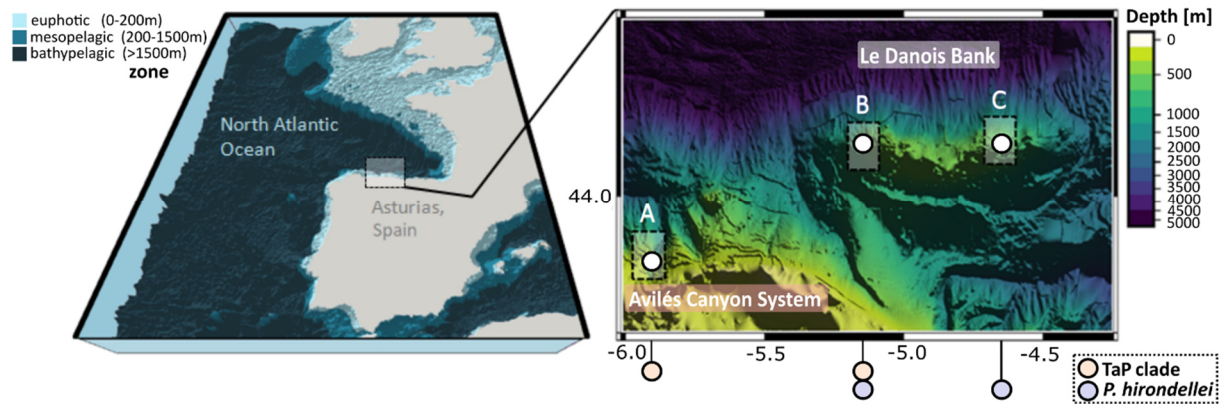


FIGURE 1 | Combined bathymetric maps of the study area showing an overview of the study area in the Asturias region on the Spanish continental margin and a close-up of the three focus regions at the Avilés Canyon System (A: ‘Canyon’) and the Le Danois Bank (B: ‘West-Bank’ and C: ‘East-Bank’). The colored circles below the map indicate the spatial distribution of the fan-shaped sponge samples included in this study.

Sponges are prominent and abundant benthic organisms in the deep North Atlantic Ocean (Klitgaard et al., 1997; Klitgaard and Tendal, 2004; Murillo et al., 2012). Dense aggregations of sponges in the deep-sea are commonly referred to as ‘sponge grounds’. Sponge grounds are often characterized as Vulnerable Marine Ecosystems (FAO 2009; FAO 2016) as they can be affected by anthropogenic activities such as bottom trawling (Kazanidis et al., 2019; Busch et al., 2020b) and oil and gas exploitation (Vad et al., 2018). These threats stand opposed to the high ecological value of deep-sea sponge grounds. The ecological value of deep-sea sponge grounds includes enhancement of local biodiversity (Beazley et al., 2013), formation of biomass hot-spots, and representation of a crucial role in biogeochemical cycling (Cathalot et al., 2015).

Sponges live in close association with dense and diverse symbiotic microbial consortia (Easson and Thacker, 2014; Hentschel et al., 2012; Thomas et al., 2016). The microbiomes of sponges are represented by diverse bacterial and archaeal clades with > 63 prokaryotic phyla having been identified so far (Thomas et al., 2016; Moitinho-Silva et al., 2017a). Microbial symbionts were shown to contribute to the metabolism of the sponge, i.e., via carbon and nitrogen cycling as well as vitamin production and defense (reviewed in Pita et al., 2018). Sponges and their associated microbial communities (hereafter termed ‘holobionts’) are considered fine-tuned entities

where both partners are tailored to function together. At the population-level, previous studies found sponge host population genetics and location to be related with the structure of the associated microbiomes at variable degrees, depending on the level of gene flow between the hosts (Griffiths et al., 2019; Díez-Vives et al., 2020; Easson et al., 2020).

The present study aimed to explore the host genetic and microbial connectivity of sponge holobionts on a small geographic gradient < 100 km. We focused on fan-shaped sponges that are commonly found at the Avilés Canyon System and the Le Danois Bank of the Cantabrian Sea (**Figure 1**). These sponges of the ‘Topsentia-and-Petromica’ clade (hereafter referred to as ‘TaP clade’) and *Phakellia hirondellei* are morphologically very similar even to the expert eye, but belong to different sponge orders. We sought to address whether genetic diversification can be resolved in deep-sea sponge holobionts sampled less than 100 km apart and what effect the geological setting (canyon versus bank) might have on population structure. In addition, oceanographic connectivity in the Cantabrian Sea was assessed by a virtual larvae tracking approach. Given that the continental shelf is comparably narrow in the study area, we were particularly interested to evaluate whether virtual sponge larvae would travel along the bathymetry. The presented findings contribute - in the long run - to a better management and preservation of vulnerable marine ecosystems.

MATERIAL AND METHODS

Field work

42 fan-shaped sponge specimens were collected with a rock dredge onboard RV Ángeles Alvariño from the area of the Avilés Canyon System (~ 43.9 °N; 6.0 °W) and the Le Danois Bank (~ 44.1 °N; 5.0 °W) in June 2017 during the expedition *SponGES0617*. In addition, 15 full depth CTD casts were performed at the respective and additional sampling stations. The focus of this study was on three sampling regions: region A ('Canyon'), region B ('West-Bank'), and region C ('East-Bank'), (Figure 1). Sampling at all three sampling regions was conducted within a similar depth range (region A: 695 m, region B: 653 m, and region C: 541 m). Bathymetric data of the study area were extracted from the Bathymetry Data Portal of the European Marine Observation and Data Network (EMODnet). Samples of different sponge species were collected in the same grab. The sample details, corresponding (meta-) data and raw sequence NCBI-accession numbers for all methodological approaches of all seawater and sponge samples presented in this study were deposited in the PANGAEA data repository (<https://doi.org/10.1594/PANGAEA.923271>).

After collection, sponge specimens were rinsed with fresh seawater and photographed. Tissue subsamples were taken for microbial work and instantly flash-frozen. Sponge fragments of each specimen (1-3 cm³) were preserved in 96 % EtOH for molecular analysis (see below), and immediately kept at -20 °C; EtOH was replaced once after one day preservation. Small sponge fragments were also incubated in sodium hypochlorite solution (Panreac 5 % w/v technical grade) and kept at room temperature until the end of the cruise for spicule analysis (see below). The remainder of the specimens not used for molecular and morphological analyses was preserved in 80 % EtOH and kept at room temperature in the reference collection at the Oceanographic Centre of Gijón (IEO, Spain). Seawater samples (2 L) for microbial community analyses were

derived from each CTD casts' bottom depth (~ 5 m above seafloor), filtered onto PVDF filter membranes (pore size = 0.22 µm and ø = 47 mm) and stored at -80 °C until further processing.

Sponge morphological analyses

For spicule preparation, soft sponge tissue was digested with sodium hypochlorite solution (Panreac 5 % w/v technical grad.). The remaining spicules were soaked 3 x 1 h with distilled water in a water bath at room temperature which was removed by repeated centrifugation (1 min at 12000 rpm). Finally, absolute ethanol was added. A few drops of ethanol and spicules were placed on a slide which was then flamed for dehydration. A few drops of DurcupanTM ACM embedding mixture for microscopy (Sigma-Aldrich) were placed on the coverslip which was then kept in the oven at 36 °C for 24 h before visualisation of spicules was performed using a microscope Nikon Eclipse 50i.

Sponge barcoding and phylogenetic analyses

DNA of all sponge specimens was extracted using the DNeasy Power Soil Kit (Qiagen) with approximately 0.25 g sponge tissue. The quality and quantity of resulting extracts were checked with a NanoDrop spectrophotometer. Fragments of the 18S rDNA (18S) gene were amplified for all sponges using the primer pair SP18aF-SP18gR (SP18aF: 5'-CCTGCCAGTAGTCATATGCTT-3'; SP18gR: 5'-CCTTGTTACGACTTTTACTTCTC-3'; Redmond et al., 2013). The polymerase chain reaction (PCR) program for 18S was 95 °C/2 min - (95 °C/2 min - 57 °C/45 sec - 72 °C/30 sec) x 30 cycles - 72 °C/3 min. For the amplification of a fragment of the cytochrome c oxidase subunit I (COI), we used two sets of primers: Pcant-COIF (5'-TTTGCAGGGATGATCGGAAC-3') and Pcant-COIR (5'-CCCGGGGCCCTCATATTTAA-3') to obtain a fragment of 708 base pairs (for more details see Taboada et al. (in review)) for *P. hirondellei* and LCO1490 and HCO2198 (Folmer et al., 1994) to obtain 685 bp. The PCR program

for *COI* was 94 °C/5 min - (94 °C/30 s - 58 °C/30 s - 72 °C/30 s) x 35 cycles - 72 °C/10 min. Amplification of both *18S* and *COI* was performed in 12.5 µL reactions, using 10.5 µL of VWR Red Taq DNA Polymerase 1.1x Master Mix (VWR International bvba/sprl, Belgium), 0.5 µL of the forward and reverse primers, and 1 µL of DNA template. PCR products were verified by gel electrophoresis on 1 % agarose. Sequencing was conducted with the primers mentioned above, using (dideoxy chain termination / cycle sequencing) on an ABI 3730XL DNA Analyser (Applied Biosystems, USA) by Eurofins Genomics for *18S* and at the Molecular Core Labs (Sequencing Facility) of the Natural History Museum of London (NHM) for *COI*.

Overlapping sequence fragments of *18S* and *COI* were assembled and trimmed separately into consensus sequences using the software Geneious v.10.1.3 (Kearse et al., 2012). Occasionally only forward or reverse sequences were used due to the poor quality of one of the fragments (for all *18S* and three *COI* sequences). Consensus sequences were then checked for contamination using BLAST (Altschul et al., 1990), aligned with MAFFT v.7.309 (Kato and Standley, 2013) using sequences of closely related species, and used to construct a phylogenetic hypothesis with Maximum Likelihood using RAxML 8.1.22 (Stamatakis, 2014). The evolutionary model was selected using jModelTest (Darriba et al., 2012), resulting in GAMMAGTR (GTR+G) for both markers. Phylogenetic analyses for *18S* and *COI* were run separately ten times, with 100 replicates for bootstrap recovery. Our taxonomic distinction between species is then based on *18S* and *COI* phylogenetic analyses (see below) and spicule confirmation. This procedure, of combining classical sponge taxonomy with molecular sequencing, left us with 21 unambiguously identified sponge specimens for the present study, which are comprised of 11 TaP clade individuals and 10 *P. hirondellei* individuals.

Sponge reproductive state analyses

An assessment of the sponge reproductive state was conducted as biological ground-truthing for the biophysical particle tracking model. Our main aim of this analysis was to derive a reasonable timepoint of larval release into the water column. For the assessment, small pieces (5 mm³) of five TaP clade individuals and three *P. hirondellei* were fixed in a solution of 2.5 % glutaraldehyde, 0.4M PBS and 0.34 NaCl onboard and stored at 4 °C (modified after Koutsouveli et al., 2018). In the home laboratory, samples were rinsed three times with buffer (0.4 M PBS-0.6 M NaCl) and then postfixed in 2 % osmium tetroxide for two hours at 4 °C. After rinsing the sample pieces three additional times (with an incubation time of 15 min per washing step, at 4 °C), partial dehydration was conducted with an ascending ethanol series (2x 30 %- 1x 50 %- 1x 70 %). In the next step, sponge pieces were submerged in 5 % hydrofluoric acid overnight to remove any silica remnants from their skeleton. Subsequently, samples were washed carefully (8x 70 % EtOH) and dehydration was continued with an increasing gradient of ethanol (1x 90 %- 2x 100 %). Once dehydrated, samples were embedded in LR-White (LWR) resin, using a gradient of LRW/EtOH (1x 70/30– 1x 50/50– 1x 30/70– 2x 100). Afterwards, samples were transferred into fresh LWR and polymerized inside embedding capsules in an oven (57 °C) for two days. Blocks of LWR were cut using a Reichert Ultracut ultramicrotome equipped with a diamond knife (DIATOME, Switzerland) at 0.5–2 µm (semi-thin sections) and 70–90 nm (ultrathin sections) respectively. The semi-thin sections were stained with toluidine blue or Richardson solution (latter prepared as described in (Mulisch and Welsch, 2015)), while the ultra-thin sections for transmission electron microscopy (TEM) were contrasted with uranyl acetate and Reynold's lead citrate. Ultra-thin sections were visualised with a Hitachi Transmission Electron Microscope TEM (H-7650) and a Tecnai G2 Spirit Bio Twin TEM (FEI Company) at 90 kV. Semi-thin sections

were visualised with an Axio Observer.Z1 microscope (Zeiss, Germany).

Sponge population genetic analysis

ddRADseq libraries were performed for a total of 21 individuals: 11 TaP clade specimens and 10 *P. hirondellei*. Library preparation was conducted following (Peterson et al., 2012) with the following modifications by (Combosch et al., 2017). Double-stranded genomic DNA (500 ng) was digested using the high-fidelity restriction enzymes SbfI and EcoRI (New England Biolabs). Resulting digested fragments were cleaned by manual pipetting using Agencourt AMPure beads (1.5X volume ratio; Beckham Coulter), and were subsequently quantified with a Qubit dsDNA HS assay (Life Technologies). In the next step, resulting fragments were ligated to custom-made P1 and P2 adapters containing sample-specific barcodes and primer annealing sites. Barcoded individuals were pooled into libraries, cleaned by manual pipetting using AMPure beads (1.5X volume ratio), and size-selected (range sizes 200–400 bp) using a Blue Pippin Prep (Sage Science). Each library was PCR-amplified using Phusion polymerase (Thermo Scientific) and a different set of PCR primers with barcodes in order to create multiplex libraries. The PCR program used was 98 °C/30 s – (98 °C/10 s – 65 °C/30 s – 72 °C/1.5 min) x 12 cycles – 72 °C/10 min. Resulting PCR products were cleaned by manual pipetting using Agencourt AMPure beads (1.5X volume ratio), quantified with a Qubit dsDNA HS assay, and quality-checked on a TapeStation 2200 (Agilent Technologies). Libraries were pooled normalizing their concentration, and pooled together with RNA-seq libraries in the same flow cell. Libraries pair-end sequenced (150 bp) were run on an Illumina HiSeq 4000 (Illumina) at Macrogen Inc. (South Korea).

Quality filtering of reads and locus assembly was conducted with the Stacks pipeline, v2.4.1 (Catchen et al., 2013). RAD-tags (DNA fragments with the two appropriate restriction enzyme cut sites that were selected, amplified, and sequenced) were processed

using *process_radtags*, where raw reads were quality-trimmed to remove low quality reads, reads with uncalled bases, and reads without a complete barcode or restriction cut site. The *process_radtags* rescue feature (-r) was used to recover minimally diverged barcodes and RAD-tags (--barcode_dist 3; --adapter_mm 2). The *process_radtags* trimming feature (-t) was used to trim remaining reads to 120 bp, in order to increase confidence in single-nucleotide polymorphism (SNP) calling. After performing these filtering steps in *process_radtags*, we retained a total of 74,442,388 reads from the initial 102,673,760 raw reads (72.5 % of the total reads retained), with an average of 3,544,876 reads per sample.

Preliminary tests were carried out following (Jeffries et al., 2016) and (Paris et al., 2017) in order to identify optimal Stacks parameters (including m, M, and n) for our dataset. Briefly, tests were carried out for 5 sets of 3 randomly chosen individuals and, for each test, all non-test parameters were kept as default. The *Stacks* populations module was run to filter data with $r = 0.8$ for each test, and the number of assembled loci, number of polymorphic loci, number of SNPs, and coverage was compared between the tests. Final parameter values were as follows: *ustacks*: M = 2, m = 3; *cstacks*: n = 1. Mean locus coverage among all samples was 47.1 ± 13.0 , ranging from 15.8 to 99.5.

The *Stacks* populations module was used to conduct a first filtering of the data, retaining those SNPs present in at least 70 % of the individuals ($r = 0.7$), and just retaining the first SNP from each RAD-tag using “--write_single_SNP” in order to reduce the linkage disequilibrium among loci. A subsequent more accurate filtering was performed using the *adegenet* R package (Jombart, 2008; R Development Core Team, 2008; Jombart and Ahmed, 2011), assessing SNP distributions across individual samples and sampling regions, and testing different filtering thresholds in order to maximise the number of retained SNPs and minimise missing data. This

approach provides significant help in defining final thresholds in comparison with the Stacks populations module. The resulting assessment resulted in no further filtering of samples. Given the known presence of symbiotic bacteria in all sponges, the resulting set of sequences containing variable SNPs obtained after running populations were filtered for bacterial hits. This was done using *-blastn* comparing the aforementioned set of sequences against a nr database for bacteria extracted from NCBI (accessioned on the 17/06/2018), using a e-value of 1e-6 or lower. This filtering of bacteria resulted in 0 being removed in both the *P. hirondellei* and TaP clade datasets. That said, only polymorphic SNPs were used.

We calculated population genetic diversity and demographic statistics separately for TaP clade specimens and *P. hirondellei* by grouping the samples per sampling region. Expected (H_e) and observed (H_o) heterozygosities, and inbreeding coefficients (FIS) were calculated per each sponge clade per sampling region using Genodive v.3.02 (Meirmans and Van Tienderen, 2004). To assess global inbreeding within sampling regions and differentiation among them, we also calculated global inbreeding coefficient FIS in Genodive.

We assessed the population structure for each sponge clade separately using three different methods: STRUCTURE v.2.3 (Pritchard et al., 2000); the function *snapclust* in the adegenet R package (Beugin et al., 2018); and the discriminant analysis of principal components (DAPC) as implemented in the adegenet R package (Jombart et al., 2010). We ran STRUCTURE with 50,000 MCMC iterations using the admixture model, with a burn-in of 20,000 iterations, setting the putative K from 1 to 4 with 10 replicates for each run. We used STRUCTURE HARVESTER (Earl and vonHoldt, 2012) and CLUMPP v.1.1.2 (Jakobsson and Rosenberg, 2007) to determine the most likely number of clusters and to average each individual's membership coefficient across the

K value replicates, respectively. We used the Akaike Information Criterion (AIC) to identify the optimal number of clusters in *snapclust.choose.k*, and then initial memberships for *snapclust* were chosen using the k-means algorithm (*pop.ini* = "kmeans"), allowing a maximum K (number of clusters) of 12 (*max* = 12), and a *n.start.kmeans* of 100 (*n.start.kmeans*= 100). For the DAPC analysis we grouped samples by sampling region, and retained the number of principal components analysis (PCA) axes and eigen values using the cross-validation *xvalDapc* function from the adegenet R package. Finally, Pairwise FST values were estimated to measure the differentiation between pairs of sampling regions using Genodive v.3.02 (Meirmans and Van Tienderen, 2004) with 20,000 permutations.

Sponge-associated microbial community analyses

DNA extraction of all fan-shaped sponges was performed using the DNeasy Power Soil Kit (Qiagen) with approximately 0.25 g sponge tissue or half a seawater filter (for 46 seawater samples in total) used as input material. The concentration of resulting DNA extracts was checked with a NanoDrop spectrophotometer and their quality assessed by gel electrophoresis after a polymerase chain reaction (PCR) with the universal 16 rRNA gene primers 27F and 1492R. After the quality check, a one-step PCR (30x cycles: 98 °C/30 s - 98 °C/9 s - 55 °C/1 min - 72 °C/1.30 min - 72 °C/10 min - hold at 4°C) was conducted on the extracts to amplify the V3 to V4 variable regions of the 16S rRNA gene (using the primer pair 341F 5'-CCTACGGGAGGCAGCAG-3' (Muyzer et al., 1993) & 806R 5'-GGACTACHVGGGTWTCTAAT-3' (Caporaso et al., 2011) in a dual-barcoding approach (Kozich et al., 2013)). The amplicon libraries were quality checked by gel electrophoresis, normalised with the SequalPrep Normalization Plate Kit (ThermoFisher Scientific) and pooled equimolarly. Sequencing was performed on a MiSeq platform (MiSeqFGx, Illumina) using v3

chemistry and in subsequent demultiplexing 0 mismatches were allowed in the barcode sequences.

Sequences were processed using QIIME2 (version 2018.11, (Bolyen et al., 2018), similar to the methods described in Busch et al., (2020c). Briefly, the DADA2 algorithm (Callahan et al. 2016) was applied on forward reads (truncated to 270nt) to generate Amplicon Sequence Variants (ASVs), which in turn were used to calculate phylogenetic ASV trees with the FastTree2 plugin. Classification of representative ASV sequences (on the genus, family, order, class and phylum level) was performed by a primer-specific trained Naive Bayes taxonomic classifier, using the Silva 132 99% OTUs 16S database (Quast et al., 2013).

Statistical analyses of the microbial community were performed on all 21 sponge individuals, including eight TaP clade individuals from region A ('Canyon'), six individuals from region B ('West-Bank') (comprised of three TaP clade individuals + three *P. hirondellei* individuals), and seven *P. hirondellei* specimens from region C ('East-Bank'). Further, nine seawater samples were included in the statistical analyses (three samples from each region). Alpha diversity indices (among others Shannon index), as well as beta diversity metrics (among others weighted UniFrac distances) were calculated. Non-metric multidimensional scaling (nMDS) was performed on weighted UniFrac distances to evaluate sample separation in ordination space. To assess if groups of samples differed significantly from each other, permutational multivariate analyses of variance (PERMANOVA) were conducted in a pairwise manner on weighted UniFrac distances. Phyla differing significantly between sample groups were determined and ranked using the Linear Discriminant Analysis Effect Size (LEfSe) algorithm (Segata et al., 2011). We applied a significance level of $\alpha = 0.05$ for all statistical analyses throughout this study. Plotting was performed with R (version 3.0.2,

(R Development Core Team, 2008)) and further fine-tuning of visualisations conducted using Inkscape (version 0.92.3; (Harrington and Team, 2005)) and/or GIMP (version 2.8).

Larval Dispersal Modeling and Oceanographic Observations

Before running the virtual particle tracking model simulations, the following steps were performed (i) validation of technical input parameters (number of released particles), (ii) further validation of biological input parameters (timepoint of larval release), (iii) comparison of model output (TS data) with in situ measurements.

Validations with biological and physical oceanographic observations

To further validate the timepoint of larval release into the water column (as determined by the sponge reproductive state analyses), we conducted an assessment of pelagic productivity. This seemed reasonable as several previous studies have described a strong link between pulses of primary productivity and the larval release of benthic organisms (e.g. Highfield et al., 2010). We assessed productivity by two means: (a) based on in situ biological data of bacterioplankton and chlorophyll-a concentrations (snapshot in time, relying on sampling in June 2017); and (b) based on remote sensing (assessment of temporal variations within the year 2017). The latter was performed with the help of publicly available resources, as we derived monthly chlorophyll-a concentrations of the year 2017 from satellite data accessed via the GlobColor web portal (provided by the European Space Agency, ESA). In situ productivity data was retrieved by flow cytometry using a Beckman Coulter Gallios machine. Samples for flow cytometry were derived in triplicates from bottom depths of all 15 performed CTD casts. Subsamples (4 mL seawater) were fixed with 200 μ L glutaraldehyde (GDA, 25 %). Sample-containing tubes were stored in vertical position at -20 °C until processing. For measurement of bacterial cell numbers, the

samples were thawed, prefiltered (syringe filters with 50 μm pore size), and sample aliquots of 400 μL were mixed with 10 μL of a SYBR Green stock solution (10,000x concentrate in DMSO). Fluoresbrite fluorescence microspheres with a diameter of 1 μm were added, followed by a product incubation time of 15 min. For measurements of phytoplankton concentrations, no staining was performed, but detection of autofluorescence conducted instead.

To validate our physical model output with in situ physical measurements (of TS data), and to get a better understanding of prevailing water mass properties, we performed in situ sensing using a Seabird 37 CTD sensor system. The bathymetry map was created with QGIS (version 3.4.4; (QGIS Development Team, 2017)). CTD data and satellite data were visualised with Python (version 3.7.3).

Biophysical modelling

We applied a Lagrangian modelling framework using the particle tracking toolbox Parcels (version 2.0.0.beta2, Delandmeter and Van Sebille, 2019) with three-dimensional velocity data from an Atlantic model based on the Nucleus for European Modelling of the Ocean (NEMO v3.6, Madec, 2016) system. VIKING20X is a successor of the Ocean General Circulation Model (OGCM) VIKING20 (Böning et al., 2016) and combines an eddy-rich ($1/20^\circ$) grid of the whole Atlantic (69°N - 33.5°S), two-way nested into a global $1/4^\circ$ ocean-sea-ice configuration (ORCA025), forced by the JRA55-do atmospheric forcing data set of the past decades (Tsuji et al., 2018). In this study we used the experiment VIKING20X.L46-KKG36107B. VIKING20X has 46 z-levels in the vertical, for which the layer thickness increases from 6 m at the surface to 250 m in the deep ocean. At the depths of the simulated regions (i.e. 628 m), the layer thickness is 94 m. The bathymetry in the model is based on the 2-minute ETOPO2 bathymetric database and represented by partial steps. Simulated velocities are stored as daily three-dimensional

averages (to allow releases on an everyday basis to assure equal integration lengths). The VIKING20 model itself has been intensively validated in previous studies (e.g. Breckenfelder et al., 2017) and used for dispersal studies (Baltazar-Soares et al., 2014; Breusing et al., 2016). To further demonstrate the reliability of the model simulations in our area of interest, we compared in situ temperature and salinity data derived from the 15 priorly mentioned full ocean depth CTD casts in June 2017, with modelled data from VIKING20X (respective data are shown in the **Supplementary Material** and discussed further in the **Results section 3.4**). Lagrangian simulations were performed on $\sim 90,000$ particles in total (1,000 particles per release region and release day), which were released from three release regions (covering four model grid boxes each) during the assumed natural spawning period in June (1st of June - 30th of June). The coordinates of the three virtual release regions were covering the actual sampling locations in 2017 at region A ('Canyon'), region B ('West-Bank') and region C ('East-Bank') (consider the **Supplementary Material** for further insights into the representation of regions in model topography as well as model velocities). The number of particles was determined based on the saturation plateau of computed rarefaction curves (respective data are shown in the **Supplementary Material** and discussed further in the **Results section 3.4**). In our study, release positions of sponge larvae were close to the seafloor (in detail a larval release was performed in the vertical centre of the last gridbox above the bottom). After release, virtual larvae were allowed to drift passively at any depth with the three-dimensional time-varying ocean velocities, and no assumptions about mortality were made at any time of the simulations. A typical pelagic larval duration of 14 days was assumed and dispersal probabilities were computed as described previously (van Sebille et al., 2018). To consider temporal fluctuations in current patterns, we modeled releases and dispersal over a period

of 10 years (between 2009 and 2018). In-depth analyses on seasonality were performed for one year, in which case we chose the year of the conducted expedition (i.e. 2017). For plotting larval trajectories and probabilities of dispersal, all particle positions deeper than 2000 m (determined by a depth tracer) were flagged particularly. Those particles which, on drifting day 14 (i.e. the last day of the typical pelagic larval duration phase) were found in areas where the seafloor lies below 2000 m, were considered as 'lost' to the open ocean. In this study we present '100 %-areas' and '95 %-areas' of larval drift. In the latter case, the area covered by the larvae drift refers to the area including 95 % of all observed particle positions within 14 days when positions are binned in boxes of 5 km x 5 km onto a geographic grid. Due to the chosen bin size the number of particle positions considered to calculate this area is, however, up to 0.6 % lower than 95 %. The connectivity matrix shows the probability that a particle is present at least once in one of the three regions during the last seven days of its drift relative to the number of all particles released. All oceanographic modelling and visualising was performed within Python (version 3.7.3) running (inside Jupyter notebooks; Kluyver et al., 2016) on an Unix system.

RESULTS

Sponge taxonomy

Phakellia hirondellei and the TaP clade individuals are morphologically very similar in their external appearance and can be difficult to distinguish even to the expert eye. Although similar, the spicule content of both lineages differed slightly: TaP clade individuals possessed two types of strongyles and two sizes of styles, and *P. hirondellei* contained one type of strongyles, two sizes of oxeas and one size of styles (**Supplementary Table S1**). Only two individuals within *P. hirondellei* showed small differences in the spicule content (lack of strongyles, **Supplementary Table S1**). The

molecular markers *18S* and *COI* were then used to investigate their phylogenetic placement and to confirm their taxonomic assignment in combination with traditional spicule analyses. In the *COI* analysis, the ten individuals identified based on the spicule content as *P. hirondellei* (Topsent, 1890; Bubarida order) were grouped in a monophyletic clade showing relatively moderate support (**Supplementary Material S1B**), with *Phakellia robusta* as sister clade (**Supplementary Material S1B**). However, these same individuals were grouped together in a clade with *P. robusta* based on the phylogenetic analysis of *18S* (**Supplementary Material S1A**) because of the low resolution power of this marker for sponge phylogeny at species level. Despite multiple sequencing attempts, generation of clean host sequences turned out to be difficult for the two sponge groups. Working with one single sequence direction for *18S*, we observed a few nucleotide mismatches within species, which could not be properly checked, but most likely do not reflect a true variation of the marker (as concluded based on the combination of the four different approaches used to decipher sponge taxonomy). The spicule analysis for all these ten individuals (**Supplementary Material S1C**) supported the assignment to *P. hirondellei*, despite small differences in the spicule content of two individuals. Within the same clade as *P. hirondellei* and *P. robusta*, we found other species which are assigned to either the order Axinellida or Bubarida (**Supplementary Material S1B**). The other 11 sponge individuals were classified as belonging to the TaP clade. *18S* sequencing suggested that this clade included *Topsentia* and *Petromica* species (**Supplementary Material S1A**). In the *COI* analysis, the robustly supported TaP clade appeared as sister to another clade containing *Petromica*, *Cymbastela*, and *Ciocalypta* (**Supplementary Material S1B**). The remaining 21 sponges could not be unambiguously taxonomically classified, which is why we discarded the respective samples from further analyses of connectivity.

Sponge microbial community composition

The microbiomes of TaP clade specimens and *P. hirondellei* clustered apart from seawater in a non-metric multidimensional scaling plot of weighted UniFrac distances (**Figure 2A**). In addition, microbiomes of TaP clade individuals and *P. hirondellei* fell into two distinct clusters based on weighted UniFrac distances. Alpha diversity indices were highest in seawater, intermediate in TaP clade specimens and

lowest in *P. hirondellei* (**Figure 2A**, small insert). Further, the microbial community composition was significantly different on ASV-level between TaP clade specimens and *P. hirondellei*, and also different from seawater (PERMANOVA, $p=0.001$, **Table 1**). Notably, the seawater reference microbiomes did not differ between the three sampling regions in terms of diversity and community composition (**Supplementary Material S2A-B**).

TABLE 1 | Overview of pairwise-comparisons (PERMANOVAs) based on weighted UniFrac matrixes of microbial communities between sponge clades (TaP clade and *P. hirondellei*) and sample types (TaP clade, *P. hirondellei*, seawater).

Group 1	Group 2	Sample size	Per-mutations	pseudo-F	p-value
TaP clade	<i>Phakellia hirondellei</i>	21	999	107.04	0.001
TaP clade	Seawater	20	999	26.56	0.001
<i>Phakellia hirondellei</i>	Seawater	19	999	94.13	0.001

In terms of bacterial abundance and composition, the microbiomes of the TaP clade specimens and *P. hirondellei* were characteristic of low microbial abundance (LMA) sponges (according to Hentschel et al., 2006; Gloeckner et al., 2014; Moitinho-Silva et al., 2017b) with dominant *Gammaproteobacteria* and *Bacteroidetes*

(**Figure 2B**). However, the relative abundances of these microbial phyla differed between TaP clade specimens and *P. hirondellei*. While the phylum *Nitrospirae* was significantly enriched in *P. hirondellei*, TaP clade individuals showed a significant enrichment of *Alphaproteobacteria* and *Bacteroidetes* (**Supplementary Material S3**).

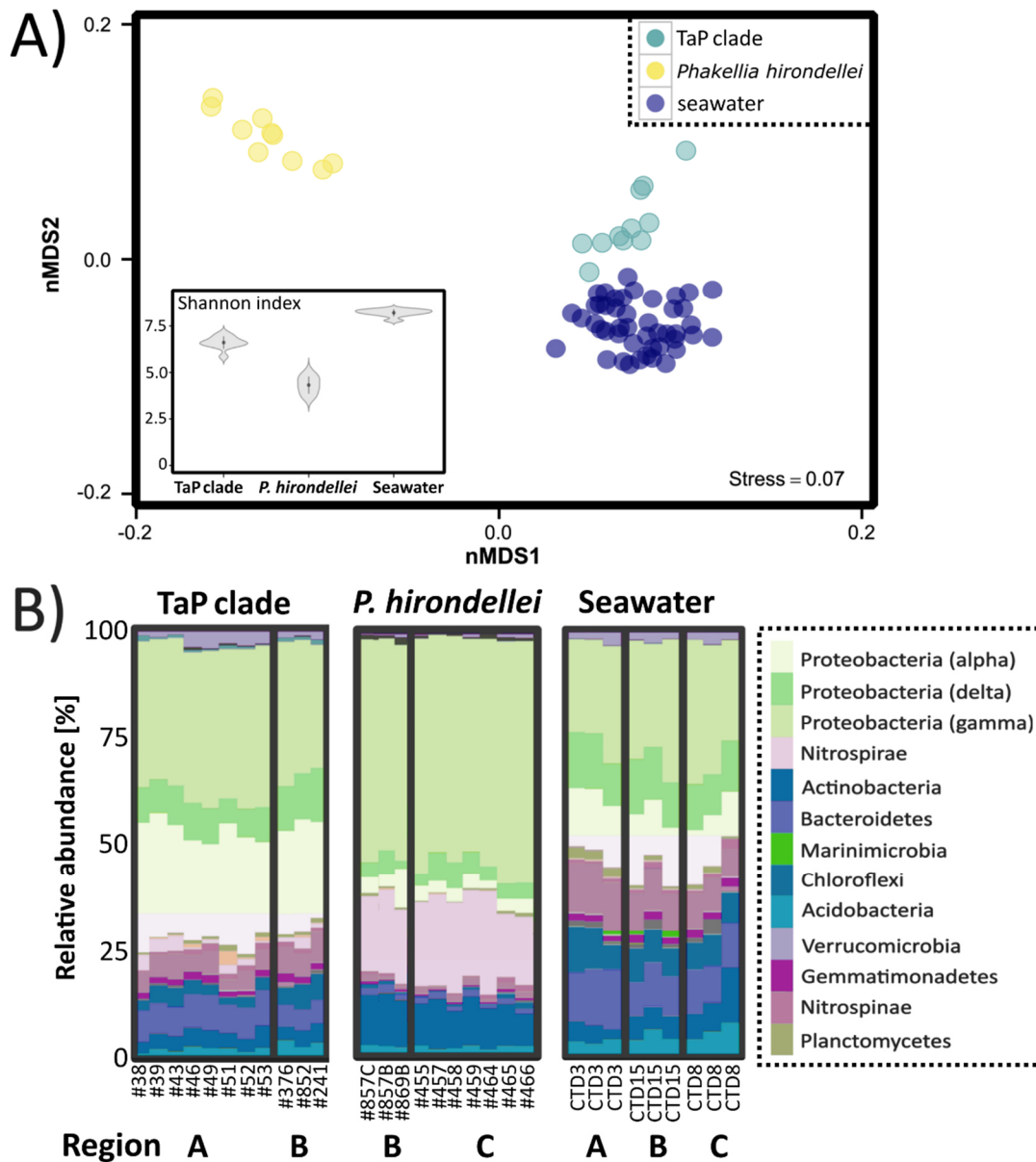


FIGURE 2 | Microbial community composition and richness of different sample types. **A)** Non-metric multidimensional scaling plot on weighted UniFrac distances. Each marker is one microbial community, with colors indicating sample type (the TaP clade, *P. hironellei* and seawater). Data of 21 sponge individuals and seawater samples from 15 stations are presented. Small plot inside **Fig 3A** presents Shannon indices of TaP clade individuals, *P. hironellei* and seawater. Dots show means and whiskers represent standard deviations. In **Fig.3B** the microbial community composition is shown on the phylum level across all three sampling sites (A: 'Canyon', B: 'West-Bank', C: 'East-Bank') for the TaP clade, *P. hironellei* and seawater.

In a nMDS approach, we observed a substructuring of TaP clade microbiomes between region A ('Canyon') and region B ('West-Bank') (**Figure 3A**), which was statistically significant (PERMANOVA, $p=0.011$, **Table 2**). *Phakellia hironellei* microbiomes showed no significant spatial differences

between region B ('West-Bank') and region C ('East-Bank') (**Figure 3B**). In terms of alpha diversity indices (Shannon index) the TaP clade and *P. hironellei* did not show any spatial differences between sampling region (**Figures 3A-B**, small inserts).

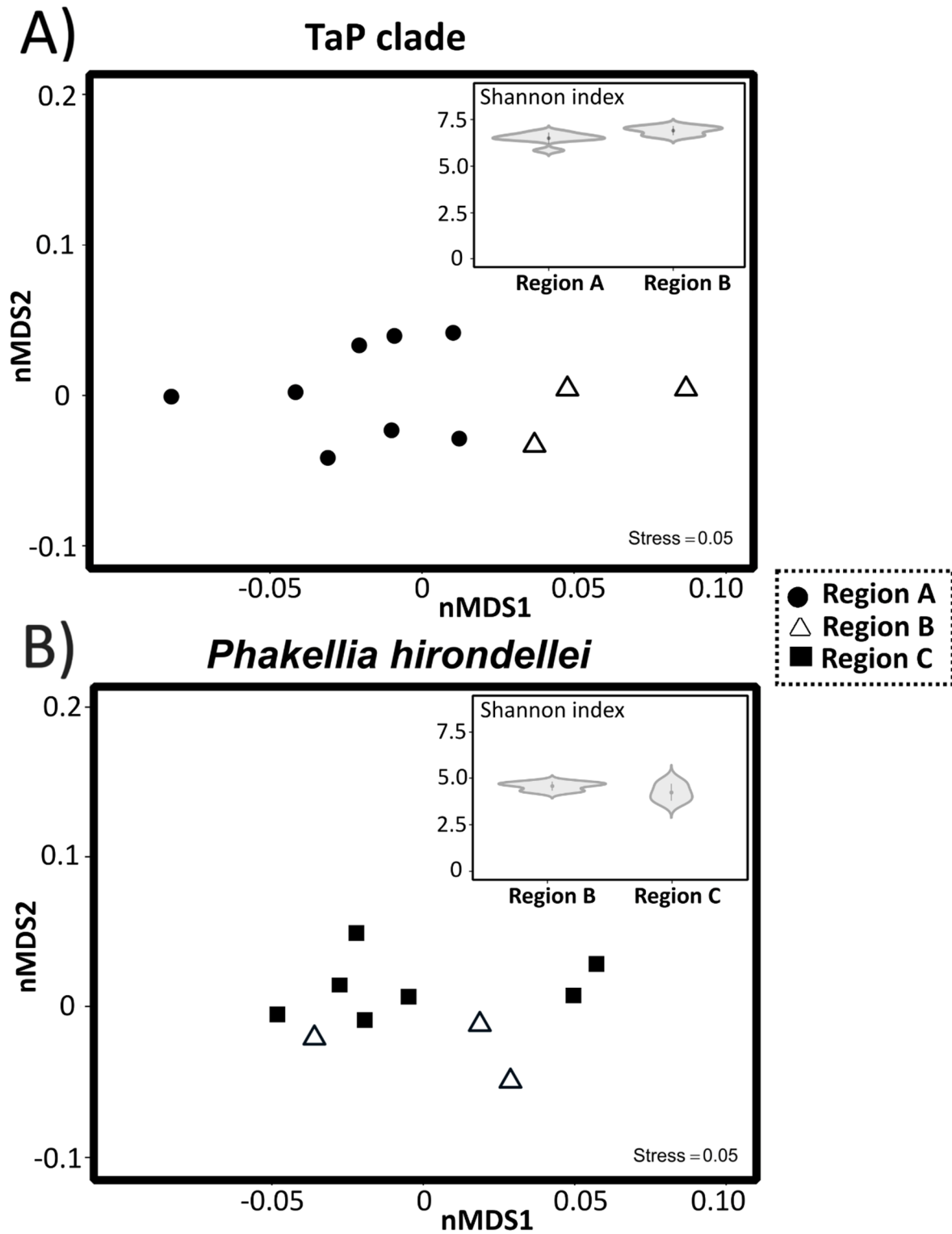


FIGURE 3 | Microbial community composition of the TaP clade and *P. hirondellei* across the three sampling regions (A: 'Canyon', B: 'West-Bank', C: 'East-Bank'). Non-metric multidimensional scaling (**Figs 3 A-B**) is based on weighted UniFrac distances (ASV-level) and symbols indicate the respective sampling regions. The violin plots presented within **Figs. 3A-B** show mean (dot) and standard deviation (whiskers) of Shannon indices for the two sponge clades across sites.

TABLE 2 | Overview of pairwise-comparisons (PERMANOVAs) based on weighted UniFrac matrixes of microbial communities between sampling regions (A:‘Canyon’, B:‘West-Bank’, C:‘East-Bank’) for the TaP clade and *P. hironellei*.

Group 1	Group 2	Sample size	Per-mutations	pseudo-F	p-value
TaP clade (RegionA)	TaP clade (RegionB)	11	999	3.59	0.011
<i>Phakellia hironellei</i> (RegionB)	<i>Phakellia hironellei</i> (RegionC)	10	999	1.61	0.144

Sponge population genetics

For population structure and connectivity (SNPs) analyses a total of 3,408 and 2,119 SNPs were obtained for the *P. hironellei* and TaP clade datasets, respectively. Population genetic statistics for the two sponge clades are given in **Table 3**. Overall, the expected heterozygosity (He), generally considered as a measure of genetic diversity, was slightly lower for TaP clade specimens (0.369) than for *P. hironellei* (0.400). Within sampling regions, He ranged from 0.366 (A:‘Canyon’) to 0.381 (B:‘West-Bank’) for TaP clade specimens, and from 0.394 (C:‘East-Bank’) to 0.408 (B:‘West-

Bank’) for *P. hironellei* (**Table 3**). Observed heterozygosity (HO) was again lower for the TaP clade (0.535) than for *P. hironellei* (0.601). Within sampling regions, HO ranged from 0.532 (A:‘Canyon’) to 0.557 (B:‘West-Bank’) for TaP clade specimens, and from 0.597 (B:‘West-Bank’) to 0.603 (C:‘East-Bank’) for *P. hironellei* (**Table 3**). All values for the inbreeding coefficient (FIS) were negative and significant for the two sponge clades and the different regions analyzed, indicating an excess of observed heterozygotes (**Table 3**).

TABLE 3 | Population genetic statistics for the TaP clade and *P. hironellei*, grouping samples per sampling region. The abbreviations are as follows: n (number of samples), Ho (‘observed heterozygosity’), He (‘expected heterozygosity’) (= genetic diversity), Fis (‘inbreeding coefficient’). Significant values are depicted in bold.

Sponge clade / Sampling region	n	Ho	He	F _{IS}
TaP clade				
A: ‘Canyon’	8	0.532	0.366	-0.454
B: ‘West-Bank’	3	0.557	0.381	-0.463
Total	11	0.535	0.369	-0.452
<i>Phakellia hironellei</i>				
B: ‘West-Bank’	3	0.597	0.408	-0.465
C: ‘East-Bank’	7	0.603	0.394	-0.531
Total	10	0.601	0.400	-0.504

The results of the STRUCTURE and adegenet analyses revealed no significant genetic structure among the samples for the two

sponge clades from the different regions (**Figure 4A-B**). This was also corroborated by the FST values for the pairwise comparisons

between regions, which was 0.023 for the comparison between TaP clade regions (p-value = 0.06) and 0.015 for the comparison between *P. hirondellei* regions (p-value = 0.169). Interestingly, the DAPC analysis for which samples were grouped per sampling region, did detect differences across sampling regions for the TaP clade, which showed a

subtle separation between regions A ('Canyon') and B ('West-Bank'), (Figure 4C; note that DAPC analyses are generally more sensitive than STRUCTURE and adegenet analyses). In contrast, for *P. hirondellei*, we did not observe this differentiation, but samples of the canyon region (region A) were not available for this species (Figure 4D).

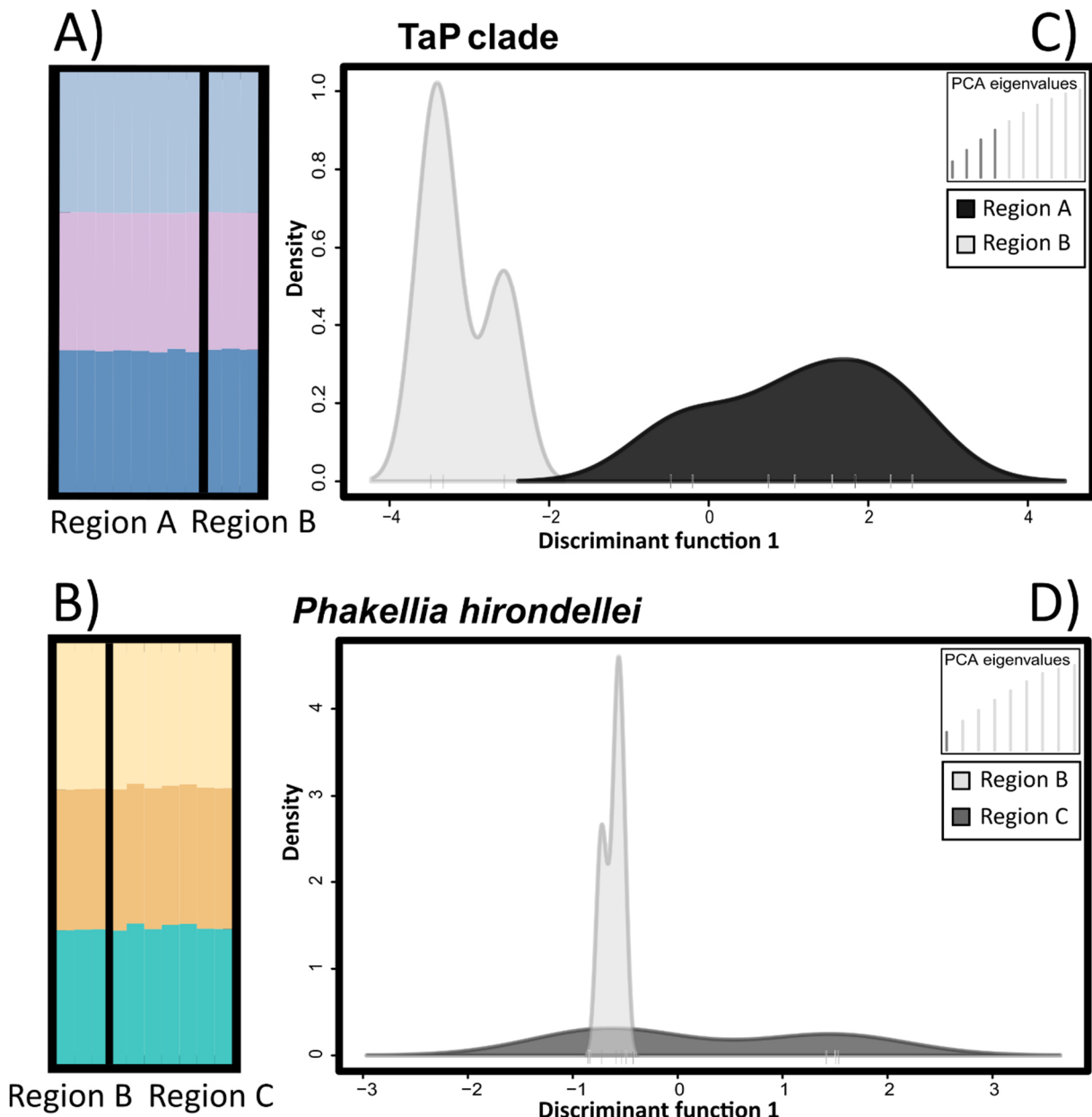


FIGURE 4 | Sponge population genetics based on individual genotype assignments of TaP clade and *P. hirondellei* individuals to clusters (K) as inferred by STRUCTURE (Fig. 4A-B). Spatial structure of the two sponge clades across the three sampling regions (A: 'Canyon', B: 'West-Bank', C: 'East-Bank'), as inferred by DAPC (Figs. 4C-D).

Validation of oceanographic modelling setup

In order to link population genetic data with oceanographic connectivity, VIKING20X was chosen as our model to simulate drifting of sponge larvae. The optimal number of particles to be released was determined as 1000 particles per grid box and day based on the saturation plateau of computed rarefaction curves (**Supplementary Material S4**). When determining the optimal model starting date, we observed that current direction and velocities seemed to be governed by mesoscale processes over the whole year (**Supplementary Material S5**). We chose June as the starting month to release virtual sponge larvae into the water column; and due to strong temporal variability we performed releases over 10 different years. The month of June was chosen because it is two months after the main phytoplankton bloom in 2017 (**Supplementary Material S6**) and it is during a period when ripe reproductive stages were found in both, TaP clade and *P. hirondellei* individuals. In this month, the in situ chlorophyll a concentrations were comparably low at all three locations (**Supplementary Material S7A-B**) and corresponded well to the satellite data. Both, TaP clade members and *P. hirondellei* were found to contain mature gametes in mid-June, when they were collected. In TaP clade individuals, vitellogenic oocytes of ~ 40-60 µm were found close to the canals, indicating that they were relatively mature and almost ready to be released in the water column (**Supplementary Material S8A**). The oocytes were full of nutrients of proteinaceous origin (**Supplementary Material S8B-C**) and heterogenous yolk, a mix of lipid and protein (**Supplementary Material S8B, D**). In *P. hirondellei*, similar vitellogenic oocytes were observed (~ 50 µm), probably mature and ready to be released in the water column (**Supplementary Material S8E**). In contrast to

TaP clade individuals, in *P. hirondellei* the nutrient reserves within the oocytes seemed to be homogeneous yolk comprised only by protein platelets (≤ 500 nm) (**Supplementary Material S8E-F**).

The three release and sampling regions (A ('Canyon'), B ('West-Bank'), C ('East-Bank')), lay within a similar depth range (**Supplementary Material S7**) and sampling at all three regions occurred within the same watermass (**Supplementary Material S9A**). In situ temperature and salinity measurements fitted overall well to the data extracted from VIKING20X in terms of their vertical profile (**Supplementary Material S9A**) and their rough characteristics at the bottom depths of the CTD sampling sites (**Supplementary Material S9B**). However, we noted a slight shift to lower salinities in the VIKING20X data in comparison to the in situ measurements. On a different note, VIKING20X does not resolve hydrodynamics inside canyons and further does not simulate important small-scale properties (such as tides), which could lead to an overestimation of connectivity between the individual regions. In our experiments we released virtual larvae at the 'opening' of the canyon (but at the depth level that corresponds to the in situ sampling depths). For our simulations we considered particles drifting on day 14 in regions where the seafloor is below 2,000 m water depth, as 'lost' to the open ocean. This configuration rests upon the evidence-based assumption that fan-shaped sponge abundance in the study area is very low below 2,000 m water depth: With the aid of a photogrammetric sledge (ROTV Politolana), fan-shaped sponges at the study area were recorded particularly within the mesopelagic zone (Sánchez et al., 2017), especially in a depth range between 425–1300 m, with highest densities between 450–550 m.

Larvae drift model

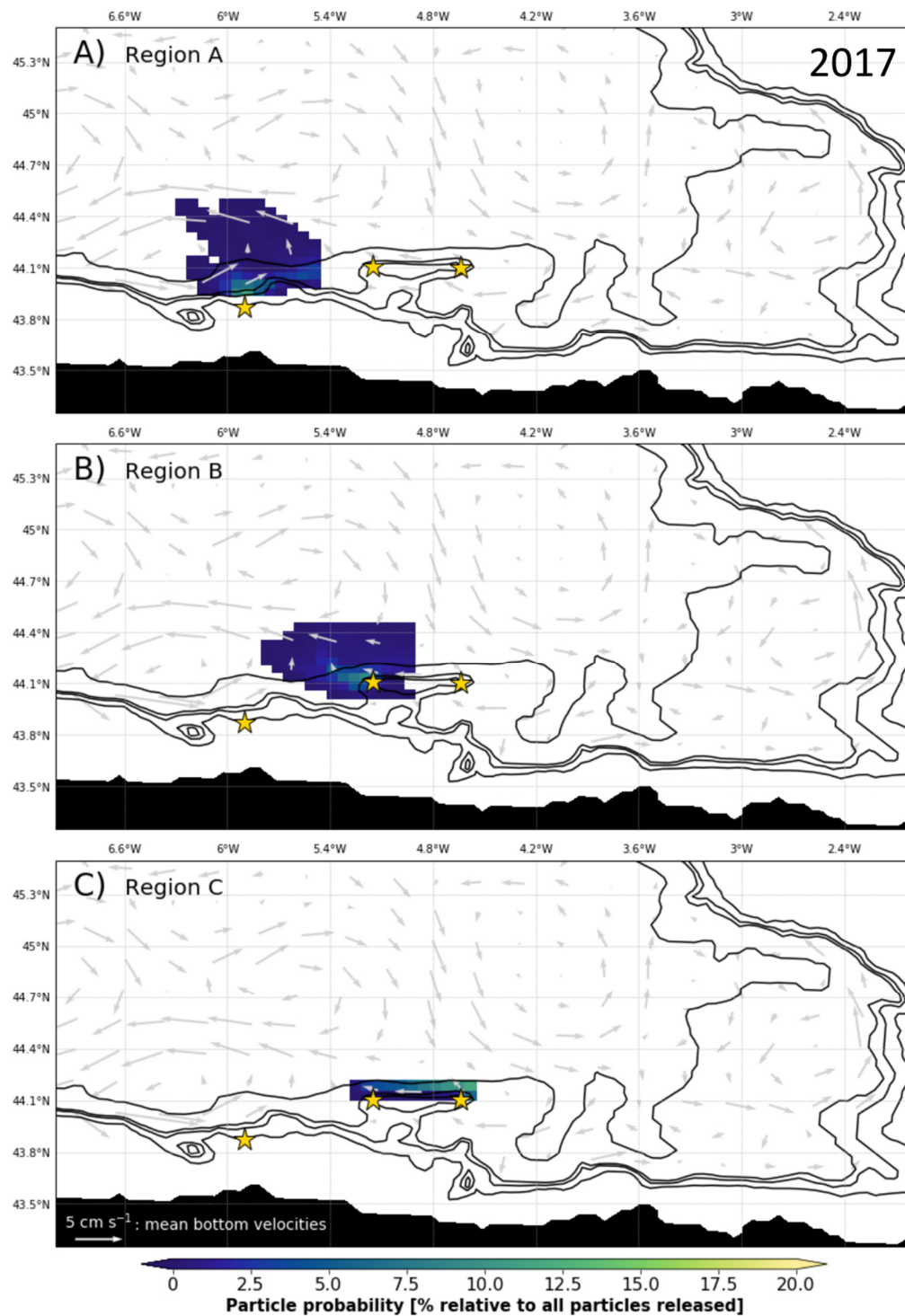


FIGURE 5 | Output of the Lagrangian simulations showing particle probabilities in space (in % relative to all particles released) after a 14 days drifting time and a daily release of larvae over a period of 30 days in June 2017 (which was determined as suitable model start point during model validation procedure) for the three regions (A: 'Canyon', B: 'West-Bank', C: 'East-Bank'). Bathymetry contours are shown by black lines and the Spanish coastline is indicated by the black area. Mean monthly current velocities for the analysed period are indicated by grey arrows.

For all three regions a considerable fraction of the released particles stayed within their release region for all analysed years and inter-regional connectivity was low. Passive-particle drift trajectories showed a preferential retention of simulated larvae within the study area (for a particular year see **Figure 5** and **Supplementary Material S10**; for multi-year data see **Supplementary Material S11**). Most virtual particles travelled along the bathymetry and were strongly influenced by the mesoscale variability, i.e. the existence and direction of ocean eddies. Across all years and sampling regions on drifting day 14, on average $44.2 \pm SE 5.7\%$ of all particles were found in areas where the seafloor is located below 2,000 m, but 0 % of the particles were actually drifting at water depths below 2,000 m. We observed strong

inter-annual variations (between 2009 and 2018) in the geographic location of eddies (**Supplementary Material S12**), as well as in the average larval drift distances, and in the 100 %-areas and 95 %- areas of larval drift (**Supplementary Material S13**). These inter-annual variations were observed at all three sampling regions. To assess inter-annual patterns of connectivity, we computed a connectivity matrix (**Figure 6**). Throughout the ten analysed years region A ('Canyon') and region C ('East-Bank') were never connected. In less than half of the analysed years, region B ('West-Bank') received particles from the other two regions. Only in two years (2009 and 2011) particles from region B were transported to another region (only region C ('East-Bank')).

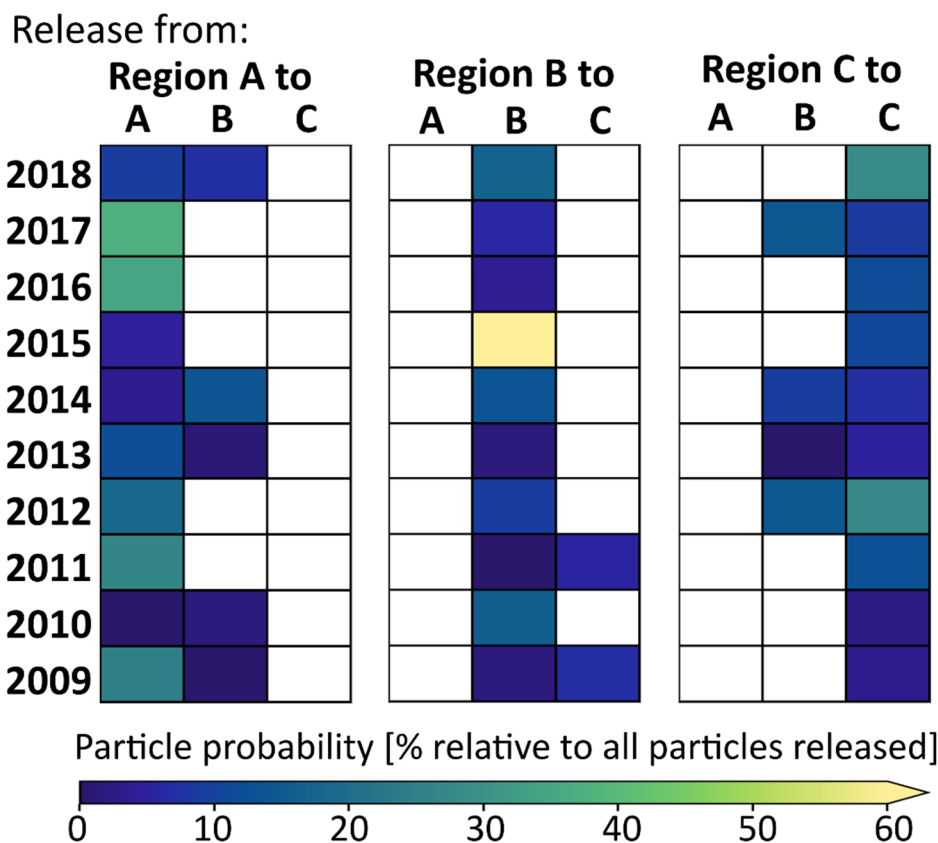


FIGURE 6 | Connectivity matrix showing the probability that a particle is present at least once in one of the three regions (A: 'Canyon', B: 'West-Bank', C: 'East-Bank') during the last seven days of its drift relative to the number of all particles released. The inter-annual variability across the ten analysed years (2009-2018) is illustrated.

DISCUSSION

Almost two-thirds of the Earth are considered deep-sea, which is also the least explored habitat on Earth (Costello et al., 2010; Ramirez-Llodra et al., 2010; Ramirez-Llodra et al., 2011). Designing population genetic studies in such understudied regions can be challenging. In the present study half of all originally collected sponge specimens could not be used for population genetic evaluation, because the taxonomic assignment remained ambiguous despite considerable efforts using both molecular (*COI*, *18S*, *16S*) and classical taxonomic markers (spicules). Further, even for specimens identified as belonging to the same species, it can be difficult to assign a taxonomic name. Indeed, sponge systematics and phylogenic reconstructions continue to be difficult using traditional spicule analyses and/or molecular techniques (Erpenbeck et al., 2006; Philippe et al., 2011). While in our study individuals belonging to the TaP clade were identified as belonging to the same species based on *COI*, *18S*, and spicule analyses, their assignment to a described order remains wanting. In order to be able to provide an official taxonomic name, a complete review of the orders to which *Topsentia* and *Petromica* (and other groups) belong will need to be accomplished.

Although the encountered difficulties in taxonomic identification led to a considerable reduction of our sampled dataset, we emphasize that a critical re-evaluation of sponge taxonomy needs to be at the basis of every population genetic analysis, as morphologically very similar specimens (i.e. specimens with a similar external body shape) may belong to different sponge orders. The microbial community composition based on *16S* amplicon data was identified as a helpful tool to support sponge taxonomy since the microbiomes were host-specific. Indeed, for many organisms phylogenetic relatedness of the host can strongly affect the associated microbial community composition (reviewed in Moran et al., 2019). For sponges, species-

specific prokaryotic communities are commonly found (e.g. Easson and Thacker, 2014; Thomas et al., 2016; Steinert et al., 2017). Besides host-specific microbiomes, both, TaP clade individuals and *P. hirondellei* showed microbial signatures typical for Low Microbial Abundance (LMA) sponges, as described in (Moitinho-Silva et al., 2017b), which differed in terms of relative abundances of microbial phyla. Results of our study further imply that on a subspecies level (genotypes obtained using SNPs), the sponge microbiome seems to be a sensitive parameter for evaluating emerging taxonomic structures of sponge holobionts, and particularly so when the ASV – level is considered for microbial analyses.

Choosing the appropriate spatial scale to identify potential genetic structures of deep-sea sponges can be difficult since knowledge on population genetics/genomics of deep-sea organisms is sparse. In order to cover a maximum geographic range, the population genetic patterns of the fan-shaped sponges *P. robusta* and *P. hirondellei* are examined over a few thousand kilometers in a complementary paper by Taboada et al. (in review), while the present study focussed on distances of < 100 km. For deep-sea organisms, generally low genetic structuring is reported when organisms occur at similar depths (McClain and Hardy, 2010; Taylor and Roterman, 2017), while other studies highlight the importance of biophysical barriers with respect to inhibition of gene flow. A recent meta-analysis on meiofaunal biodiversity covering 21 canyons together with reference stations on the adjacent slopes, observed a high dissimilarity in taxonomic composition between canyons and slopes (Bianchelli and Danovaro, 2019). At a population genetic level, the red gorgonian *Paramuricea clavata* showed a strong genetic differentiation among populations in different submarine canyons in the Mediterranean Sea, presumably due to limited effective dispersal (Pérez-Portela et al., 2016). Similarly, another recent study found genetic differentiation among cold-water coral populations across

sites in the near neighbourhood (Boavida et al., 2019).

In the present study, we mostly observed genetically well-connected sponge populations for both, the TaP clade, and *P. hironellei*. This result was consistent across the STRUCTURE analysis (Figure 4 A-B), adegenet analysis, and F_{st} values. Interestingly, individuals belonging to the TaP clade showed a subtle separation between sampling regions A ('Canyon') and B ('West-Bank') in the DAPC analysis (Figure 4 C), despite the analysis detected an overall single panmictic population. Similarly, we observed a significant difference of TaP clade microbiomes between sampling regions A ('Canyon') and B ('West-Bank') in terms of weighted UniFrac distances. Seawater microbiomes did however not differ between the sampling regions. This supports our oceanographic in situ observations (TS data) that sampling at all three sampling regions occurred within the same watermass and under similar environmental conditions. We thus conclude that slight but distinct substructure in host population genetics and microbial community composition occur concomitantly. We further conclude that although gene flow generally occurs between both examined geologic features (the Avilés Canyon System and the Le Danois Bank), the exchange may be low enough to promote subtle genetic differentiation. It is quite remarkable that slight differentiation can even be found on scales below 100 km distance, and interestingly on both, the host and microbiome level. For *P. hironellei*, we did not observe a regional differentiation, but samples of the canyon region (region A) were not available for this species. Our results are overall consistent with a previous lack in strong genetic structuring in deep-sea organisms occurring at similar depths (McClain and Hardy, 2010; Taylor and Roterman, 2017).

Our performed simulations of Lagrangian connectivity suggest that virtual sponge larvae stay in the study area and ocean

currents generally allow local gene flow events, despite inter-annual variations in the connectivity between the three regions. For sessile invertebrates, 10–100 km distance is the typical range within which the majority of off-spring is transported away from its parents (Palumbi, 2004). Particle trajectories calculated within our study reveal a similar range for TaP clade individuals and *P. hironellei*. Larval drifting occurred in swirled patterns and exhibited a variable strength across different years, as the major circulation patterns seemed to be governed by mesoscale processes. Overall, ocean currents seem to promote virtual larvae to drift along the bathymetry in the study area. Our results thus suggest that, although the continental shelf is comparably narrow around the northern shores of Spain, ocean circulation may prevent the loss of sponge larvae towards the open ocean. Whether this will still be the case under future climate scenarios remains to be tested in future studies. Interestingly, in a recent study by Fox et al., (2016) trajectories of cold-water coral larvae were found to be strongly correlated to the North Atlantic Oscillation. The authors of that study conclude that the connectivity of the existing MPA network is vulnerable to atmospheric-driven changes in ocean circulation.

Besides the need of further studies on the variability of the physical settings, also further biological data is urgently needed to improve understanding of connectivity in the deep-sea, especially for key organisms like sponges and the ecosystems they form. As current understanding of deep-sea larval behavior is extremely limited (in general) to non-existent (for deep-sea sponges), our particle trajectory calculations are based on the assumption of passive dispersal only. However, recent studies have shown that active larval movements can have an order of magnitude impact on dispersal (Gary et al., 2020). It will therefore be crucial for future studies to gather more information on the ecology and behaviour of deep-sea sponge larvae.

In summary, we show that canyons can promote slight genetic divergence of sponges and also of their microbiomes, although more data would be needed to provide further statistical support for this observation. We further conclude that the specific prevailing ocean currents of the Cantabrian Sea dictate sponge host genetics by promoting a settlement in suitable areas within kilometers from their parent sponges and consequently allowing gene flow. However oceanographic connectivity showed a distinct variability between the simulated years. While genetic uniformity would occur in times of connectivity between the Avilés Canyon and the Le Danois Bank, genetic differentiation might happen in years of disconnection, when the deep-water currents promote drifting of larvae in different directions. These findings contribute to a better understanding of connectivity, genetic and microbial variability of deep-sea organisms. Our results imply that a design of ecologically coherent networks of MPAs is no trivial task for deep-sea sponge holobionts. We conclude for our study site (on scales < 100 km) that (i) physical connectivity between close-by sites via ocean currents has a strong temporal component, and that (ii) differences between close-by sites are stronger reflected in the sponge-associated microbial community than in the concomitant sponge-host genetics. These points reveal that it is important for a design of ecologically coherent networks of MPAs, to consider the naturally existing variability of connectivity (e.g. physical data from multiple years) and to consider the occurring microbial diversity in addition to the macrobial diversity.

REFERENCES

- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. (1990). Basic local alignment search tool. *J Mol Biol* 215: 403-410.
- Andrello, M., Guilhaumon, F., Albouy, A., Parravicini, V., Scholtens, J., Verley, P., Barange, M., et al. (2017). Global mismatch between fishing dependency and larval supply from marine reserves. *Nat. Commun.* 8: 16039.
- Baltazar-Soares, M., Biastoch, A., Harrod, C., Hanel, R., Marohn, L., Prigge, E., et al. (2014). Recruitment collapse and population structure of the european eel shaped by local ocean current dynamics. *Curr Biol* 24: 104–108.

ACKNOWLEDGEMENTS

We dedicate this study to Hans Tore Rapp, friend, mentor and esteemed colleague whose outstanding expertise in deep-sea sponge taxonomy will be dearly missed. We greatly acknowledge the crew and scientific party of RV Ángeles Alvariño cruise *SponGES0617* for their valuable support at sea. As well as members of the IEO Gijón for logistical support during a guest stay of KB before the cruise. We appreciate Andrea Hethke's and Ina Clefsen's, as well as Thomas Hansen's assistance in the laboratory after the cruise while generating the microbial amplicon and flow cytometry data. We thank Willi Rath for support in technical issues with the modelling part, and Lara Schmittmann and Ina Clefsen for technical support with *18S* and *COI* sequencing. Paco Cárdenas provided valuable input with regard to the sponge taxonomic identification and on a first draft of the manuscript, which was very much appreciated. We further thank two anonymous reviewers for their valuable input during the review process.

AUTHOR CONTRIBUTIONS

UH, AR, AB, KB designed the study. KB, AR, ST, VK, PR, JC participated in sampling. PR, JC, AR conducted sponge taxonomic analysis. ST, AR performed sponge population genetic analyses, while KB, UH performed microbial analyses. CS, KB, AB designed and performed Lagrangian modelling. KG provided VIKING20X data. VK and KB did microscopy of sponge tissue. KB performed assessment of environmental conditions in situ and remotely. AF was involved in sequencing of microbial samples. KB and UH wrote the manuscript, and the other authors reviewed and edited it.

- Beazley, L., Kenchington, E., Murillo, F.J., del Mar Sacau, M. (2013). Deep-sea sponge grounds enhance diversity and abundance of epibenthic megafauna in the Northwest Atlantic. *ICES Journal of Marine Science* 70: 1471-1490.
- Beugin, M.P., Gayet, T., Pontier, D., Devillard, S., and Jombart, T. (2018). A fast likelihood solution to the genetic clustering problem. *Methods Ecol Evol* 9: 1006–1016.
- Bianchelli, S. and Danovaro, R. (2019). Meiofaunal biodiversity in submarine canyons of the Mediterranean Sea: A meta-analysis. *Prog Oceanogr* 170: 69–80.

- Boavida, J., Becheler, R., Choquet, M., Frank, N., Taviani, M., Bourillet, J.-F., Meistertzheim, A.-L., Grehan, A., Savini, A., Arnaud-Haond, S. (2019). Out of the Mediterranean? Post-glacial colonization pathways varied among cold-water coral species. *Journal of Biogeography* 46: 915–931.
- Bolyen, E., Rideout, J.R., Dillon, M.R. et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 37: 852–857.
- Böning, C.W., Behrens, E., Biastoch, A., Getzlaff, K., and Bamber, J.L. (2016). Emerging impact of Greenland meltwater on deep-water formation in the North Atlantic Ocean. *Nat Geosci* 9: 523–527.
- Breckenfelder, T., Rhein, M., Roessler, A., Böning, C.W., Biastoch, A., Behrens, E., and Mertens, C. (2017). Flow paths and variability of the North Atlantic Current: A comparison of observations and a high-resolution model. *J Geophys Res Ocean* 122: 2686–2708.
- Breusing, C., Biastoch, A., Drews, A., Metaxas, A., Jollivet, D., Vrijenhoek, R.C., et al. (2016). Biophysical and population genetic models predict the presence of “phantom” stepping stones connecting mid-Atlantic ridge vent ecosystems. *Curr Biol* 26: 2257–2267.
- Busch, K., Hanz, U., Mienis, F., Mueller, B., Franke, A., Roberts, E.M., Rapp, H.T., Hentschel, U. (2020a). On giant shoulders: how a seamount affects the microbial community composition of seawater and sponges. *Biogeosciences* 17: 3471–3486.
- Busch, K., Beazley, L., Kenchington, E., Whoriskey, F., Slaby, B.M., Hentschel, U. (2020b). Microbial diversity of the glass sponge *Vazella pourtalesii* in response to anthropogenic activities. *Conservation Genetics*, doi:10.1007/s10592-020-01305-2.
- Busch, K., Wurz, E., Rapp, H.T., Bayer, K., Franke, A., Hentschel, U. (2020c). *Chloroflexi* dominate the deep-sea golf ball sponges *Craniella zetlandica* and *Craniella infrequens* throughout different life stages. *Front. Mar. Sci.* 7: 674.
- Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Knight, R., Lozupone, C.A., et al. (2011). Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc Natl Acad Sci* 108: 4516–4522.
- Catchen, J., Hohenlohe, P.A., Bassham, S., Amores, A., and Cresko, W.A. (2013). Stacks: An analysis tool set for population genomics. *Mol Ecol* 22: 3124–3140.
- Cathalot, C., van Oevelen, D., Cox, T. J. S., Kutti, T., Lavaleye, M., Duineveld, G., et al. (2015). Cold-water coral reefs and adjacent sponge grounds: hot spots of benthic respiration and organic carbon cycling in the deep sea. *Front. Mar. Sci.* 2: 37.
- Combosch, D.J., Lemer, S., Ward, P.D., Landman, N.H., and Giribet, G. (2017). Genomic signatures of evolution in Nautilus-An endangered living fossil. *Mol Ecol* 26: 5923–5938.
- Costello, M.J., Cheung, A., and De Hauwere, N. (2010). Surface area and the seabed area, volume, depth, slope, and topographic variation for the world’s seas, oceans, and countries. *Environ Sci Technol* 44: 8821–8828.
- Cowen, R.K. and Sponaugle, S. (2009). Larval dispersal and marine population connectivity. *Ann Rev Mar Sci* 1: 443–466.
- Darriba, D., Taboada, G.L., Doallo, R. and Posada, D. (2012). jModelTest 2: more models, new heuristics and parallel computing. *Nature methods* 9(8):772–772.
- Delandmeter, P. and Van Sebille, E. (2019). The Parcels v2.0 Lagrangian framework: New field interpolation schemes. *Geosci Model Dev* 12: 3571–3584.
- Earl, D.A. and vonHoldt, B.M. (2012). STRUCTURE HARVESTER: A website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4: 359–361.
- Easson, C.G. and Thacker, R.W. (2014). Phylogenetic signal in the community structure of host-specific microbiomes of tropical marine sponges. *Front Microbiol* 5: 532.
- Erpenbeck, D., Breeuwer, J.A.J., Parra-Velandia, F.J., and Van Soest, R.W.M. (2006). Speculation with spiculation? - Three independent gene fragments and biochemical characters versus morphology in demosponge higher classification. *Mol Phylogenet Evol* 38: 293–305.
- De Forges, B.R., Koslow, J.A., and Poore, G.C.B. (2000). Diversity and endemism of the benthic seamount fauna. *Nature* 405: 26–29.
- Díez-Vives, C., Taboada, S., Leiva, C., Busch, K., Hentschel, U., Riesgo, A. (2020). On the way to specificity - Microbiome reflects sponge genetic cluster primarily in highly structured populations. *Molecular Ecology*. doi:10.1111/mec.15635.
- Easson, C. G., Chaves-Fonnegra, A., Thacker, R. W., Lopez, J. V. (2020). Host population genetics and biogeography structure the microbiome of the sponge *Cliona delitrix*. *Ecology and Evolution* 205(22):3505–14.
- FAO (2009). The FAO international guidelines for the management of deep-sea fisheries in the High Seas. *FAO Fisheries and Aquaculture Department*, Rome.
- FAO (2016). Vulnerable Marine Ecosystems: Processes and practices in the High Seas. In: Thompson, A., Sanders, J., Tandstad, M., Carocci, F., Fuller, S. (eds), *FAO Fisheries and Aquaculture Technical Paper* 595, Rome.
- Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenhoek, R. (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* 3(5): 294–299.
- Fox, A.D., Henry, L.-A., Corne, D.W., Roberts, J.M. (2016). Sensitivity of marine protected area network connectivity to atmospheric variability. *Royal Society Open Science* 3:160494.
- Gallego, A., Gibb, F.M., Tullet, D., Wright, P.J. (2017). Bio-physical connectivity patterns of benthic marine species used in the designation of Scottish nature conservation marine protected areas. *ICES J. Mar. Sci.* 74:1797–1811.
- Gary S.F., Fox A.D., Biastoch A., Roberts J.M., Cunningham S.A. (2020). Larval behaviour, dispersal and population connectivity in the deep sea. *Scientific Reports* 10:10675.
- Gloeckner, V., Wehrl, M., Moitinho-Silva, L., Schupp, P., Pawlik, J.R., Lindquist, N.L., et al. (2014). The HMA-LMA dichotomy revisited: an electron microscopical survey of 56 sponge species. *Biol Bull* 227: 78–88.
- Gómez-Ballesteros, M., Druet, M., Muñoz, A., Arrese, B., Rivera, J., Sánchez, F., et al. (2014). Geomorphology of the Avilés Canyon System, Cantabrian Sea (Bay of Biscay). *Deep Res Part II Top Stud Oceanogr* 106: 99–117.
- Griffiths, S. M., Antwis, R. E., Lenzi, L., Lucaci, A., Behringer, D. C., Butler IV, M. J., Preziosi, R. F. (2019). Host genetics and geography influence microbiome composition in the sponge *Ircinia campana*. *Journal of Animal Ecology* 88(11):1684–1695.
- Harrington, B. and the Inkscape Development Team (2005). Inkscape. Available online at: <http://www.inkscape.org/>.
- Hentschel, U., Usher, K.M., and Taylor, M.W. (2006). Marine sponges as microbial fermenters. *FEMS Microbiol Ecol* 55: 167–177.
- Hentschel, U., Piel, J., Degnan, S.M., Taylor, M.W. (2012). Genomic insights into the marine sponge microbiome. *Nat Rev Microbiol* 10: 641–654.
- Hickey, B., Baker, E., and Kachel, N. (1986). Suspended particle movement in and around Quinault submarine canyon. *Mar Geol* 71: 35–83.
- Highfield, J.M., Eloire, D., Conway, D. V. P., Lindeque, P. K., Attrill, M. J., Somerfield, P. J. (2010). Seasonal dynamics of meroplankton assemblages at station L4. *Journal of Plankton Research* 32(5):681–691.
- Indraningrat, A. A. G., Micheller, S., Runderkamp, M., Sauerland, I., Becking, L. E., Smidt, H., Sipkema, D. (2019). Cultivation of sponge-associated bacteria from *Agelas sventres* and *Xestospongia muta* collected from different depths. *Mar. Drugs* 17:1–17.
- Jakobsson, M. and Rosenberg, N.A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23: 1801–1806.
- Jeffries, D.L., Copp, G.H., Handley, L.L., Håkan Olsén, K., Sayer, C.D., and Hänfling, B. (2016). Comparing RADseq and microsatellites to infer complex phylogeographic patterns, an empirical perspective in the Crucian carp, *Carassius carassius*, L. *Mol Ecol* 25: 2997–3018.
- Jombart, T. (2008). ADEGENET: A R package for the multivariate analysis of genetic markers. *Bioinformatics* 24: 1403–1405.
- Jombart, T. and Ahmed, I. (2011). adegenet 1.3-1: New tools for the analysis of genome-wide SNP data. *Bioinformatics* 27: 3070–3071.

- Jombart, T., Devillard, S., and Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* 11: 94.
- Katoh, K. and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol Biol Evol* 30: 772–780.
- Kazanidis, G., Vad, J., Henry, L.-A., Neat, F., Berx, B., Georgoulas, K., Roberts, J.M. (2019). Distribution of deep-sea sponge aggregations in an area of multisectoral activities and changing oceanic conditions. *Front. Mar. Sci.* 6: 163.
- Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., et al. (2012). Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28: 1647–1649.
- Kennington, E., Wang, Z., Lirette, C., Murillo, F.J., Guijarro, J., Yashayaev, I., Maldonado, M. (2019). Connectivity modelling of areas closed to protect vulnerable marine ecosystems in the northwest Atlantic. *Deep-Sea Res. Part I* 143: 85–103.
- Klitgaard, A.B. and Tendal, O.S. (2004). Distribution and species composition of mass occurrences of large-sized sponges in the northeast Atlantic. *Prog Oceanogr* 61: 57–98.
- Klitgaard, A.B., Tendal, O.S., and Westerberg, H. (1997). Mass occurrences of large sponges (Porifera) in Faroe Island (NE Atlantic) Shelf and slope areas: Characteristics, distribution and possible causes. In Hawkins, L.E. and Hutchinson, S. (eds), Responses of marine organisms to their environment. Southampton, England: *Proceedings of the 30th European Marine Biology Symposium*, 129–142.
- Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B.E., Bussonnier, M., Frederic, J., et al. (2016). Jupyter Notebooks - a publishing format for reproducible computational workflows. In Loizides, F. and Schmidt, B. (eds), Positioning and power in academic publishing: players, agents and agendas. Amsterdam, The Netherlands: *IOS Press*, 87–90.
- Koutsouveli, V., Taboada, S., Moles, J., Cristobo, J., Ríos, P., Bertran, A., Solà, J., Avila, C., Riesgo, A. (2018). Insights into the reproduction of some Antarctic dendroceratid, poecilosclerid, and haplosclerid demosponges. *PLoS ONE* 13(2): e0192267.
- Kozich, J.J., Westcott, S.L., Baxter, N.T., Highlander, S.K., and Schloss, P.D. (2013). Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform. *Appl Environ Microbiol* 79: 5112–5120.
- Levin, L.A. (2006). Recent progress in understanding larval dispersal: New directions and digressions. *Integr Comp Biol* 46: 282–297.
- Madec, G. (2016). NEMO ocean engine. Note Du Pôle Modélisation, *Inst Pierre-Simon Laplace* 406.
- Martín, J., Palanques, A., and Puig, P. (2006). Composition and variability of downward particulate matter fluxes in the Palamós submarine canyon (NW Mediterranean). *J Mar Syst* 60: 75–97.
- McClain, C.R. and Hardy, S.M. (2010). The dynamics of biogeographic ranges in the deep sea. *Proc R Soc B Biol Sci* 277: 3533–3546.
- Meirmans, P.G. and Van Tienderen, P.H. (2004). GENOTYPE and GENODIVE: Two programs for the analysis of genetic diversity of asexual organisms. *Mol Ecol Notes* 4: 792–794.
- Moitinho-Silva, L., Nielsen, S., Amir, A., Gonzalez, A., Ackermann, G.L., Cerrano, C., Astudillo-García, C., Easson, C., Sipkema, D., Liu, F., Steinert, G., Kotoulas, G., McCormack, G.P., Feng, G., Bell, J.J., Vicente, J., Björk, J.R., Montoya, J.M., Olson, J.B., Reveillaud, J., Steindler, L., Pineda, M.-C., Marra, M.V., Ilan, M., Taylor, M.W., Polymenakou, P., Erwin, P.M., Schupp, P.J., Simister, R.L., Knight, R., Thacker, R.W., Costa, R., Hill, R.T., Lopez-Legentil, S., Dailianis, T., Ravasi, T., Hentschel, U., Li, Z., Webster, N.S., Thomas, T. (2017a) The sponge microbiome project. *Gigascience* 6: gix077.
- Moitinho-Silva, L., Steinert, G., Nielsen, S., Hardoim, C.C.P., Wu, Y.C., McCormack, G.P., et al. (2017b). Predicting the HMA-LMA status in marine sponges by machine learning. *Front Microbiol* 8: 1–14.
- Moran, N.A., Ochman, H., and Hammer, T.J. (2019). Evolutionary and ecological consequences of gut microbial communities. *Annu Rev Ecol Syst* 50: 451–475.
- Morato, T., Hoyle, S.D., Allain, V., and Nicol, S.J. (2010). Seamounts are hotspots of pelagic biodiversity in the open ocean. *Proc Natl Acad Sci* 107: 9707–9711.
- Mulisch, M. and Welsch, U. (2015). Romeis Mikroskopische Technik Aufl. 19. *Springer Spektrum*, Heidelberg, Germany.
- Murillo, F.J., Muñoz, P.D., Cristobo, J., Ríos, P., González, C., Kennington, E., and Serrano, A. (2012). Deep-sea sponge grounds of the Flemish Cap, Flemish Pass and the Grand Banks of Newfoundland (Northwest Atlantic Ocean): Distribution and species composition. *Mar Biol Res* 8: 842–854.
- Muyzer, G., Waal, E.C.D.E., and Uitierlinden, A.G. (1993). Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. *Appl Environ Microbiol* 59: 1–6.
- Palumbi, S.R. (2004). Marine reserves and ocean neighborhoods: The spatial scale of marine populations and their management. *Annu Rev Environ Resour* 29: 31–68.
- Paris, J.R., Stevens, J.R., and Catchen, J.M. (2017). Lost in parameter space: A road map for Stacks. *Meth Ecol Evol* 8:1360-1373.
- Pérez-Portela, R., Cerro-Gálvez, E., Taboada, S., Tidu, C., Campillo-Campbell, C., Mora, J., and Riesgo, A. (2016). Lonely populations in the deep: genetic structure of red gorgonians at the heads of submarine canyons in the north-western Mediterranean Sea. *Coral Reefs* 35: 1013–1026.
- Peterson, B.K., Weber, J.N., Kay, E.H., Fisher, H.S., and Hoekstra, H.E. (2012). Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* 7: e37135.
- Philippe, H., Brinkmann, H., Lavrov, D. V., Littlewood, D.T.J., Manuel, M., Wörheide, G., and Baurain, D. (2011). Resolving difficult phylogenetic questions: Why more sequences are not enough. *PLoS Biol* 9: e1000602.
- Pita, L., Rix, L., Slaby, B.M., Franke, A., and Hentschel, U. (2018). The sponge holobiont in a changing ocean: from microbes to ecosystems. *Microbiome* 6: 46.
- Pritchard, J.K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155: 945–959.
- QGIS Development Team (2017). Geographic Information System. Open Source Geospatial Foundation Project. Available online at: <http://qgis.osgeo.org>.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res* 41: 590–596.
- R Development Core Team (2008). R: A language and environment for statistical computing.
- Ramirez-Llodra, E., Clark, M.R., Smith, C.R., Tyler, P.A., Rowden, A.A., Bergstad, O.A., et al. (2011). Man and the last great wilderness: Human impact on the deep sea. *PLoS One* 6: e22588.
- Ramirez-Llodra, E., Brandt, A., Danovaro, R., De Mol, B., Escobar, E., German, C.R., Levin, L.A., Martínez Arbizu, P., Menot, L., Buhl-Mortensen, P., Narayanaswamy, B.E., Smith, C.R., Tittensor, D.P., Tyler, P.A., Vanreusel, A., Vecchione, M. (2010). Deep, diverse and definitely different: unique attributes of the world's largest ecosystem. *Biogeosciences* 7(9):2851-2899.
- Redmond, N.E., Morrow, C.C., Thacker, R.W., Diaz, M.C., Boury-Esnault, N., Cárdenas, P., et al. (2013). Phylogeny and systematics of demospongiae in light of new small-subunit ribosomal DNA (18S) sequences. *Integr Comp Biol* 53: 388–415.
- Rumín-Caparrós, A., Sanchez-Vidal, A., González-Pola, C., Lastras, G., Calafat, A., and Canals, M. (2016). Particle fluxes and their drivers in the Avilés submarine canyon and adjacent slope, central Cantabrian margin, Bay of Biscay. *Prog Oceanogr* 144: 39–61.
- Sánchez, F., Gómez-Ballesteros, M., González-Pola, C., and Punzón, A. (2014). Sistema de Cañones Submarinos de Avilés, Madrid, Spain.
- Sánchez, F., Rodríguez Basalo, A., García-Alegre, A., and Gómez-Ballesteros, M. (2017). Hard-bottom bathyal habitats and key-stone epibenthic species on Le Danois Bank (Cantabrian Sea). *J Sea Res* 130: 134–153.

- Sánchez, F., Serrano, A., Parra, S., Ballesteros, M., and Cartes, J.E. (2008). Habitat characteristics as determinant of the structure and spatial distribution of epibenthic and demersal communities of Le Danois Bank (Cantabrian Sea, N. Spain). *J Mar Syst* 72: 64–86.
- Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30(9):1312–1313.
- Taboada, S., Ríos, P., Mitchell, A., Cranston, A., Busch, K., Tonzo, V., Cárdenas, P., Sánchez, F., Leiva, C., Koutsouveli, V., Cristobo, J., Xavier, J., Hentschel, U., Rapp, H.T., Morrow, C., Drewery, J., Romero, P., Arias, M.B., Riesgo, A. (in review). Genetic diversity, gene flow and hybridization in fan-shaped sponges *Phakellia* spp. in the North-East Atlantic deep sea. submitted to *Molecular Ecology*.
- Topsent, E. (1890). Notice préliminaire sur les spongiaires recueillis durant les campagnes de l'Hirondelle. *Bulletin de la Société zoologique de France*. 15: 26-32, 65-71.
- van Sebille, E., Griffies, S.M., Abernathy, R., Adams, T.P., Berloff, P., Biastoch, A., et al. (2018). Lagrangian ocean analysis: Fundamentals and practices. *Ocean Model* 121: 49–75.
- Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W.S., and Huttenhower, C. (2011). Metagenomic biomarker discovery and explanation. *Genome Biol* 12: R60.
- Van Soest and de Voogd (2015). Sponge species composition of north-east Atlantic cold-water coral reefs compared in a bathyal to inshore gradient. *Journal of the Marine Biological Association of the United Kingdom* 95(7):1461-1474.
- Steinert, G., Rohde, S., Janussen, D., Blaurock, C., and Schupp, P.J. (2017). Host-specific assembly of sponge-associated prokaryotes at high taxonomic ranks. *Sci Rep* 7: 1–9.
- Taylor, M.L. and Roterman, C.N. (2017). Invertebrate population genetics across Earth's largest habitat: The deep-sea floor. *Mol Ecol* 26: 4872–4896.
- Thomas, T., Moitinho-Silva, L., Lurgi, M., Björk, J.R., Easson, C., Astudillo-García, C., et al. (2016). Diversity, structure and convergent evolution of the global sponge microbiome. *Nat Commun* 7: 11870.
- Tsujino, H., Urakawa, S., Nakano, H., Small, R.J., Kim, W.M., Yeager, S.G., et al. (2018). JRA-55 based surface dataset for driving ocean–sea-ice models (JRA55-do). *Ocean Model* 130: 79–139.
- Vad, J., Kazanidis, G., Henry, L.-A., Jones, D.O.B., Tendal, O.S., Christiansen, S., Henry, T.B., Roberts, J.M. (2018). Potential impacts of offshore oil and gas activities on deepsea sponges and the habitats they form. *Adv. Mar. Biol.* 79:33-60.
- White, J.W., Schroeger, J., Drake, P.T., Edwards, C.A. (2014). The value of larval connectivity information in the static optimization of marine reserve design. *Conserv. Lett.* 7:533–544.

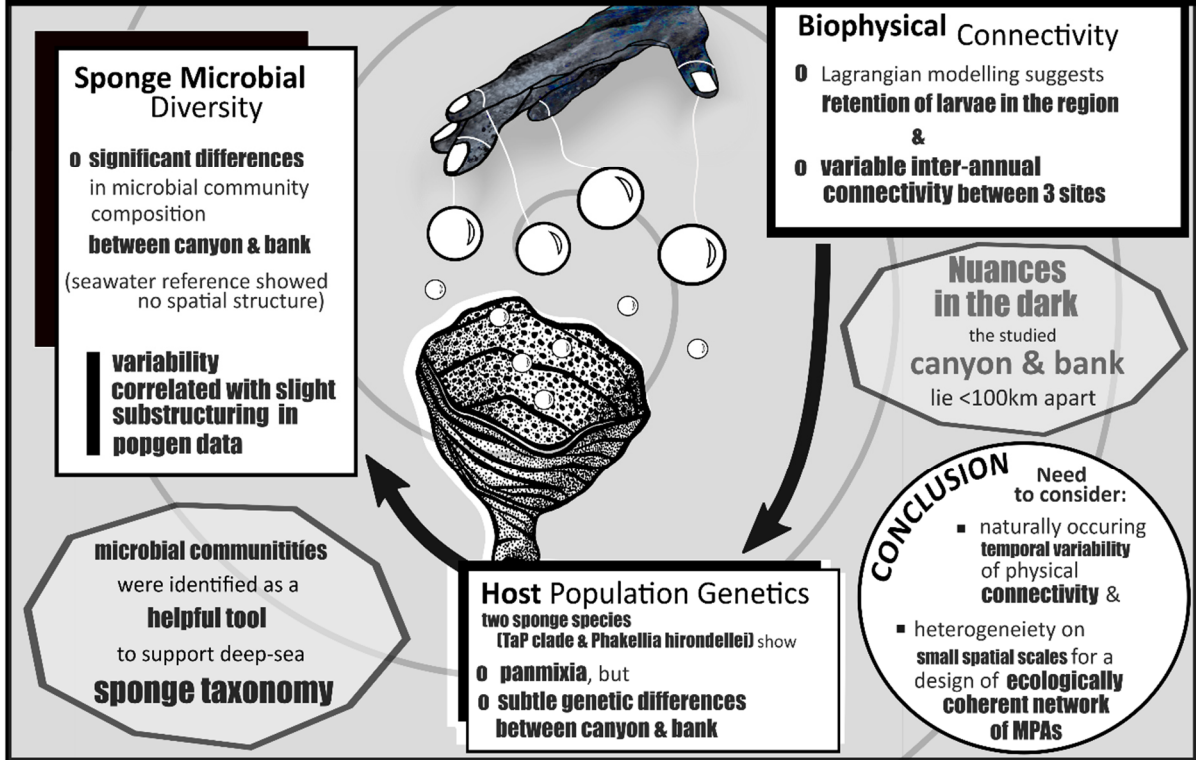
GRAPHICAL SUMMARY

Population connectivity of fan-shaped sponge holobionts in the deep Cantabrian Sea



§ by SponGES; illustrated by K. Busch

Busch · Taboada · Riesgo · Koutsouveli · Ríos · Cristobo · Franke · Getzlaff · Schmidt · Biastoch · Hentschel



Chapter 6

On giant shoulders: how a seamount affects the microbial community composition of seawater and sponges

© 2020 Busch et al.

Original publication by *Biogeosciences* (Copernicus)

Reprinted under the terms of the Creative Commons Attribution License (CC-BY 4.0).

doi: 10.5194/bg-17-3471-2020.

On giant shoulders: how a seamount affects the microbial community composition of seawater and sponges

Kathrin Busch¹, Ulrike Hanz², Furu Mienis², Benjamin Mueller³, Andre Franke⁴, Emyr Martyn Roberts⁵, Hans Tore Rapp^{5,†}, and Ute Hentschel^{1,6,*}

¹Research Unit Marine Symbioses, Research Division Marine Ecology, GEOMAR Helmholtz Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany, ²Department of Ocean Systems, NIOZ Royal Netherlands Institute for Sea Research, and Utrecht University, 1790 AB Den Burg, Texel, the Netherlands, ³Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Science Park 904, P.O. Box 94248, Amsterdam, the Netherlands, ⁴Department of Genetics and Bioinformatics, Institute of Clinical Molecular Biology (IKMB), Rosalind-Franklin-Straße 12, 24105 Kiel, Germany, ⁵Department of Biological Sciences, University of Bergen, K.G. Jebsen Centre for Deep Sea Research, P.O. Box 7803, 5020 Bergen, Norway, ⁶Faculty of Mathematics and Natural Sciences, Christian-Albrechts-Universität zu Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany. [†]deceased, 7 March 2020. *Correspondence author.

Seamounts represent ideal systems to study the influence and interdependency of environmental gradients at a single geographic location. These topographic features represent a prominent habitat for various forms of life, including microbiota and macrobiota, spanning benthic as well as pelagic organisms. While it is known that seamounts are globally abundant structures, it still remains unclear how and to which extent the complexity of the sea floor is intertwined with the local oceanographic mosaic, biogeochemistry and microbiology of a seamount ecosystem. Along these lines, the present study aimed to explore whether and to what extent seamounts can have an imprint on the microbial community composition of seawater and of sessile benthic invertebrates, sponges. For our high-resolution sampling approach of microbial diversity (16S rRNA gene amplicon sequencing) along with measurements of inorganic nutrients and other biogeochemical parameters, we focused on the Schulz Bank seamount ecosystem, a sponge ground ecosystem which is located on the Arctic Mid-Ocean Ridge. Seawater samples were collected at two sampling depths (mid-water, MW, and near-bed water, BW) from a total of 19 sampling sites. With a clustering approach we defined microbial microhabitats within the pelagic realm at Schulz Bank, which were mapped onto the seamount's topography and related to various environmental parameters (such as suspended particulate matter, SPM; dissolved inorganic carbon, DIC; silicate, SiO_4^- ; phosphate, PO_4^{3-} ; ammonia, NH_4^+ ; nitrate, NO_3^{2-} ; nitrite, NO_2^- ; depth; and dissolved oxygen, O_2). The results of our study reveal a "seamount effect" (sensu stricto) on the microbial mid-water pelagic community at least 200 m above the sea floor. Further, we observed a strong spatial heterogeneity in the pelagic microbial landscape across the seamount, with planktonic microbial communities reflecting oscillatory and circulatory water movements, as well as processes of benthic-pelagic coupling. Depth, NO_3^{2-} , SiO_4^- , and O_2 concentrations differed significantly between the determined pelagic microbial clusters close to the sea floor (BW), suggesting that these parameters were presumably linked to changes in microbial community structures. Secondly, we assessed the associated microbial community compositions of three sponge species along a depth gradient of the seamount. While sponge-associated microbial communities were found to be mainly species-specific, we also detected significant intra-specific differences between individuals, depending on the pelagic near-bed cluster they originated from. The variable microbial phyla (i.e. phyla which showed significant differences across varying depth, NO_3^{2-} , SiO_4^- , O_2 concentrations and different from local seawater communities) were distinct for every sponge-species when considering average abundances per species. Variable microbial phyla included representatives of both those taxa traditionally counted for the

variable community fraction and taxa counted traditionally for the core community fraction. Microbial co-occurrence patterns for the three examined sponge species *Geodia hentscheli*, *Lissodendoryx complicata*, and *Schaudinnia rosea* were distinct from each other. Over all, this study shows that topographic structures such as the Schulz Bank seamount can have an imprint (seamount effect sensu lato) on both the microbial community composition of seawater and sessile benthic invertebrates such as sponges by an interplay between the geology, physical oceanography, biogeochemistry, and microbiology of seamounts.

INTRODUCTION

Seamounts and mid-ocean ridges are prominent geologic features that add to the complexity of the sea floor. In the traditional sense, seamounts are defined as isolated submarine volcanic features with a minimum height of 1000 m from base to summit (Menard, 1964; Wessel et al., 2010). However, geologic features with a height of 50-100 m may also be considered as seamounts (Staudigel et al., 2010, Smith and Cann, 1992; Wessel et al., 2010). There may be up to 100 000 to > 25 million seamounts present in the oceans (IUCN, 2013), although the error rate associated with these estimations is high (IUCN, 2013). Despite the lack of accurate numbers, there is no doubt that with an estimated 10 million km² coverage, the area occupied by these habitats is globally significant. Elevated topographic features in the open ocean are often hotspots of biological diversity and productivity (IUCN, 2013; Morato et al., 2010). It appears that interaction of the topography with the hydrography creates a combination of amplified tidal flow, increased current speed, and the formation of internal waves, which strongly enhances vertical mixing around seamounts (Lavelle and Mohn, 2010; Van Haren et al., 2017; Roberts et al., 2018). Consequential upwelling of nutrient-rich deep waters stimulates primary productivity in this layer of enhanced mixing (IUCN, 2013). In addition to vertical mixing processes, horizontal fluxes of organic matter may also be affected by the presence of seamounts, as they may promote enclosed or semi-enclosed oceanographic circulation patterns, like Taylor caps or columns (Chapman and Haidvogel, 1992; Roberts et al., 2018, and references therein), leading to a retention of organic and inorganic matter.

The above-mentioned processes make seamounts important habitats for pelagic as well as benthic species (Morato et al., 2010; Rogers, 2018) due to beneficial prevailing conditions. Particularly areas with strong water flows (evoked by interactions of currents and tides with elevated topography), in combination with a steep and irregular hard substrate, represent suitable habitats for benthic suspension feeders, which indeed densely populate most seamounts (Genin et al., 1986; IUCN, 2013). Sponges (Porifera; Grant, 1836) often dominate these suspension feeder communities and are increasingly recognised as key components of shallow and deep marine ecosystems (DeGoeij et al., 2017; Maldonado et al., 2016). Due to their high filtering capacity and association with diverse microbial communities, sponges are considered to substantially influence the carbon, nitrogen, and silicate cycling in marine systems (Taylor et al., 2007; Maldonado et al., 2012, 2019; De Goeij et al., 2013; Rix et al., 2016a) and to contribute to benthic-pelagic coupling by actively removing particulate organic matter (POM) from the water column (Pile et al., 1997; Reiswig, 1971; Ribes et al., 1999). In addition to their influence on particulate organic matter pools, many sponges have been identified to primarily feed on dissolved organic matter (DOM) (De Goeij et al., 2008; Mueller et al., 2014; Hoer et al., 2018; Gantt et al., 2019). Energy and nutrients stored in this DOM are then transferred into particulate detritus, which fuels benthic food webs (De Goeij et al., 2013; Rix et al., 2016b).

Intimate sponge-microbe associations have been observed throughout diverse habitats, reaching from coastal shallow sites in tropical and temperate regions to the deep sea and polar seas (Helber et al., 2019; Kennedy et

al., 2014; Moitinho-Silva et al., 2014; Naim et al., 2014; Schmitt et al., 2012; Steinert et al., 2019; Thomas et al., 2016). According to their microbiome, sponges can be classified to either feature high microbial abundance (HMA) or low microbial abundance (LMA) (Hentschel et al., 2003; Moitinho-Silva et al., 2017; Weisz et al., 2008). The dichotomy between HMA and LMA sponges is considered a main driver of the microbial community structure associated with shallow-water sponges (Moitinho-Silva et al., 2017). In comparison to shallow waters, few studies have been conducted on the microbiology of deep-sea sponges (Borchert et al., 2017; Jackson et al., 2013; Kennedy et al., 2014; Reveillaud et al., 2014; Steinert et al., 2020; Tian et al., 2016). However, for example for deep-sea sponges of the genus *Geodia* (*G. barretti*, *G. macandrewii*, *G. phlegraei*, *G. atlantica*), similar microbial phyla have been observed as in HMA shallow-water sponges, such as Acidobacteria, Poribacteria and Chloroflexi (Luter et al., 2017; Radax et al., 2012; Schöttner et al., 2013). In addition to the HMA-LMA dichotomy, an important factor in structuring the microbiomes of shallow-water sponges is host taxonomy, which is manifested in ubiquitous species-specific sponge microbiomes (Easson and Thacker, 2014; Steinert et al., 2017; Thomas et al., 2016). Systematic analyses of the influence of biogeochemical parameters (particularly dissolved inorganic substances) on sponge-associated microbial diversity and interactivity are still lacking, particularly in deep-sea sponges. Seamounts provide an ideal study system in this regard, as they offer the potential of examining steep environmental gradients over small spatial scales. Sponge ground ecosystems are areas harbouring high densities of structure-forming sponge individuals. The Arctic Schulz Bank seamount has been observed to host a rich and diverse community of sponges (Roberts et al., 2018; Meyer et al., 2019) and may be considered a sponge ground ecosystem harbouring a reservoir of yet unexamined microbial biodiversity.

The present study aimed to characterize the microbial community composition of

seawater surrounding the Schulz Bank seamount ecosystem, located on the Arctic Mid-Ocean Ridge. Seawater samples were collected at two sampling depths from a total of 19 sampling sites and the corresponding microbiome data were mapped onto the topography of the Schulz Bank seamount ecosystem. Secondly, we assessed the associated microbial community compositions of three sponge species along a depth gradient of the seamount. Diversity metrics, as well as changes in the abundance of individual microbial taxa, were related with a set of biogeochemical parameters. This study explores whether topographic structures such as the Schulz Bank seamount, can have an imprint on the microbial community composition both of seawater and of sessile benthic invertebrates, sponges.

METHODS

Description of the Schulz Bank seamount

Schulz Bank is located on the Arctic Mid-Ocean Ridge (73.8 °N, 7.5 °E) between the Greenland and Norwegian seas (**Fig. 1**). It is exposed to three main water masses: (i) the Norwegian Deep Water (NwDW) that is present at the base and flanks of the seamount; (ii) the intermediate water mass (NwArIW), which is most likely Norwegian Arctic Intermediate Water and occurs at the summit and shallower areas; and (iii) the warmer surface water mass (NwAtW) which is Norwegian Atlantic Water. Notably near-bed water masses at Schulz Bank's summit have unusually low temperatures of around 0 to -1 °C (Roberts et al., 2018). Estimations of the seamount's basal dimensions state conservative values of 10 km x 4 km to 15 km x 6 km (Roberts et al., 2018), but it may however also be larger as Schulz Bank belongs to a ridge system. The summit of the seamount is located at around 600 m below the water surface, and the base depth is at more than 2500 m below the water surface. In a two-dimensional view, Schulz Bank has a broadly elliptical shape (Roberts et al., 2018). Bathymetry data presented in this study were derived from the Bathymetry Data Portal of the European Marine Obser-

vation and Data Network (EMODnet), and spatial analyses were performed in QGIS (version 3.4.4) as well as ArcGIS (version 10.6).

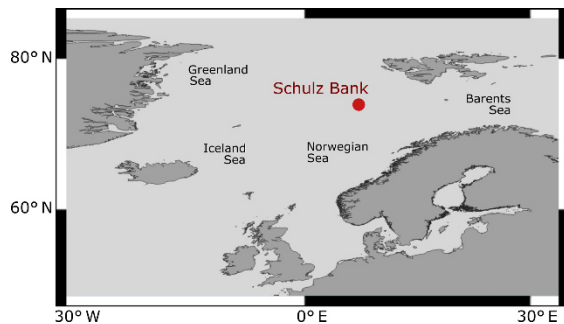


FIGURE 1 | Overview map showing the location of the Schulz Bank seamount between the Greenland and Norwegian seas.

Sampling procedures

Three cruises were undertaken onboard RV *G.O. Sars* (campaign names: GS2016109A, GS2017110, and GS2018108) during Northern Hemisphere summer in the years 2016-2018. Seawater samples were collected with a rosette water sampler equipped with 12 x 10 L Niskin bottles combined with a conductivity-temperature-depth (CTD) sensor system (SBE-9, Sea-Bird Electronics Inc., Washington, USA). In total, 19 CTD stations were covered, carried out along transects aligned with the seamounts' minor and major axes, and also with the 73.8 °N line of latitude. At each of the 19 stations, seawater samples for microbial analyses and biogeochemical parameters were collected at two water depths during the CTD upcast (**Fig. 3c** and PANGAEA for metadata): (i) 400 m below the seawater surface (mid-water) and (ii) correspondingly at 10 m above the sea floor (near-bed water). Naturally, the near-bed depths varied along with seamount topography and ranged from 575 to 2966 m. Sponges were sampled between 580 and 2184 m water depth, along the CTD transects (and therefore in close spatial vicinity to the seawater samples), by a remotely operated vehicle (ROV *Ægir 6000*, University of Bergen). A total of 36 sponge individuals representing the most abundant species were randomly selected from a larger collection effort. This

subset included 16 *Geodia hentscheli* (Cárdenas et al., 2010) (Demospongiae), 8 *Lissodendoryx complicata* (Hansen, 1885) (Demospongiae), and 12 *Schaudinna rosea* (Fristedt, 1887) (Hexactinellida). The sponges were taxonomically identified by visual inspection on board the ship. In addition, whole specimens and additional sponge samples were fixed in 99 % EtOH for deposition in the collections of the University of Bergen.

Biogeochemical analyses and measurements of environmental parameters

The following nine environmental parameters were analysed: depth, suspended particulate matter (SPM), dissolved inorganic carbon (DIC), silicate (Si), phosphate (PO_4^{3-}), ammonia (NH_4^+), nitrate (NO_3^{2-}), nitrite (NO_2^-), and dissolved oxygen (O_2). Depth and dissolved O_2 data were recorded during in situ water column profiling. Depth (pressure) was recorded with the CTD sensor system mentioned above. O_2 concentrations were derived from a dissolved oxygen sensor (SBE-43, Sea-Bird Electronics Inc., Washington, USA) that was attached to the rosette water sampler. For the analysis of suspended particulate matter (SPM), 2 x 10 L of water were filtered over pre-weighed combusted GF/F filters, which were rinsed with demineralized water to remove salts (47 mm Whatman™ GF/F filters pre-combusted at 450 °C, stored at -20 °C). Filters were freeze-dried and weighed before further analysis. For the analysis of inorganic nutrients (ammonia, NH_4^+ ; phosphate, PO_4^{3-} ; nitrate, NO_3^{2-} ; nitrite, NO_2^- ; and silicate, Si), seawater samples were filtered over 0.2 µm filters. Water samples for NH_4^+ , PO_4^{3-} , NO_x analysis were stored at -20 °C and for Si analyses at 4 °C. Nutrients were measured with a QuAatro gas segmented continuous flow analyser (Seal Analytical, Norderstedt, Germany). Measurements were made simultaneously on four channels for PO_4^{3-} (Murphy and Riley, 1962), NH_4^+ (Helder and de Vries, 1979), and NO_3^{2-} combined with NO_2^- (Grasshoff et al., 2009) and separately for Si (Strickland and Parsons, 1972). A freshly diluted mixed-nutrient standard

containing Si, PO_4^{3-} , and NO_3^{2-} was added to each run. The cocktail served as a guide to monitor the performance of the standards. All measurements were calibrated with standards diluted in low-nutrient seawater (LNSW). For the analysis of dissolved inorganic carbon (DIC), seawater samples were transferred into a glass vial containing 15 μL HgCl_2 (mercury chloride) and analysed on a TechniconTraacs800 auto-analyser (Technicon Instruments Corporation, Tarrytown, USA) following the methodology of Stoll et al. (2001). Analyses of variance (ANOVAs) were performed to test for statistical differences in the biogeochemical and physical parameters between the determined microbial near-bed water clusters (see below). As numbers of samples per mid-water cluster and per near-bed water cluster were unequal, we calculated Type III sums of squares for ANOVAs (unbalanced ANOVAs). We further calculated Spearman's rank correlations between depth and those biogeochemical parameters which turned out to differ significantly across the determined near-bed water clusters in the ANOVA analyses (see below).

Amplicon sequencing

Seawater samples were collected in triplicates from different Niskin bottles, yielding a total of 114 samples from all stations (i.e. 19 stations x 2 sampling depths x 3 biological replicates per station and depth = 114 seawater samples in total). A total of 2 L of seawater sample was filtered onto polyvinylidene fluoride (PVDF) filter membranes (Merck Millipore) with a pore size of 0.22 μm and a diameter of 47 mm and stored at -80°C . For sponge collection, cubes of approximately 1 cm^3 were cut from the mesohyl with a scalpel, rinsed (sterile seawater), flash-frozen in liquid nitrogen, and stored at -80°C . DNA was extracted from half a seawater filter or ~ 0.25 g of sponge tissue by using the DNeasy Power Soil Kit (Qiagen, Venlo, the Netherlands). The quality of the DNA extraction was assessed based on the 260/280 ratio using a NanoDrop spectrophotometer as well as by polymerase chain reaction with uni-

versal 16S primers and subsequent gel electrophoresis. The V3-V4 variable regions of the 16S rRNA gene were then amplified in a one-step polymerase chain reaction (PCR) using the primer pair 341F-806R (dual-barcoding approach; Kozich et al., 2013; primer sequences: 5'-CCTACGGGAGGCAGCAG-3' and 5'-GGAC-TACHVGGGTWTCTAAT-3'). After verification of the presence of PCR products by gel electrophoresis, normalization (SequalPrep normalisation plate kit; ThermoFisher Scientific, Waltham, USA) and equimolar pooling were performed. Sequencing was conducted on the MiSeq platform (MiSeqFGx; Illumina, San Diego, USA) with v3 chemistry. The settings for demultiplexing were 0 mismatches in the barcode sequences.

Bioinformatic analyses

For computation of microbial core-diversity metrics, sequences were processed within the QIIME2 environment (version 2018.11; Bolyen et al., 2019). Amplicon sequence variants (ASVs) were generated from forward reads (truncated to 270nt) with the DADA2 algorithm (Callahan et al., 2016). Phylogenetic trees were calculated based on resulting ASVs with the FastTree2 plugin. Representative ASVs were classified using the Silva 132 99 % OTUs 16S database (Quast et al., 2013) with the help of a primer-specific trained naïve Bayes taxonomic classifier. Alpha and beta diversity indices (e.g. Faith's phylogenetic diversity and weighted UniFrac distances, respectively) were calculated within QIIME2. To evaluate sample separation in ordination space, non-metric multidimensional scaling (NMDS) was performed on weighted UniFrac distances for seawater and sponge-associated microbiomes separately.

A machine learning approach was used to define microbial microhabitats within the pelagic realm. Seawater microbiomes were clustered based on weighted UniFrac distances. The NbClust function was applied in R (version 3.0.2; R Development Core Team, 2008) to generate 30 indices to identify the best number of clusters based on the majority rule. A coordinate grid was set up as a basis for

a georeferenced extrapolation of sampling points. In detail, we used the sampling coordinates and the determined cluster affiliation of our 114 in situ measured seawater samples as the input for our predictions of seawater microbial cluster affiliation in space. Clustering regions were set up with the help of the *k*-nearest-neighbor-algorithm. The machine learning approach was fine-tuned in several ways: (i) the algorithm was trained in a way that in situ measured data points always belong to the cluster actually determined based on the sequencing data; (ii) a normalisation was applied with the help of a distance-weighted function meaning that closer data points have a higher weight; (iii) the probability of class membership was calculated and plotted as indication of confidence. Permutational multivariate analyses of variance (PERMANOVAs) were performed with 999 permutations to determine whether microbiomes of selected clusters were statistically significantly different from each other. In detail, pairwise tests across the determined clusters were conducted for the following samples separately: mid-water samples, near-bed water samples, *G. hentscheli*, *L. complicata*, and *S. rosea*. A significance level of $\alpha=0.05$ was applied for all statistical analyses in this study.

To evaluate co-occurrence patterns between microbial taxa across environmental gradients (i.e. determined near-bed water clusters), networks were constructed separately for every sponge species and seawater. Mean relative abundances of microbial phyla were calculated for all biological replicates of each sample type and for the corresponding near-bed water cluster. Microbial phyla, which showed significantly different enrichment between clusters, were determined and ranked using the linear discriminant analysis effect size (LEfSe) algorithm (Segata et al., 2011). A correlation matrix was established for those taxa that differed significantly between clusters, to assess co-occurrences. In particular, the direction and strength of correlations were characterised for any significant phylum with

all other significant taxa (as well as the relations with depth).

RESULTS

Structure and composition of seawater microbial communities

A NMDS plot on weighted UniFrac distances separated the microbial communities of mid-water and near-bed water samples in ordination space with few exceptions (**Fig. 2**).

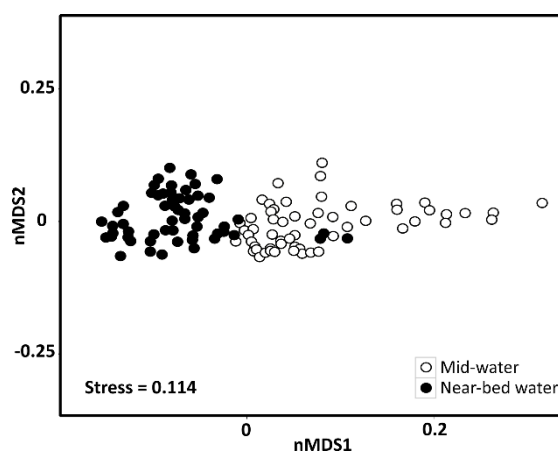


FIGURE 2 | Seawater microbial community composition of mid-water and near-bed water samples visualized by a non-metric multidimensional scaling plot on weighted UniFrac distances. Each marker is one microbial community, with colours indicating the sample subtype (i.e. mid-water or near-bed water).

Cluster analysis based on weighted UniFrac distances revealed two distinct clusters in the mid-water samples, of which one (MW1) was located precisely above the summit of Schulz Bank, while the other (MW2) covered the wider seamount region and vicinity (**Fig. 3a**). Four distinct microbiome clusters were detected in the near-bed water samples (BW1-4). In terms of similarity, cluster BW1 was most distinct from all other clusters while clusters BW2 and BW3 were most similar to each other (**Fig. 3b**). Moreover, BW1 cluster samples separated in ordination space in that they grouped with mid-water rather than near-bed water samples (i.e. consider the few black dots grouped together with the white dots in **Fig. 2**).

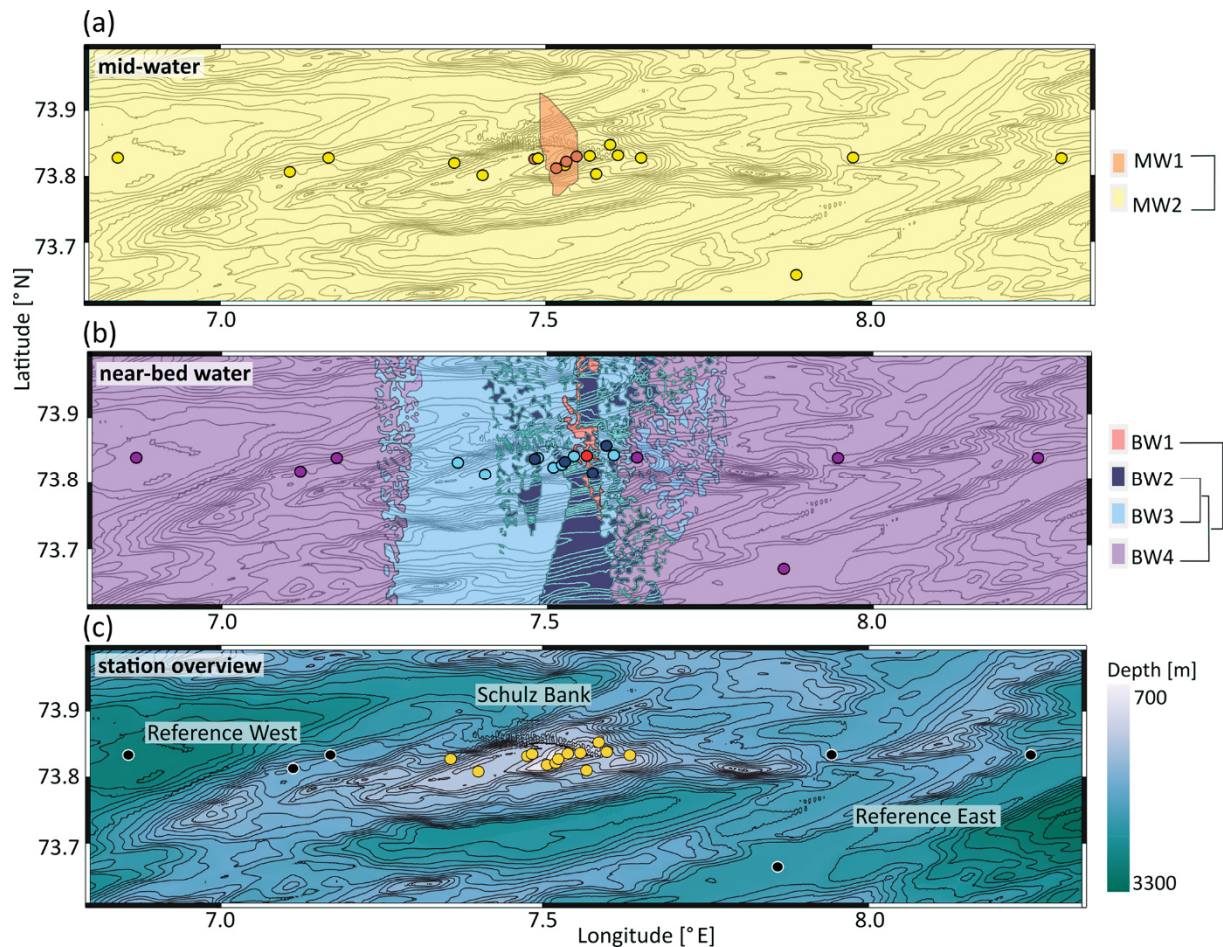


FIGURE 3 | Seawater microbial community structure across Schulz Bank. Contour lines in all three subplots represent the underlying topography. Colours in panels (a) and (b) represent clusters based on weighted UniFrac distances, where coloured dots indicate stations with in situ sampling and filled areas represent extrapolations based on machine learning. The further away predicted areas are from actual sample points, the higher the associated uncertainty of these predictions is. Panel (a) includes all mid-water samples derived during the CTD transects. Panel (b) includes all near-bed water samples. Here, the degree of cluster similarity can be deduced from the dendrogram to the right of the plot(s). Panel (c) provides an overview of the sampling area, showing the locations of all 19 CTD stations. Stations directly located on the Schulz Bank are coloured yellow, while reference stations (west and east of Schulz Bank) are indicated in black. Colouring in panel (c) was done according to depth.

Plotting the clusters on a spatial map revealed that near-bed water cluster BW1 was located near the summit of Schulz Bank seamount (average depth = 575 m), while clusters BW2 and BW3 covered its flanks (average depth \pm SE = 919 ± 106 and 922 ± 142 m, respectively), and cluster BW4 represented the vicinity close to the seamount (average depth \pm SE = 1836 ± 376 m) (Fig. 3b). Statistical testing of the individual depth data points contributing to a given cluster revealed a significant difference in the

depth parameter between the clusters (ANOVA, $p = 0.01$, $df1 = 3$; $df2 = 10$).

Figure 3c shows the bathymetry highlighting the contour lines of Schulz Bank seamount and its vicinity (reference west and east) as well as the 19 sampling stations. As our samples originated from three different cruises, we also double-checked the extent of temporal variability in our dataset. We observe frequently that seawater samples of different

cruises belong to the same cluster (Fig. S1a in the Supplement). From this we conclude that temporal variations are not the main driver of our observed similarities or dissimilarities in pelagic microbial community compositions.

Microbial richness was overall slightly lower in the mid-water samples (mean Faith's phylogenetic diversity \pm standard error = 45.5 ± 0.8) than in the near-bed water samples (54.4 ± 0.9) (Fig. S2a). Near-bed water samples from the summit (BW1) displayed a slightly lower microbial richness than the other near-bed water samples (Fig. S2b). The mid-water samples collected above Schulz summit also showed a slightly lower microbial richness than the other mid-water samples. Pairwise comparisons (PERMANOVA) revealed that the seawater microbial community clusters within the mid-water and near-bed water samples were significantly different from each other in terms of their microbial community composition (Table S1). Furthermore, the pool of mid-water samples (MW1 - MW2) was significantly different from the pool of near-bed water samples (BW1 - BW4). Overall, the eight most dominant seawater microbial phyla, sorted in descending order of mean relative abundance, were Proteobacteria (54 % of total community), Bacteroidetes (17 %), Verrucomicrobia (7 %), Marinimicrobia (SAR406 clade) (6 %), Actinobacteria (5 %), Chloroflexi (4 %), Acidobacteria (2 %), and Planctomycetes (1 %).

Seawater biogeochemistry at Schulz Bank seamount

When comparing the biogeochemical parameters of the mid-water clusters, only dissolved O₂ concentrations differed significantly (ANOVA, $p = 0.02$, $df_1 = 1$; $df_2 = 17$) with slightly higher concentrations in MW1 (6.90 ± 0.04 mL L⁻¹) compared to MW2 (6.75 ± 0.03 mL L⁻¹) (Table S2). All other tested biogeochemical

parameters SPM (2.22 ± 1.39 mg L⁻¹), DIC (2269.07 ± 26.63 μ mol L⁻¹), SiO₄⁻ (5.66 ± 0.06 μ mol L⁻¹), PO₄³⁻ (0.86 ± 0.01 μ mol L⁻¹), NH₄⁺ (0.11 ± 0.03 μ mol L⁻¹), NO₃²⁻ (12.99 ± 0.08 μ mol L⁻¹), and NO₂⁻ (0.02 ± 0.01 μ mol L⁻¹) were not statistically different between MW1 and MW2. The values for mid-water samples are reported as average \pm standard error.

Of the eight biogeochemical parameters tested, the following three differed significantly between the near-bed water clusters. These were NO₃⁻ (ANOVA, $p = 0.04$, $df_1 = 3$; $df_2 = 10$), SiO₄⁻ (ANOVA, $p = 0.03$, $df_1 = 3$; $df_2 = 10$), and dissolved O₂ (ANOVA, $p = 0.01$, $df_1 = 3$; $df_2 = 15$). Nitrate (range= 13.00-14.78 μ mol L⁻¹) and SiO₄⁻ (range = 6.00-10.65 μ mol L⁻¹) increased with depth, with lowest concentrations at the summit (BW1), intermediate concentrations at the flanks (BW2, BW3), and highest concentrations in the seamount vicinity (BW4) (Fig. 4). Dissolved oxygen (range= 6.48-6.99 mL L⁻¹) showed the reverse pattern in that its concentration was highest at the summit (BW1), intermediate at the flanks (BW2, BW3) and lowest in the seamount vicinity sites (BW4). Spearman's rank correlations calculated between depth and the three other significant parameters indeed revealed significant correlations in all cases (NO₃⁻ : $\rho = 0.77$, $p < 0.01$; SiO₄⁻ : $\rho = 0.85$, $p < 0.01$; dissolved O₂: $\rho = -0.77$, $p < 0.01$) (Fig. S3). The other biogeochemical parameters SPM (range= 0.49-1.87 mg L⁻¹), DIC (range= 2248.00-2265.67 μ mol L⁻¹), PO₄³⁻ (range= 0.90-0.97 μ mol L⁻¹), NH₄⁺ (range= 0.10-0.17 μ mol L⁻¹), and NO₂⁻ (range= 0-0.05 μ mol L⁻¹) were not significantly different between the near-bed water clusters. At the summit, no pronounced differences in biogeochemical parameters were observed between the near-bed water (BW1) and mid-water samples (MW1) (Table S2).

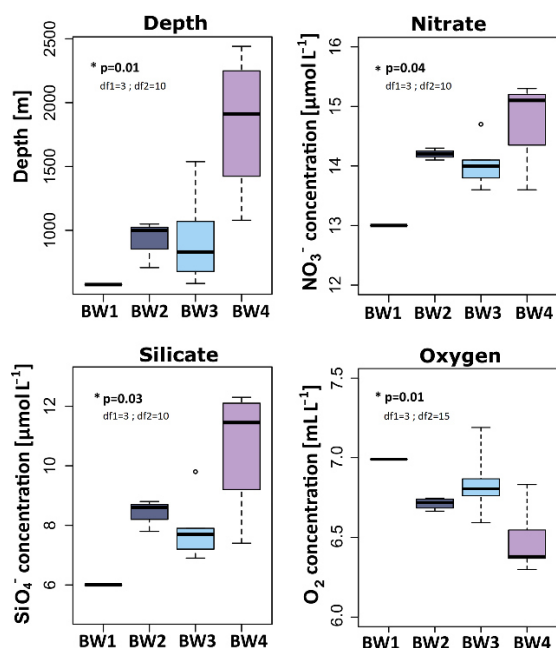


FIGURE 4 | Concentrations and measurements of significant (ANOVA, $\alpha = 0.05$) biogeochemical parameters for near-bed water samples, across the determined near-bed water clusters. The p values as well as degrees of freedom ($df1$ and $df2$) for these parameters are written into the respective graphs. Colouring is the same as for Figs. 2 and 3.

Structure and composition of sponge microbial communities

In order to analyze structure and microbial community composition of the sponges, we randomly selected at least four biological replicates per sponge species per BW cluster for statistical analysis. As sponges were sampled in the close vicinity but not exactly at the same locations as seawater samples, BW cluster affiliations of the sponge sampling locations were determined based on spatial extrapolations of the seawater data as described before (see Fig. S1b for exact sponge sampling locations). Overall, the three deep-sea sponge

species *S. rosea*, *G. hentscheli*, and *L. complicata* showed host species-specific microbiomes, as indicated by a clear separation of their microbial communities in ordination space (Fig. 5). Sub-structuring based on near-bed water clusters in the non-metric multidimensional scaling plot as well as pairwise comparisons (PERMANOVA) revealed that the sponge microbial communities within each species differed significantly depending on the near-bed water clusters from which they were collected (Table S1). The only exception was *S. rosea*, for which specimens from the flank (BW3) showed a microbial community composition that was intermediate between the summit (BW1) and the other flank cluster (BW2).

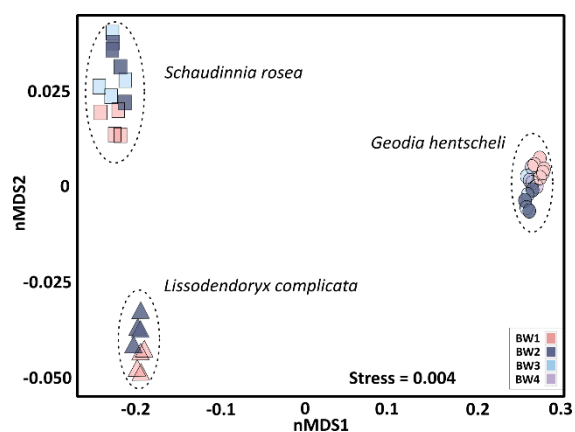


FIGURE 5 | Sponge microbial community composition visualized by a non-metric multidimensional scaling plot on weighted UniFrac distances. Each marker is one microbial community, with symbols representing the sample sub-type (i.e. sponge species) and colours indicating the near-bed water cluster present at the respective sponge sampling location.

The dominant microbial phylum in *S. rosea* and *L. complicata* was Proteobacteria (Fig. 6a and Fig. 6b), whereas *G. hentscheli* microbiomes were dominated by Chloroflexi, Acidobacteria and Proteobacteria (Fig. 6c).

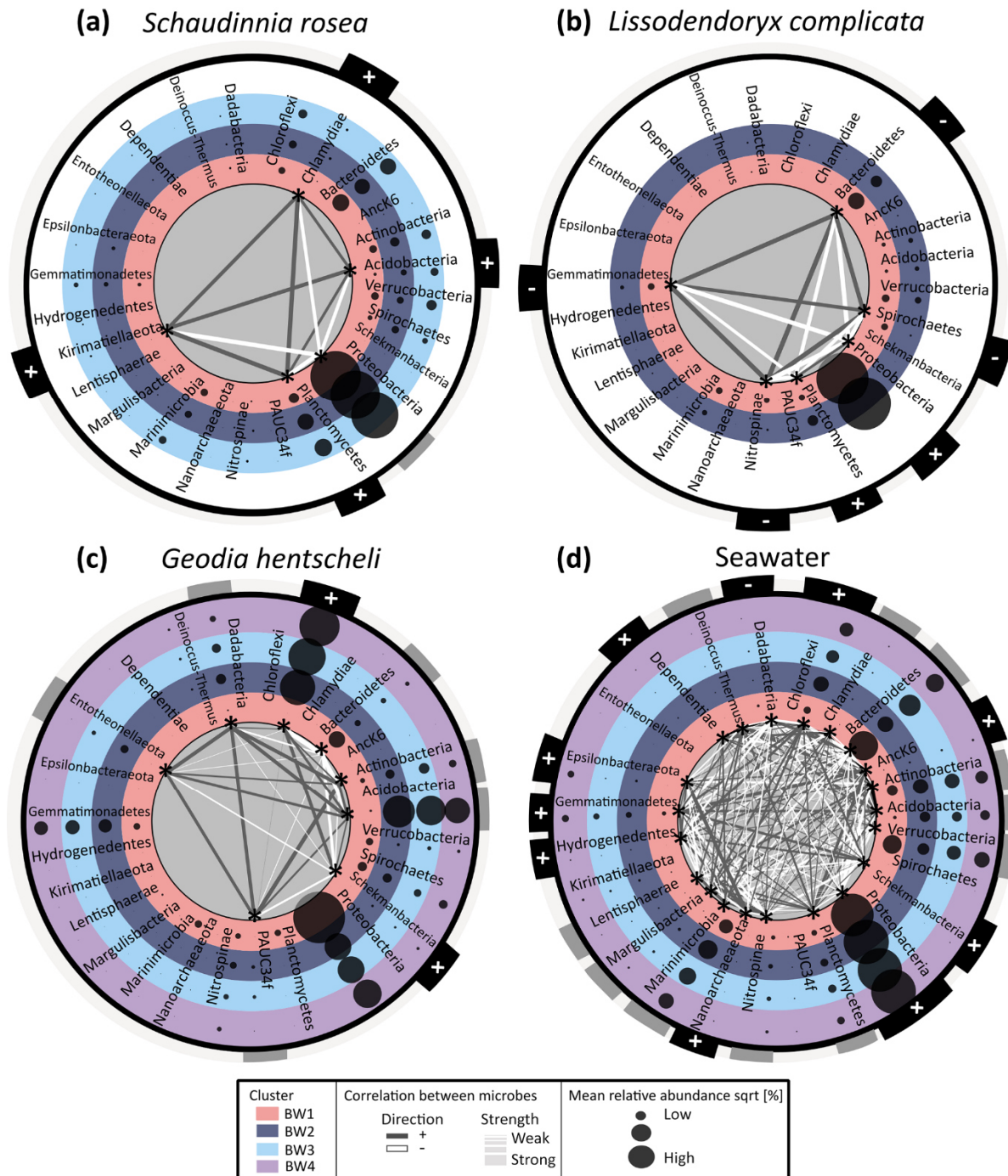


FIGURE 6 | Co-occurrence network and differential abundance of microbial phyla across the four determined near-bed water clusters. Panels (a)–(c) show sponge data, with panel (a) showing average *Schaudinnia rosea* data, panel (b) presenting average *Lissodendoryx complicata* data, and panel (c) illustrating *Geodia hentscheli* data. Panel (d) contains seawater data. Near-bed water clusters are represented by differently coloured rings. Each ring contains a list with microbial phyla which are alphabetically sorted. Average relative abundances of each of the respective phyla for the samples within a given cluster are indicated by bubble sizes. Those microbial phyla which are statistically significantly enriched or depleted across the four clusters (LEfSe analysis) are marked with an asterisk inside the inner most ring. Only for those taxa where the difference is significant are correlation strength (indicated by size of connecting lines) and direction (represented by colour of connecting lines: white is negative correlation; dark grey is positive correlation) with all other significant taxa plotted. For all microbial phyla relation with depth is indicated in the outer ring of each plot by + (meaning significant positive relation) or - (meaning significant negative relation).

Sponge microbiomes were more stable than seawater communities with fewer phyla exhibiting significant differences across the four near-bed water clusters or positively relating with depth (Fig. 6). For the hexactinellid *S. rosea*, the relative abundances of five bacterial phyla (Acidobacteria, Chlamydiae, Kirimatiellaota, Planctomyces and Proteobacteria) were significantly different between individuals that were sampled from different near-bed water clusters. Out of these five phyla, Acidobacteria, Chlamydiae, Kirimatiellaota, and Planctomyces were positively related with depth while for the Proteobacteria neither a positive nor negative relation with depth was discernable. Consequently, the Proteobacteria showed a negative relation with the four other phyla in the network analysis. For the demosponge *L. complicata*, the relative abundances of Bacteroidetes, Gemmatimonadetes, Nitrospinae, Planctomyces, Proteobacteria and Spirochaetes were significantly different between sponge individuals that were sampled from the different near-bed water clusters. For this sponge species, samples were only available from near-bed water clusters 1 and 2. Of the six phyla, Planctomyces and Proteobacteria were positively related with depth, while the other four were negatively related with depth, which is also reflected in the network analysis. For the demosponge *G. hentscheli*, the relative abundances of eight phyla (Acidobacteria, Actinobacteria, Bacteroidetes, Chloroflexi, Dadabacteria, Entotheonellota, PAUC34f, and Schekmanbacteria) were significantly different between sponge individuals sampled from the near-bed water clusters BW1-BW4. Of those, Chloroflexi and Schekmanbacteria were positively related with depth, while the others showed variable patterns over depth. The network analysis showed both positive and negative correlations between taxa for those

increasing with depth as well as those displaying a variable response.

When analysing host-associated microbiomes, ambient seawater microbiomes are valuable references for comparison. In this study, a total of 21 microbial clades were identified in ambient seawater, whose relative abundances varied significantly between the four near-bed water clusters. A total of nine taxa showed a positive relation with depth, one (Dadabacteria) showed a negative relation with depth, and the remaining 11 taxa showed a variable response to depth. For seawater more phyla varied between near-bed water clusters than for the sponge samples. Overall, more microbial taxa showed significant positive relations with depth, NO_3^{2-} , and SiO_4^- and negative relations with O_2 than vice versa. The microbial taxa showing a significant difference in relative abundance between the near-bed water clusters (as determined by LEfSe) were different between sponges and seawater and also between sponge species. Further, significant differences in relative abundances were observed for both abundant and less abundant sponge symbiont lineages.

DISCUSSION

Research records about seamount microbiology are sparse and comparably few studies have been conducted on deep-sea sponge microbiomes in general (Borchert et al., 2017; Jackson et al., 2013; Kennedy et al., 2014; Reveillaud et al., 2014). Our main aim was to assess whether and via which potential mechanisms a seamount can affect the community structure of pelagic and benthos (sponge)-associated microbial communities, using the Schulz Bank seamount as an exemplary field site. A total of 19 CTD sampling stations, each with two sampling depths, on and around Schulz Bank were analysed towards this goal and combined with sponge-associated microbial data gained during additional ROV dives.

A seamount imprint on seawater microbial communities

In this study we observed a pronounced similarity between the microbial community composition of the mid-water cluster located precisely above Schulz Bank's summit (MW1) and the microbial community composition of the near-bed water cluster at the summit (BW1). This is evident in **Fig. 2**, where few black dots representing the BW1 cluster group with mid-water samples rather than near-bed water samples. In addition, the microbial community in the mid-water cluster above Schulz Bank's

summit (MW1) was distinct from the community in the mid-water cluster covering the wider seamount region and vicinity (MW2), despite similar prevailing biogeochemical conditions in both mid-water clusters (MW1 vs MW2; an exception is the significant difference in O₂ concentrations between both clusters). From these two observations we conclude that the presence of a seamount can have an imprint on the microbial community structure in the overlying water column (seamount effect *sensu stricto*). In

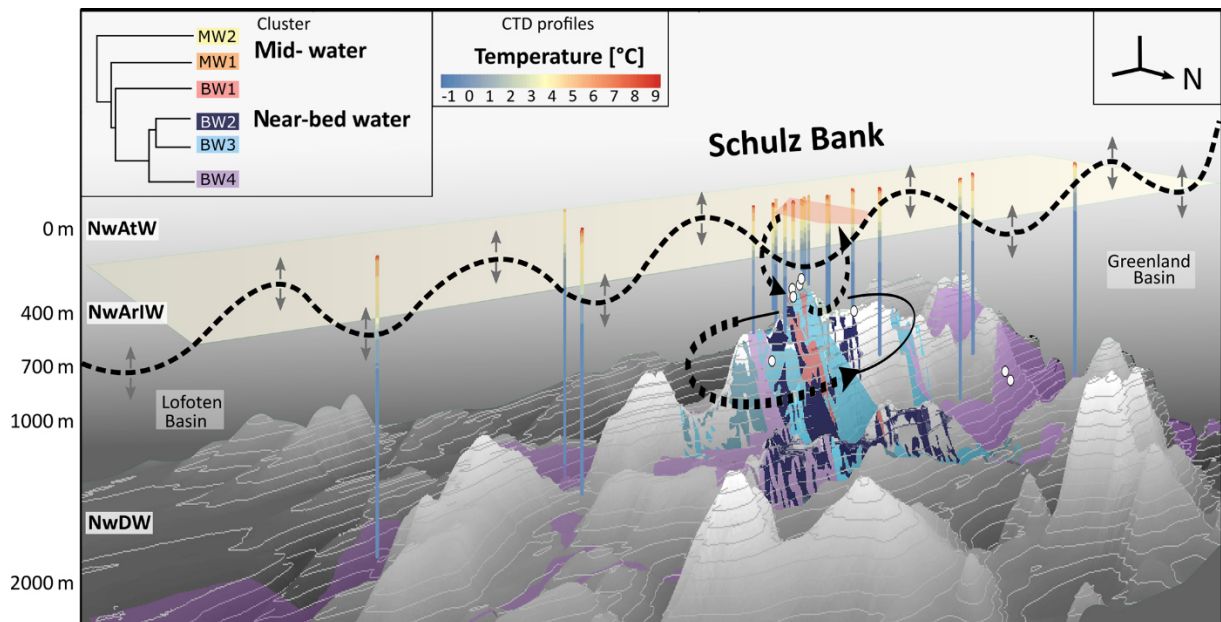


FIGURE 7 | Conceptual overview and vertical 3D section showing spatial distribution of microbial clusters and oceanographic patterns on the Schulz Bank seamount. Extrapolated seawater microbial clusters are indicated by coloured polygons: mid-water clusters are marked in orange (MW1) and yellow (MW2), while near-bed water clusters are marked in red (BW1), dark blue (BW2), light blue (BW3), and purple (BW4). The degree of cluster similarity can be deduced from the dendrogram in the left corner of the plot. Whole water column CTD profiles are indicated, showing the measured temperature values from surface to bottom at the respective sampling locations. Sponge sampling locations visible on this side of the seamount are indicated by white balls. Vertical positions of the major water masses Norwegian Atlantic Water (NwAtW), Norwegian Arctic Intermediate Water (NwArIW), and Norwegian Deep Water (NwDW) are indicated. To give a broad orientation in space, a north arrow is depicted, and the major geologic features (Lofoten Basin and Greenland Basin). For Schulz Bank, water flows, such as a potential Taylor column circulating around the seamount, mixing between summit and pelagic realm, as well as tidally driven internal motions (black horizontal line with bidirectional arrows), are indicated by dashed arrows and lines.

particular we suspect that topography-induced vertical mixing processes occur at Schulz Bank seamount, which reallocate microbial communities within the water column and in turn influence the pelagic microbial diversity as far as

at least 200 m above the seamount's summit. In support of these interpretations, oscillating currents relating to the barotropic and baroclinic (internal) tide have been reported previ-

ously at the summit of the Schulz Bank seamount (Roberts et al., 2018) and other seamounts (Van Haren et al., 2017). Although these mentioned studies suggest that respective oceanographic processes are particularly pronounced at the summit part of seamounts, we cannot exclude the following two aspects based on our own study: (i) since we only sampled one mid-water depth, vertical mixing of microbial communities could in theory extend far higher than 200 m above the seamount's summit. (ii) As the vertical distances between the mid-water and the near-bed water samples were larger at other parts of the seamount compared to the summit, vertical mixing of microbial communities may in fact not only be restricted to the summit of Schulz Bank.

In addition to tide-induced vertical hydrodynamic processes, horizontal flow patterns can also help to explain the presence of seamount-specific microbial communities. Roberts et al. (2018) calculated that a Taylor cap or Taylor column may be (temporarily) present at Schulz Bank. This oceanographic phenomenon describes an isolated anti-cyclonic flow circulation pattern over a seamount and hence may promote temporary spatial isolation of a seamount ecosystem from adjacent waters. Our conceptual schematic overview of hydrodynamics at Schulz Bank seamount is shown in **Fig. 7** (dashed lines). In this figure, a digital elevation model of Schulz Bank seamount is depicted in combination with the overlying water column structure and oceanographic context. Temperature profiles derived from whole water column sensing by CTD casts are plotted. Based on these profiles the vertical distributions of the surface water (NwAtW), intermediate water (NwArIW), and Norwegian Deep Water (NwDW) were deduced in combination with the identified water masses as described in Roberts et al. (2018) (**Fig. 7**). In addition to oceanographic insights, **Fig. 7** includes a 3D visualization of our determined microbiome clusters at and around Schulz Bank seamount. This figure recapitulates that similarities and dissimilarities in microbial signatures of seawater in this study

(cluster dendrogram **Fig.7**) were consistent with the oscillatory water movements (i.e., due to internal tide-induced mixing) and possible circulatory flows (Taylor column) as predicted by Roberts et al. (2018).

Microbial richness was overall slightly lower at the summit of the seamount (BW1) than at the deeper locations (i.e. BW2-BW4). A similar trend was observed for the mid-water samples, where microbial richness was slightly lower for the microbial community above Schulz's summit in comparison to samples in its vicinity (MW1 vs MW2). On a macroscopic level, Morato et al. (2010) and others have described seamounts as hotspots of pelagic biodiversity; much less is however known about microbial diversity at seamounts. Our results of a lower microbial richness above the seamount summit might seem contradictory at first. However, Schulz Bank is a recognized sponge ground ecosystem, with a peak in sponge density and diversity at the seamount summit (Roberts et al., 2018; Meyer et al., 2019). Sponges are very efficient suspension feeders and are known for removing large amounts of particulate organic matter including prokaryotes and small eukaryotes from the water column (Leys et al., 2018). Benthic-pelagic coupling mediated by selective feeding of sponges on seawater microorganisms (McMurray et al., 2016; Van Oevelen et al., 2018) in combination with the discussed hydrodynamic patterns (vertical mixing) might explain the slightly reduced microbial richness of the water body residing directly above a sponge ground (BW1 and MW1). Sponge density and community composition changes along the topography of Schulz Bank seamount in that density is highest at the summit (Meyer et al., 2019; Roberts et al., 2018). This natural variation can further explain the observed differences in microbial community composition between the other near-bed water samples (BW2-4). For these samples, we observed distinct pelagic microbial communities at a finer resolution than can be explained by the pure water masses distribution (consider depth of intermediate (NwArIW) and deep water (NwDW) layers in

Fig. 7). In particular the near-bed clusters BW2 and BW3, originated from a similar depth range and both were located at the seamount's flanks. Besides ecologically rooted explanations (i.e. variable presence of benthic organisms that influence biogeochemical cycles), hydrodynamic processes (i.e. local flow direction linked to small-scale topography and/or spatial orientation of the seamount's flanks) can also explain the observation of distinct microbial community compositions within the near-bed water. We hence conclude that besides patterns related to water masses, we observe a much higher spatial heterogeneity of pelagic microbial communities than previously recognized. We call this kind of imprint on the pelagic microbial community composition, which is probably based on the topography combined with benthic-pelagic coupling processes and hydrodynamics, a "seamount effect sensu lato". These observations suggest that the presence of a seamount can have profound impacts on the distribution of microbial landscapes in the open ocean.

Seamounts are recognized as unique habitats in terms of ecosystem dynamics (Genin and Boehlert, 1985) and macroecology (Morato et al., 2010). The present study reveals that seamounts also have a unique microbial signature that extends hundreds of meters up into the overlying water column. In addition, we detected three distinct microbial clusters in seawater samples taken near the seabed directly above Schulz Bank which were distinctly different from those of seawater collected in the vicinity of the seamount (near-bed water clusters BW1-3 vs BW4). NO_3^- , SiO_4^- , and dissolved O_2 concentrations differed significantly between the four near-bed water clusters. The observation that the seamount is intersecting with different biogeochemical properties and microbial communities, is particularly interesting in regard to benthic organisms. Along these lines, the Schulz Bank seamount provides a platform for sponges and their associated microbial communities to respond to topography-enabled environmental gradients (also a seamount effect sensu lato).

A seamount imprint on sponge-associated microbial communities

The investigated sponge species *S. rosea*, *G. hentscheli* and *L. complicata* were selected for this study as they represent key taxa of the sponge community at Schulz Bank. Microbiomes of these three species clustered clearly apart from each other in ordination space, indicating a dominant host species effect on the associated microbial community structure. *S. rosea* and *L. complicata* showed characteristic microbial signatures of LMA sponges (as defined in Moitinho-Silva et al., 2017) that are being dominated by Proteobacteria. On the contrary, *G. hentscheli* displayed a microbial signature characteristic of HMA sponges with dominant clades such as Chloroflexi and Acidobacteria, which is consistent with previous reports on sponges of the genus *Geodia* (Luter et al., 2017; Radax et al., 2012; Schöttner et al., 2013).

When analysing each sponge species separately, sponge specimen microbiomes differed significantly between each other and depended on the near-bed water clusters to which they belonged. This finding suggests that an environmental signature is also detectable in sponge-associated microbial communities (seamount signature sensu lato). This observation is striking, as sponge-associated microbial communities are considered to be highly stable associations (Cárdenas et al., 2014; Erwin et al., 2012, 2015; Pita et al., 2013; Steinert et al., 2016).

Previous studies have shown that abiotic factors (i.e. depth, geographical location) influence the microbial community structure in shallow-water sponges, but stated that the core community is shaped by the intimate interaction with the sponge host (Lurgi et al., 2019). Interestingly, in our study the major microbial players in terms of abundance, such as Chloroflexi in *G. hentscheli*, show significant enrichment-depletion patterns across the four clusters. Traditionally (shallow water) sponge-associated microbes have been classified into

core, variable and species-specific communities (Schmitt et al., 2012). The present study reveals that for the three investigated deep-sea sponges at Schulz Bank seamount the variable community overlaps with the core community when considering high taxonomy ranks. In addition, network analyses showed both positive and negative correlations between taxa for those increasing with depth as well as those displaying a variable response. We suspect that primary responders to environmental parameters have cascading effects on microbial lineages that are not directly affected by water biogeochemistry.

In this study, silicate, oxygen and nitrate concentrations, as well as depth, differed significantly between the four near-bed water clusters. While SiO_4^- and NO_3^- were positively correlated with depth, O_2 showed a negative relationship. An increase in nutrient concentrations with depth is consistent with our previous expectations (consider, e.g. Bristow et al., 2017) and can be explained by remineralization processes of sinking marine snow within the deep water layers. Decreasing O_2 concentrations from the intermediate water (NwArIW) to the deep water (NwDW) are also consistent with our expectation based on physical oceanography, as water layers more recently oxygenated at the ocean surface at their site of formation typically carry more oxygen (Jeansson et al., 2017). In general, the absolute differences in the concentrations of all three significant environmental parameters were comparably small (especially NO_3^- and O_2). However, as microbial communities were significantly different between the clusters, we posit that the observed biogeochemical differences, albeit small, should be considered drivers of sponge microbial community composition. In support of our hypothesis, the process of denitrification is for example known to be highly sensitive to nanomolar concentrations of O_2 concentrations (Dalsgaard et al., 2013). In addition, a previous study on the sponge *Xestospongia muta* demonstrated that changing NO_x concentrations over depth contribute

to shaping the microbial community composition (Morrow et al., 2016). Furthermore, several studies have noted that depth is an important factor in structuring sponge-associated microbiomes (Indraningrat et al., 2019; Lesser et al., 2020; Steinert et al., 2016).

This is, to our knowledge, the first study that explores the impact of seawater biogeochemistry on deep-sea sponge microbiomes. We have used LEfSe analyses to identify sponge symbiont taxa whose relative abundance varies with depth, this being used as a proxy for selected biogeochemical parameters. We have further used co-occurrence networks to identify positively or negatively co-varying microbial clades. In the following we discuss one representative example of a sponge-associated (and seawater-associated) microbial phylum which was significantly enriched-depleted across the near-bed water clusters: *Chloroflexi* in *G. hentscheli*. Chloroflexi (among several other phyla) relative abundance differed significantly between the four near-bed water clusters for *G. hentscheli* and seawater, showing a positive relation with NO_3^- , SiO_4^- , and depth, and a negative relation with O_2 . Members of the phylum Chloroflexi have been attributed to a relevant role in the degradation of organic matter, particularly in the deep ocean pelagic realm and within HMA sponges (Bayer et al., 2018; Landry et al., 2017). High degradation rates of organic matter are often related with low O_2 and high nutrient concentrations, owing to biogeochemical feedbacks where nutrients enhance oxygen demand by increasing biological production and oxygen consumption during decomposition. Taken together, the differences in relative abundances of Chloroflexi in *G. hentscheli* and seawater could be driven by the NO_3^- , SiO_4^- , and O_2 concentrations in ambient seawater. While sponge microbiomes are generally considered to be highly stable in time and space, we provide a first evidence that small differences in water biogeochemistry may affect sponge microbiome composition. However, no uniform shifts in relative abundances of microbial taxa were observed for *G. hentscheli*,

L. complicata, and *S. rosea* but rather an individual response of each host species related to biogeochemical parameters. One explanation is that biological interactions between the sponge host and its microbiome, or between the microbes themselves, might have masking effects.

Conclusions

We provide insights into the variability of pelagic and benthic (sponge-associated) microbiomes at the Arctic Schulz Bank seamount based on the microbiome analyses of 114 seawater and 36 sponge samples. Interestingly, a seamount signature is detected within the microbial community composition of samples originating as far as 200 m above the seamount summit. Further, our correlation-based results suggest that the biogeochemistry of seawater which varies over depth (NO_3^- , SiO_4^- , and O_2 concentrations) has a detectable but variable influence on the composition of sponge-associated microbiomes. This study provides new perspectives on the influence of seamounts on the microbial diversity in their vicinity. We conclude that the geology, physical oceanography, biogeochemistry, and microbiology of seamounts and similar structures are even more closely linked than currently appreciated.

DATA AVAILABILITY

Sample metadata and biogeochemical data were deposited in the PANGAEA database: <https://doi.pangaea.de/10.1594/PANGAEA.911304> (Busch et al., 2020). Raw sequences were archived in the NCBI Sequence Read Archive under BioProject ID PRJNA600711.

SUPPLEMENTARY MATERIAL

The supplement related to this article is available online at: <https://doi.org/10.5194/bg-17-3471-2020-supplement>.

AUTHOR CONTRIBUTION

KB, FM, BM, and UHe designed the study. KB, UHa, FM, EMR, and HTR participated in sampling. UHa, FM, and KB focused on biogeochemical parameters. HTR conducted the sponge taxonomic analysis. UHe and KB were responsible for the microbial pipeline. AF was involved in sequencing. KB performed the data analysis (bioinformatics and visualizations). KB and UHe wrote the manuscript. BM, FM, EMR, UHa, and HTR reviewed and edited the manuscript.

ACKNOWLEDGEMENTS

This work is dedicated to Hans Tore Rapp, sponge taxonomist, deep-sea explorer, colleague, and friend. Samples included in this study were collected in compliance with the Nagoya Protocol. We thank the crews and scientific parties of RV *G. O. Sars* cruises GS2016109A, GS2017110, and GS2018108 for great technical support while at sea. We are grateful for sponge sampling by Stig Vågenes and his ROV *Ægir 6000* team (UiB), Christine Rooks (UiB) for sampling assistance on board the ship during GS2016109A, and Jasper de Goeij (UvA) for interesting discussions. We further acknowledge Ina Clefsen, Andrea Hethke, Ilona Urbach, and Tonio Hauptmann (Kiel, Germany) for excellent laboratory support with the amplicon pipeline. Corinna Bang (IKMB, Kiel) provided valuable support with the sample sequencing and revised the manuscript before submission. We also thank the two anonymous reviewers for their comments, which helped to improve this paper.

FINANCIAL SUPPORT

This research has been supported by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement no. 679849 (the SponGES project). The article processing charges for this open-access publication were covered by a Research Centre of the Helmholtz Association.

REFERENCES

- Bayer, K., Jahn, M. T., Slaby, B. M., Moitinho-Silva, L., and Hentschel, U.: Marine sponges as *chloroflexi* hot spots: genomic insights and high-resolution visualization of an abundant and diverse symbiotic clade, *mSystems*, 3, 1–19, <https://doi.org/10.1128/mSystems.00150-18>, 2018.
- Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., Alexander, H., Alm, E. J., Arumugam, M., Asnicar, F., Bai, Y., Bisanz, J. E., Bittinger, K., Brejnrod, A., Brislawn, C. J., Brown, C. T., Callahan, B. J., Caraballo-Rodríguez, A. M., Chase, J., Cope, E. K., Da Silva, R., Diener, C., Dorrestein, P. C., Douglas, G. M., Durall, D. M., Duvallet, C., Edwardson, C. F., Ernst, M., Estaki, M., Fouquier, J., Gauglitz, J. M., Gibbons, S. M., Gibson, D. L., Gonzalez, A., Gorlick, K., Guo, J., Hillmann, B., Holmes, S., Holste, H., Huttenhower, C., Huttley, G. A., Janssen, S., Jarmusch, A. K., Jiang, L., Kaehler, B. D., Kang, K. B., Keefe, C. R., Keim, P., Kelley, S. T., Knights, D., Koester, I., Kosciulek, T., Kreps, J., Langille, M. G. I., Lee, J., Ley, R., Liu, Y.-X., Loftfield, E., Lozupone, C., Maher, M., Marotz, C., Martin, B. D., McDonald, D., McIver, L. J., Melnik, A. V., Metcalf, J. L., Morgan, S. C., Morton, J. T., Naimey, A. T., Navas-Molina, J. A., Nothias, L. F., Orchanian, S. B., Pearson, T., Peoples, S. L., Petras, D., Preuss, M. L., Pruesse, E., Rasmussen, L. B., Rivers, A., Robeson II, M. S., Rosenthal, P., Segata, N., Shaffer, M., Shiffer, A., Sinha, R., Song, S. J., Spear, J. R., Swafford, A. D., Thompson, L. R., Torres, P. J., Trinh, P., Tripathi, A., Turnbaugh, P. J., Ul-Hasan, S., van der Hooft, J. J. J., Vargas, F., Vázquez-Baeza, Y., Vogtmann, E., von Hippel, M., Walters, W., Wan, Y., Wang, M., Warren, J., Weber, K. C., Williamson, C. H. D., Willis, A. D., Xu, Z. Z., Zaneveld, J. R., Zhang, Y., Zhu, Q., Knight, R., and Caporaso, J. G.: Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2, *Nat. Biotechnol.*, 37, 852–857, <https://doi.org/10.1038/s41587-019-0209-9>, 2019.
- Borchert, E., Selvin, J., Kiran, S. G., Jackson, S. A., O’Gara, F., and Dobson, A. D. W.: A novel cold active esterase from a deep sea sponge *Stelletta normani* metagenomic library, *Front. Mar. Sci.*, 4, 1–13, <https://doi.org/10.3389/fmars.2017.00287>, 2017.
- Bristow, L. A., Mohr, W., Ahmerkamp, S., and Kuypers, M. M. M.: Nutrients that limit growth in the ocean, *Curr. Biol.*, 27, 474–478, <https://doi.org/10.1016/j.cub.2017.03.030>, 2017.
- Busch, K., Hanz, U., Mienis, F., Müller, B., Franke, A., Roberts, E. M., Rapp, H. T., and Hentschel, U.: Microbial diversity and biogeochemical parameters at Schulz Bank seamount (Arctic Mid-Ocean Ridge) measured in summer 2016, 2017 and 2018, *PANGAEA*, <https://doi.org/10.1594/PANGAEA.911304>, 2020.
- Callahan, B. J., McMurdie, P. J., Rosen, M. J., Han, A. W., Johnson, A. J. A., and Holmes, S. P.: DADA2: High-resolution sample inference from Illumina amplicon data, *Nat. Methods*, 13, 581–583, <https://doi.org/10.1038/nmeth.3869>, 2016.
- Cárdenas, C. A., Bell, J. J., Davy, S. K., Hoggard, M., and Taylor, M. W.: Influence of environmental variation on symbiotic bacterial communities of two temperate sponges, *FEMS Microbiol. Ecol.*, 88, 516–527, <https://doi.org/10.1111/1574-6941.12317>, 2014.
- Cárdenas, P., Rapp, H. T., Schander, C., and Tendal, O. S.: Molecular Taxonomy and Phylogeny of the *Geodiidae* Gray, 1867 (*Porifera*, *Demospongiae*, *Astraphorida*) – combining Phylogenetic and Linnaean Classification, *Zool. Scr.*, 39, 89–106, 2010.
- Chapman, D. C. and Haidvogel, D. B.: Formation of Taylor caps over a tall isolated seamount in a stratified ocean, *Geophys. Astro. Fluid.*, 64, 31–65, <https://doi.org/10.1080/03091929208228084>, 1992.
- Dalsgaard, T., De Brabandere, L., and Hall, P. O. J.: Denitrification in the water column of the central Baltic Sea, *Geochim. Cosmochim. Ac.*, 106, 247–260, <https://doi.org/10.1016/j.gca.2012.12.038>, 2013.
- De Goeij, J. M., Van Den Berg, H., Van Oostveen, M. M., Epping, E. H. G., and Van Duyl, F. C.: Major bulk dissolved organic carbon (DOC) removal by encrusting coral reef cavity sponges, *Mar. Ecol. Prog. Ser.*, 357, 139–151, <https://doi.org/10.3354/meps07403>, 2008.
- De Goeij, J. M., Van Oevelen, D., Vermeij, M. J. A., Osinga, R., Middeburg, J. J., De Goeij, A. F. P. M., and Admiraal, W.: Surviving in a marine desert: The sponge loop retains resources within coral reefs, *Science*, 342, 108–110, <https://doi.org/10.1126/science.1241981>, 2013.
- DeGoeij, J. M., Lesser, M. P., and Pawlik, J. R.: Nutrient Fluxes and Ecological Functions of Coral Reef Sponges in a Changing Ocean, in: *Climate Change, Ocean Acidification and Sponges: Impacts Across Multiple Levels of Organization*, edited by: Carballo, J. L. and Bell, J. J., Springer International Publishing, Cham, Switzerland, 373–410, 2017.
- Easson, C. G. and Thacker, R. W.: Phylogenetic signal in the community structure of host-specific microbiomes of tropical marine sponges, *Front. Microbiol.*, 5, 1–11, <https://doi.org/10.3389/fmicb.2014.00532>, 2014.
- Erwin, P. M., Pita, L., López-Legentil, S., and Turon, X.: Stability of sponge-associated bacteria over large seasonal shifts in temperature and irradiance, *Appl. Environ. Microbiol.*, 78, 7358–7368, <https://doi.org/10.1128/AEM.02035-12>, 2012.
- Erwin, P. M., Coma, R., López-Sendino, P., Serrano, E., and Ribes, M.: Stable symbionts across the HMA-LMA dichotomy: Low seasonal and interannual variation in sponge-associated bacteria from taxonomically diverse hosts, *FEMS Microbiol. Ecol.*, 91, 1–11, <https://doi.org/10.1093/femsec/fiv115>, 2015.
- Fristedt, K.: Sponges from the Atlantic and Arctic Oceans and the Behring Sea. Vega-Expeditionenens Vetenskap, iakttagelser (Nordenskiöld), plates 22-31, 4, 401–471, 1887.
- Gant, S. E., McMurray, S. E., Stubler, A. D., Finelli, C. M., Pawlik, J. R., and Erwin, P. M.: Testing the relationship between microbiome composition and flux of carbon and nutrients in Caribbean coral reef sponges, *Microbiome*, 7, 1–13, <https://doi.org/10.1186/s40168-019-0739-x>, 2019.
- Genin, A. and Boehlert, G. W.: Dynamics of temperature and chlorophyll structures above a seamount: an oceanic experiment, *J. Mar. Res.*, 43, 907–924, <https://doi.org/10.1357/002224085788453868>, 1985.
- Genin, A., Dayton, P. K., Lonsdale, P. F., and Spiess, F. N.: Corals on seamount peaks provide evidence of current acceleration over deep-sea topography, *Nature*, 322, 59–61, <https://doi.org/10.1038/322059a0>, 1986.
- Grant, R. E.: Animal Kingdom, in: *The Cyclopaedia of Anatomy and Physiology*, edited by: Todd, R. B., Sherwood, Gilbert and Piper, London, UK, 107–118, 1836.
- Grasshoff, K., Kremling, K., and Ehrhardt, M.: *Methods of Seawater Analysis*, 3rd edn., WILEY-VCH Verlag GmbH, Weinheim, Germany, 2009.
- Hansen, G. A.: Spongidae. The Norwegian North-Atlantic Expedition 1876–1878, plates I–VII, *Zoology*, 13, 1–26, 1885.
- Helber, S. B., Steinert, G., Wu, Y. C., Rohde, S., Hentschel, U., Muihondo, C. A., and Schupp, P. J.: Sponges from Zanzibar host diverse prokaryotic communities with potential for natural product synthesis, *FEMS Microbiol. Ecol.*, 95, f1z026, <https://doi.org/10.1093/femsec/fiz026>, 2019.
- Helder, W. and de Vries, R. T. P.: An automatic phenol-hypochlorite method for the determination of ammonia in sea- and brackish waters, *Neth. J. Sea Res.*, 13, 154–160, 1979.
- Hentschel, U., Fieseler, L., Wehrl, M., Gernert, C., Steinert, M., Hacker, J., and Horn, M.: Microbial diversity of marine sponges, in: *Sponges (Porifera)*, edited by: Mueller, W. E. G., Springer, Berlin, Heidelberg, Germany, 59–88, 2003.
- Hoer, D. R., Gibson, P. J., Tommerdahl, J. P., Lindquist, N. L., and Martens, C. S.: Consumption of dissolved organic carbon by Caribbean reef sponges, *Limnol. Oceanogr.*, 63, 337–351, <https://doi.org/10.1002/lno.10634>, 2018.
- Indraningrat, A. A. G., Micheller, S., Runderkamp, M., Sauerland, I., Becking, L. E., Smidt, H., and Sipkema, D.: Cultivation of sponge-associated bacteria from *Agelas sventres* and *Xestospongia muta* collected from different depths, *Mar. Drugs*, 17, 1–17, <https://doi.org/10.3390/md17100578>, 2019.
- IUCN: Seamounts Project: An Ecosystem Approach to Management of Seamounts in the Southern Indian Ocean, *Int. Union Conserv. Nat. Nat. Resour.*, Gland, Switzerland, 1, 1–60, 2013.
- Jackson, S. A., Flemer, B., McCann, A., Kennedy, J., Morrissey, J. P., O’Gara, F., and Dobson, A. D. W.: Archaea appear to dominate the microbiome of *Inflatella pellicula* deep sea sponges, *PLoS One*, 8, 1–8, <https://doi.org/10.1371/journal.pone.0084438>, 2013.

- Jeansson, E., Olsen, A., and Jutterström, S.: Arctic Intermediate Water in the Nordic Seas, 1991–2009, *Deep-Sea Res. Pt. I*, 128, 82–97, <https://doi.org/10.1016/j.dsr.2017.08.013>, 2017.
- Kennedy, J., Flemer, B., Jackson, S. A., Morrissey, J. P., O’Gara, F., and Dobson, A. D. W.: Evidence of a putative deep sea specific microbiome in marine sponges, *PLoS One*, 9, 1–13, <https://doi.org/10.1371/journal.pone.0091092>, 2014.
- Kozich, J. J., Westcott, S. L., Baxter, N. T., Highlander, S. K., and Schloss, P. D.: Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the miseq illumina sequencing platform, *Appl. Environ. Microbiol.*, 79, 5112–5120, <https://doi.org/10.1128/AEM.01043-13>, 2013.
- Landry, Z., Swa, B. K., Herndl, G. J., Stepanauskas, R., and Giovannoni, S. J.: SAR202 genomes from the dark ocean predict pathways for the oxidation of recalcitrant dissolved organic matter, *MBio*, 8, 1–19, <https://doi.org/10.1128/mBio.00413-17>, 2017.
- Lavelle, J. W. and Mohn, C.: Motion, commotion, and biophysical connections at deep ocean seamounts, *Oceanography*, 23, 90–103, <https://doi.org/10.5670/oceanog.2010.64>, 2010.
- Lesser, M. P., Mueller, B., Pankey, M. S., Macartney, K. J., Slattery, M., and de Goeij, J. M.: Depth-dependent detritus production in the sponge, *Halisarca caerulea*, *Limnol. Oceanogr.*, 65, 1200–1216, 2020.
- Leys, S. P., Kahn, A. S., Fang, J. K. H., Kutti, T., and Bannister, R. J.: Phagocytosis of microbial symbionts balances the carbon and nitrogen budget for the deep-water boreal sponge *Geodia barretti*, *Limnol. Oceanogr.*, 63, 187–202, <https://doi.org/10.1002/lno.10623>, 2018.
- Lurgi, M., Thomas, T., Wemheuer, B., Webster, N. S., and Montoya, J. M.: Modularity and predicted functions of the global sponge-microbiome network, *Nat. Commun.*, 10, 1–12, <https://doi.org/10.1038/s41467-019-08925-4>, 2019.
- Luter, H. M., Bannister, R. J., Whalan, S., Kutti, T., Pineda, M.-C., and Webster, N. S.: Microbiome analysis of a disease affecting the deep-sea sponge *Geodia barretti*, *FEMS Microbiol. Ecol.*, 93, 1–6, <https://doi.org/10.1093/femsec/fix074>, 2017.
- Maldonado, M., Ribes, M., and van Duyl, F. C.: Nutrient fluxes through sponges. Biology, budgets, and ecological implications, *Adv. Mar. Biol.*, 62, 113–182, <https://doi.org/10.1016/B978-0-12-394283-8.00003-5>, 2012.
- Maldonado, M., Aguilar, R., Bannister, R. J., Bell, J. J., Conway, K. W., Dayton, P. K., Diaz, C., Gutt, J., Kelly, M., Kenchington, E. L. R., Leys, S. P., Pomponi, S. A., Rapp, H. T., Rützler, K., Tendal, O. S., Vacelet, J., and Young, C. M.: Sponge grounds as key marine habitats: a synthetic review of types, structure, functional roles, and conservation concerns, in: *Marine animal forests*, edited by: Rossi, S., Bramanti, L., Gori, A., and Orejas, C., Springer International Publishing Switzerland, Cham, Switzerland, 2016.
- Maldonado, M., López-Acosta, M., Sitjà, C., García-Puig, M., Galobart, C., Ercilla, G., and Leynaert, A.: Sponge skeletons as an important sink of silicon in the global oceans, *Nat. Geosci.*, 12, 815–822, <https://doi.org/10.1038/s41561-019-0430-7>, 2019.
- McMurray, S. E., Johnson, Z. I., Hunt, D. E., Pawlik, J. R., and Finelli, C. M.: Selective feeding by the giant barrel sponge enhances foraging efficiency, *Limnol. Oceanogr.*, 61, 1271–1286, <https://doi.org/10.1002/lno.10287>, 2016.
- Menard, H. W.: *Marine Geology of the Pacific*, McGraw-Hill, New York, USA, 271 pp., 1964.
- Meyer, H. K., Roberts, E. M., Rapp, H. T., and Davies, A. J.: Spatial patterns of arctic sponge ground fauna and demersal fish are detectable in autonomous underwater vehicle (AUV) imagery, *Deep-Sea Res. Pt. I*, 153, 103137, <https://doi.org/10.1016/j.dsr.2019.103137>, 2019.
- Moitinho-Silva, L., Bayer, K., Cannistraci, C. V., Giles, E. C., Ryu, T., Seridi, L., Ravasi, T., and Hentschel, U.: Specificity and transcriptional activity of microbiota associated with low and high microbial abundance sponges from the Red Sea, *Mol. Ecol.*, 23, 1348–1363, <https://doi.org/10.1111/mec.12365>, 2014.
- Moitinho-Silva, L., Steinert, G., Nielsen, S., Hardoim, C. C. P., Wu, Y. C., McCormack, G. P., López-Legentil, S., Marchant, R., Webster, N., Thomas, T., and Hentschel, U.: Predicting the HMALMA status in marine sponges by machine learning, *Front. Microbiol.*, 8, 1–14, <https://doi.org/10.3389/fmicb.2017.00752>, 2017.
- Morato, T., Hoyle, S. D., Allain, V., and Nicol, S. J.: Seamounts are hotspots of pelagic biodiversity in the open ocean, *P. Natl. Acad. Sci. USA*, 107, 9707–9711, <https://doi.org/10.1073/pnas.0910290107>, 2010.
- Morrow, K. M., Fiore, C. L., and Lesser, M. P.: Environmental drivers of microbial community shifts in the giant barrel sponge, *Xestospongia muta*, over a shallow to mesophotic depth gradient, *Environ. Microbiol.*, 18, 2025–2038, <https://doi.org/10.1111/1462-2920.13226>, 2016.
- Mueller, B., De Goeij, J. M., Vermeij, M. J. A., Mulders, Y., Van Der Ent, E., Ribes, M., and Van Duyl, F. C.: Natural diet of coral-excavating sponges consists mainly of dissolved organic carbon (DOC), *PLoS One*, 9, 1–7, <https://doi.org/10.1371/journal.pone.0090152>, 2014.
- Murphy, J. and Riley, J. P.: A modified single solution method for the determination of phosphate in natural waters, *Anal. Chim. Act.*, 27, 31–36, 1962.
- Naim, M. A., Morillo, J. A., Sørensen, S. J., Waleed, A. A. S., Smidt, H., and Sipkema, D.: Host-specific microbial communities in three sympatric North Sea sponges, *FEMS Microbiol. Ecol.*, 90, 390–403, <https://doi.org/10.1111/1574-6941.12400>, 2014.
- Pile, A. J., Patterson, M. R., Savarese, M., Chernykh, V. I., and Fialkov, V. A.: Trophic effects of sponge feeding within Lake Baikal’s littoral zone. Sponge abundance, diet, feeding efficiency, and carbon flux, *Limnol. Oceanogr.*, 42, 178–184, 1997.
- Pita, L., Turon, X., López-Legentil, S., and Erwin, P. M.: Host rules: Spatial stability of bacterial communities associated with marine sponges (*Ircinia* spp.) in the western Mediterranean sea, *FEMS Microbiol. Ecol.*, 86, 268–276, <https://doi.org/10.1111/1574-6941.12159>, 2013.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schwaer, T., Yarza, P., Peplies, J., and Glöckner, F. O.: The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools, *Nucleic Acids Res.*, 41, 590–596, <https://doi.org/10.1093/nar/gks1219>, 2013.
- Radax, R., Rattei, T., Lanzen, A., Bayer, C., Rapp, H. T., Urlich, T., and Schleper, C.: Metatranscriptomics of the marine sponge *Geodia barretti*: Tackling phylogeny and function of its microbial community, *Environ. Microbiol.*, 14, 1308–1324, <https://doi.org/10.1111/j.1462-2920.2012.02714.x>, 2012.
- R Development Core Team: R: A language and environment for statistical computing, available at: <http://www.r-project.org> (last access: 10 October 2013), 2008.
- Reiswig, H. M.: Particle feeding in natural populations of three marine demosponges, *Biol. Bull.*, 141, 568–591, 1971.
- Reveillaud, J., Maignien, L., Eren, M. A., Huber, J. A., Apprill, A., Sogin, M. L., and Vanreusel, A.: Host-specificity among abundant and rare taxa in the sponge microbiome, *ISME J.*, 8, 1198–1209, <https://doi.org/10.1038/ismej.2013.227>, 2014.
- Ribes, M., Coma, R., and Gili, J.: Natural diet and grazing rate of the temperate sponge *Dysidea avara* (Demospongiae, Dendroceratida) throughout an annual cycle, *Mar. Ecol. Prog. Ser.*, 176, 179–190, 1999.
- Rix, L., De Goeij, J. M., Mueller, C. E., Struck, U., Middelburg, J. J., Van Duyl, F. C., Al-Horani, F. A., Wild, C., Naumann, M. S., and Van Oevelen, D.: Coral mucus fuels the sponge loop in warm- and cold-water coral reef ecosystems, *Sci. Rep.*, 6, 1–11, <https://doi.org/10.1038/srep18715>, 2016a.
- Rix, L., Goeij, J. M. De, Oevelen, D. Van, Struck, U., Al-horani, F. A., Wild, C., and Naumann, M. S.: Differential recycling of coral and algal dissolved organic matter via the sponge loop, *Funct. Ecol.*, 31, 1–12, <https://doi.org/10.1111/1365-2435.12758>, 2016b.
- Roberts, E. M., Mienis, F., Rapp, H. T., Hanz, U., Meyer, H. K., and Davies, A. J.: Oceanographic setting and short-timescale environmental variability at an Arctic seamount sponge ground, *Deep Sea Res. Pt. I*, 98–113, <https://doi.org/10.1016/j.dsr.2018.06.007>, 2018.
- Rogers, A. D.: The Biology of Seamounts: 25 Years on, *Adv. Mar. Biol.*, 79, 137–224, <https://doi.org/10.1016/bs.amb.2018.06.001>, 2018.
- Schmitt, S., Tsai, P., Bell, J., Fromont, J., Ilan, M., Lindquist, N., Perez, T., Rodrigo, A., Schupp, P. J., Vacelet, J., Webster, N., Hentschel, U., and Taylor, M. W.: Assessing the complex sponge microbiota: core, variable and species-specific bacterial communities in marine sponges, *ISME J.*, 6, 564–576, <https://doi.org/10.1038/ismej.2011.116>, 2012.

- Schöttner, S., Hoffmann, F., Cárdenas, P., Rapp, H. T., Boetius, A., and Ramette, A.: Relationships between host phylogeny, host type and bacterial community diversity in cold-water coral reef sponges, *PLoS One*, 8, 1–11, <https://doi.org/10.1371/journal.pone.0055505>, 2013.
- Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., and Huttenhower, C.: Metagenomic biomarker discovery and explanation, *Genome Biol.*, 12, R60, <https://doi.org/10.1186/gb-2011-12-6-r60>, 2011.
- Smith, D. K. and Cann, J. R.: The role of seamount volcanism in crustal construction at the Mid-Atlantic Ridge (24°–30° N), *J. Geophys. Res.-Sol. Ea.*, 97, 1645–1658, 1992.
- Staudigel, H., Koppers, A. A. P., William Lavelle, J., Pitcher, T. J., and Shank, T. M.: Defining the word “seamount”, *Oceanography*, 23, 20–21, 2010.
- Steinert, G., Taylor, M. W., Deines, P., Simister, R. L., De Voogd, N. J., Hoggard, M., and Schupp, P. J.: In four shallow and mesophotic tropical reef sponges from Guam the microbial community largely depends on host identity, *PeerJ*, 2016, 1–25, <https://doi.org/10.7717/peerj.1936>, 2016.
- Steinert, G., Rohde, S., Janussen, D., Blaurock, C., and Schupp, P. J.: Host-specific assembly of sponge-associated prokaryotes at high taxonomic ranks, *Sci. Rep.*, 7, 1–9, <https://doi.org/10.1038/s41598-017-02656-6>, 2017.
- Steinert, G., Wemheuer, B., Janussen, D., Erpenbeck, D., Daniel, R., Simon, M., Brinkhoff, T., and Schupp, P. J.: Prokaryotic diversity and community patterns in antarctic continental shelf sponges, *Front. Mar. Sci.*, 6, 1–15, <https://doi.org/10.3389/fmars.2019.00297>, 2019.
- Steinert, G., Busch, K., Bayer, K., Kodami, S., Arbizu, P. M., Kelly, M., Mills, S., Erpenbeck, D., Dohrmann, M., Wörheide, G., Hentschel, U., and Schupp, P. J.: Compositional and quantitative insights into bacterial and archaeal communities of South Pacific deep-sea sponges (Demospongiae and Hexactinellida), *Front. Microbiol.*, 11, 716, <https://doi.org/10.3389/fmicb.2020.00716>, 2020.
- Stoll, M. H. C., Bakker, K., Nobbe, G. H., and Haese, R. R.: Continuous-flow analysis of dissolved inorganic carbon content in seawater, *Anal. Chem.*, 73, 4111–4116, <https://doi.org/10.1021/ac010303r>, 2001.
- Strickland, J. and Parsons, T.: A practical handbook of seawater analysis, *Fish. Res. Board Canada*, 167, 65–70, 1972.
- Taylor, M. W., Radax, R., Steger, D., and Wagner, M.: Sponge-Associated Microorganisms: Evolution, Ecology, and Biotechnological Potential, *Microbiol. Mol. Biol. Rev.*, 71, 295–347, <https://doi.org/10.1128/MMBR.00040-06>, 2007.
- Thomas, T., Moitinho-Silva, L., Lurgi, M., Björk, J. R., Easson, C., Astudillo-García, C., Olson, J. B., Erwin, P. M., López-Legentil, S., Luter, H., Chaves-Fonnegra, A., Costa, R., Schupp, P. J., Steindler, L., Erpenbeck, D., Gilbert, J., Knight, R., Ackermann, G., Victor Lopez, J., Taylor, M. W., Thacker, R. W., Montoya, J. M., Hentschel, U., and Webster, N. S.: Diversity, structure and convergent evolution of the global sponge microbiome, *Nat. Commun.*, 7, 1–12, <https://doi.org/10.1038/ncomms11870>, 2016.
- Tian, R. M., Sun, J., Cai, L., Zhang, W. P., Zhou, G. W., Qiu, J. W., and Qian, P. Y.: The deep-sea glass sponge *Lophophysema eversa* harbours potential symbionts responsible for the nutrient conversions of carbon, nitrogen and sulfur, *Environ. Microbiol.*, 18, 2481–2494, <https://doi.org/10.1111/1462-2920.13161>, 2016.
- Van Haren, H., Hanz, U., de Stigter, H., Mienis, F., and Duineveld, G.: Internal wave turbulence at a biologically rich Mid-Atlantic seamount, *PLoS One*, 12, 1–16, <https://doi.org/10.1371/journal.pone.0189720>, 2017.
- Van Oevelen, D., Mueller, C. E., Lundälv, T., Van Duyl, F. C., De Goeij, J. M., and Middelburg, J. J.: Niche overlap between a cold-water coral and an associated sponge for isotopically enriched particulate food sources, *PLoS One*, 13, 1–16, <https://doi.org/10.1371/journal.pone.0194659>, 2018.
- Weisz, J. B., Lindquist, N., and Martens, C. S.: Do associated microbial abundances impact marine demosponge pumping rates and tissue densities?, *Oecologia*, 155, 367–376, <https://doi.org/10.1007/s00442-007-0910-0>, 2008.
- Wessel, P., Sandwell, D. T., and Kim, S.-S.: Seamount Census, *Oceanography*, 23, 24–33, 2010.

GRAPHICAL SUMMARY

On giant shoulders: How a **seamount** affects the microbial **community composition** of **seawater** and **sponges**

Busch Hanz Mienis Müller Franke Roberts Rapp Hentschel

illustrated by K.Busch



Seamounts are globally abundant submarine structures, which offer great potential to study the **impacts and interactions of environmental gradients at a single geographic location.**

We observed that **a seamount can have an imprint on the microbial community structure** of the watercolumn (reaching at least 200m above the seamount summit). Further, near-bed water and sponge-associated microbial communities showed signs of an environmental imprint along the seamount topography.

We conclude that the **geology, physical oceanography, biogeochemistry, and microbiology of seamounts** are even more **closely linked** than currently appreciated.

Chapter 7

How to sum up individuals? Nestedness and connectivity of the global deep-sea sponge microbiome

How to sum up individuals?

Nestedness and connectivity of the global deep-sea sponge microbiome

Kathrin Busch^{1*}, Beate Slaby¹, Lindsay Beazley², Ina Clefsen¹, Ana Colaço³, Javier Cristobo⁴, Luisa Federwisch⁵, Andrea Hethke¹, Ellen Kenchington², Furu Mienis⁶, Ana Riesgo⁷, Pilar Ríos⁴, Lucía Pita¹, Peter Schupp⁸, Joana Xavier⁹, Hans Tore Rapp⁹, Ute Hentschel^{1,10*}

¹GEOMAR Helmholtz Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany, ²Department of Fisheries and Oceans, Bedford Institute of Oceanography, Dartmouth, Nova Scotia, Canada, ³IMAR-Laboratório Marítimo da Guia, Faculdade de Ciências de Lisboa, Estrada do Guincho, 2750 Cascais, Portugal, ⁴Instituto Español de Oceanografía. Centro Oceanográfico de Gijón, C/ Príncipe de Asturias 70 bis, 33212 Gijón, Asturias, Spain, ⁵Alfred-Wegener-Institut Helmholtz-Zentrum für Polar-und Meeresforschung, Am Handelshafen 12, 27570, Bremerhaven, Germany, ⁶NIOZ Royal Netherlands Institute for Sea Research, and Utrecht University, 1790 AB Den Burg, Texel, the Netherlands, ⁷Universidad de Alcalá, Departamento de Ciencias de la Vida, Apdo. 20, 28805, Alcalá de Henares, Spain, ⁸Institute for Chemistry and Biology of the Marine Environment (ICBM), University of Oldenburg, Oldenburg, Germany, ⁹University of Bergen, K.G. Jebsen Centre for Deep Sea Research, P.O. Box 7803, 5020 Bergen, Norway, ¹⁰Christian-Albrechts University of Kiel, Christian-Albrechts-Platz 4, 24118 Kiel, Germany. *Corresponding authors.

This project investigates the biodiversity of microbial consortia associated with deep-sea sponges. 20 deep-sea expeditions were conducted to collect data for a next-generation biodiversity assessment of sponge-microbiomes that were integrated with extensive oceanographic and ecological metadata. Our study aims to determine the main drivers of microbiome variability within deep-sea sponges. The generated 16S amplicon data reveal that sponges represent diverse reservoirs of microbes in the dark ocean. We find that different host species and even individuals harbor their own, characteristic microbiomes which are different from seawater and sediment microbiomes originating from the same sampling location. We propose that the microbial diversity within deep-sea sponges is driven additionally by environmental factors such as temperature, salinity, nutrients/oxygen, and depth. With 1077 sponge individuals analysed in the present study, we are spanning various spatial scales reaching from individual sponge holobionts to an integrated global assessment. Interestingly, our results reveal a high individuality in microbial community composition. Individual sponge grounds harbor rather dissimilar microbial communities, and the pool of microbial features which occur in less than ten samples is exceptional large (>80 % up to 96 % of total number of ASV features). Despite high variability in beta diversity, alpha diversity is remarkably constant within each samplotype (but different between samptypes), and allowed establishment of a trait-based predictive framework which links sponge morphology with the sponge-associated microbial community structure.

MAIN TEXT

Individuality is the foundation of diversity. One key prerequisite in order to define individuality is that only definite entities can show individualism. But setting the boundaries for single units is no trivial task. When taking us humans as example, recent developments in microbiology have led to a challenge in what we define as our *self*. The term *metaorganism* has been established to pay tribute to the fact that we live in close association with diverse microbial communities, on which we rely for key physiological functions such as e.g. digestion (Bosch and McFall-Ngai, 2011; McFall-Ngai et al., 2013). Over the years it has become clear that not only humans, but the vast majority of animals on Earth are metaorganisms. From an ecological point of view every animal host can thus be considered an individual ecosystem. Similar to a Russian doll (Pita et al., 2018), each ecosystem is in turn nested into a larger ecological context. It is fundamental for a single ecological unit to keep separate from its immediate surrounding and to maintain barriers for disconnectedness. Sponges are basal organisms which master the challenge to persist as a distinct metaorganisms. They are in fact one of the most complex metaorganisms and can harbour a high microbial diversity on high taxonomic ranks (Thomas et al., 2016). This is remarkable as (i) sponges are effective filter feeding animals, meaning that they have on average 24 000 L kg⁻¹ of surrounding seawater (microbes) flowing through their body each day (Hentschel et al., 2003), and (ii) sponges thrive in the ocean which is generally considered the largest interconnected habitat on Earth (Ramirez-Llodra et al., 2011). Sponges in the deep-sea occur frequently in dense patches, referred to as *sponge grounds* (Klitgaard and Tendal, 2004; Xavier et al., 2015; Maldonado et al., 2016). These habitats are considered hotspots of biodiversity and ecosystem services (Maldonado et al., 2016). These aspects make sponges

ideal candidates to study the nestedness of microbial communities across multiple levels. For this study we considered three main ecological layers around sponge-associated microbiomes, namely the sponge host, sponge grounds (i.e. the sponge population level), and the environment. In order to assess nestedness and connectivity of the global deep-sea sponge microbiome, we performed 20 deep-sea expeditions (Figure 1A; Table S1). During these ship expeditions 1077 sponge-associated microbiomes were sampled mainly by remotely operated vehicles (Table S2) along with 355 seawater microbiomes and 114 sediment microbiomes from 52 sponge ground locations. In this study we observed that deep-sea sponges harboured microbiomes that were significantly different from seawater and sediment microbiomes (Figure 1C, Table S5). This observation is in accordance with previous studies, which highlighted the difference between host-associated (including sponge-associated) and free-living microbiomes (Thomas et al., 2016; Thompson et al., 2017). We conclude that also deep-sea sponges are indeed a separate ecological unit from their surrounding environment in terms of microbial diversity. We observed 29 more microbial phyla in the analysed deep-sea samples than have been previously recorded from shallow waters (in total 71 microbial phyla in deep-sea sponges vs the 42 recorded phyla from shallow waters, Thomas et al., 2016). The five most abundant microbial phyla were *Proteobacteria*, *Chloroflexi*, *Acidobacteriota*, *Bacteroidota*, and *Actinobacteriota* (Figure 1C). These taxa showed significant enrichments between sampletypes (Figure S6). Interestingly we further observed significant differences in the microbiomes of three sponge groups (Figure 1C, Table S5). With the help of machine learning (Figure S1, Figure S2) and

taxonomic identification of the 169 sampled sponge species (**Table S3**), we identified these three groups as High Microbial Abundance (HMA) sponges, Low Microbial Abundance demosponges (LMA_demo), and Low Microbial Abundance glass sponges (LMA_glass). In addition to the distinct differences in microbial beta diversity, we also observed significant differences in microbial alpha diversity between these sponge groups (but note that both LMA sponge types had an overall similar alpha diversity) and the different sampletypes (Figure 1D, Table S6). Overall these results

make clear that sponges represent a relevant resource of microbial diversity in the deep-sea. This is particularly interesting as abundances of free-living microbial taxa are known to decrease with depth (Karner et al., 2001; Busch et al., 2017); and biodiversity assessments in the deep ocean are rare (Snelgrove, 1999; Danovaro, 2009; Ramirez-Llodra et al., 2010; Danovaro et al., 2014; Corinaldesi, 2015). Our results further show that deeply rooted differences in microbial community composition exist between three sponge types.

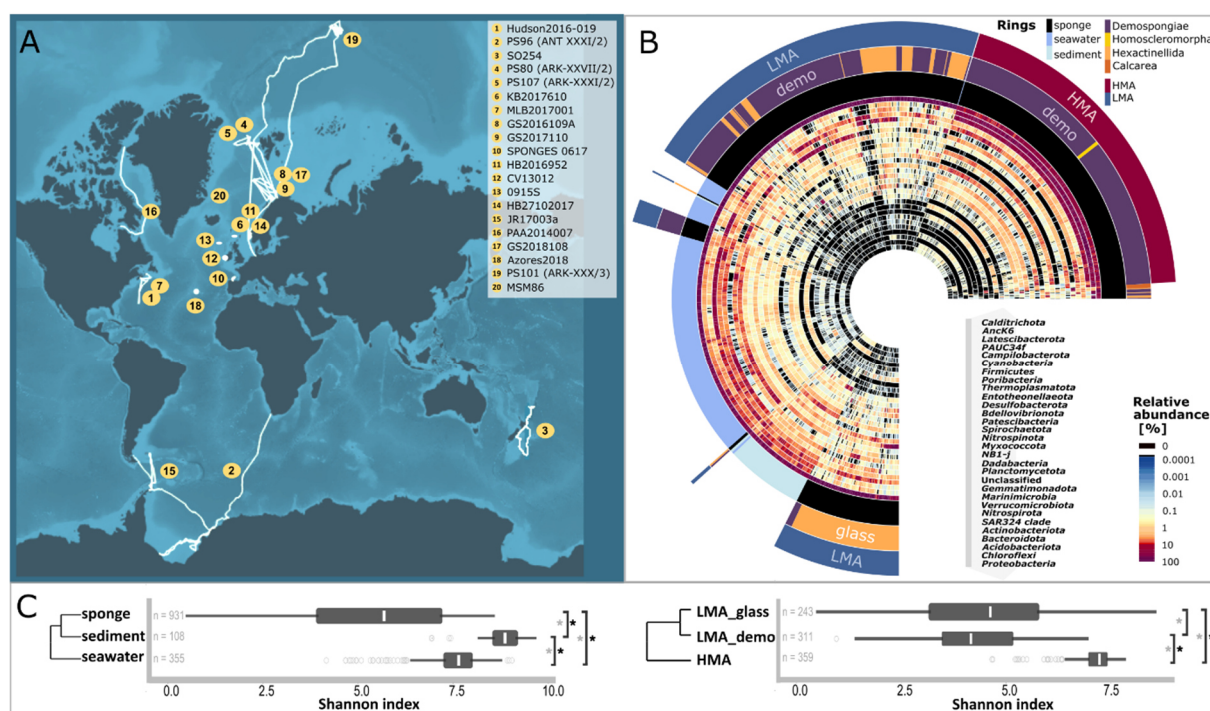


FIGURE 1 | A) Cruise tracks of the 20 conducted deep-sea expeditions. **B)** Clustering dendrogram based on weighted UniFrac distances (microbial beta-diversity). A heatmap showing relative abundances of the 30 most prominent microbial phyla is included. Phyla are labelled right next to the heatmap and sorted after decreasing abundance from outside to inside. Colored rings around the heatmap indicate sampletypes, sponge classes, and the output of the machine learning approach to predict sponge status. **C)** Shannon indices of sampletypes and sponge types (microbial alpha diversity). Sample numbers are indicated (grey text). Black asterisks indicate if output of pairwise Dunn's test was significant (i.e. difference in alpha-diversity). Grey asterisks indicate if output of pair-wise PERMANOVA was significant (i.e. difference in beta-diversity). A clustering dendrogram (based on weighted UniFrac distances) is added left to the plot to indicate similarity between groups.

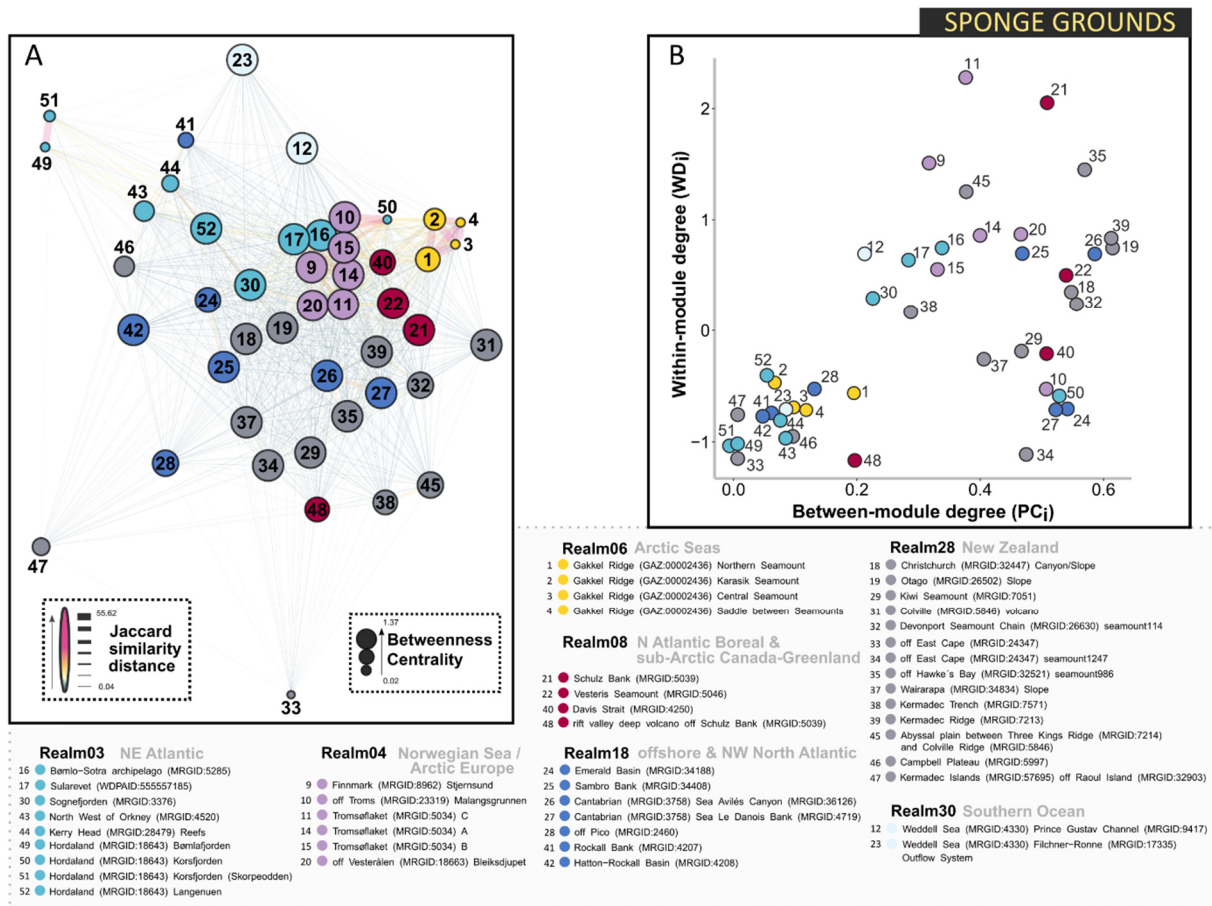


FIGURE 2 | A) Similarity network of sponge grounds based on Jaccard distances (presence-absence of microbial taxa). Size of points is according to node betweenness centrality. Numbers inside nodes encode for locations (compare text right next to plot). Coloring of nodes was done according to which realm (based on Costello et al., 2017) the respective sponge ground belongs. **B)** Within- and between-module degrees of different sponge grounds as derived from a bipartite network between microbial taxa and sponge grounds (location).

To assess connectivity between sponge grounds, we established a similarity network between locations (**Figure 2A**), and a bipartite network between locations and microbial amplicon single nucleotide variant (ASV) occurrences (**Figure 2B**). We observed an overall low within- and between- module degree suggesting that connectivity in microbial community composition within and between sponge grounds was generally low. Overall, sponge microbiomes occurring from the same ecological realm were more similar to each other than to more distant grounds

(**Figure 2A**). Nevertheless, microbial community composition was significantly different between the vast majority of sponge grounds, while alpha diversity was constant in almost all cases for each sponge type (**Figure S5, Table S16, Table S17**). We conclude that the deep-sea is a heterogeneous place. As sponge grounds are wide-spread, but vulnerable ecosystems in the deep ocean, we propose that also the associated microbial diversity should be considered when discussing about conservation strategies.

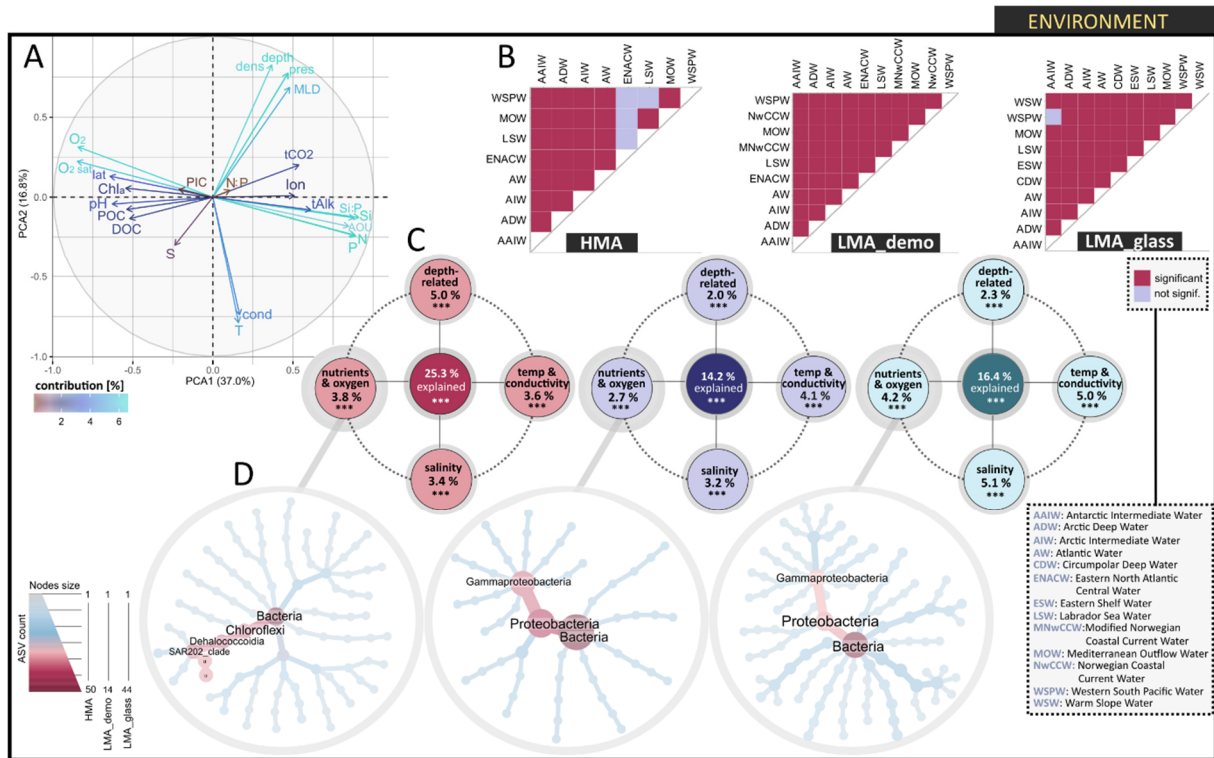


FIGURE 3 | A) Principle component analysis (PCA) of 24 environmental parameters. Coloring is added according to each parameter's contribution. **B)** Color coded statistical output of pair-wise PERMANOVAs (based on weighted UniFrac distances) between watermasses. Full names of watermass abbreviations are given in the lower right corner. **C)** Final variation partitioning models for HMA sponges, LMA_demo sponges and LMA_glass sponges. Asterisks indicate significance of models as assessed by permutations. Percentages indicate fraction of microbial variability that is explained by the four parameter groups individually, and together (center of each sub-plot). Note that only those microbial taxa which occurred in more than 10 samples of each sponge type were considered for this analysis. **D)** Heat trees of microbial community compositions occurring in the nutrient/oxygen modules of HMA sponges, LMA_demo sponges, and LMA_glass sponges. According modules were derived from weighted gene correlation networks. Only those taxa with a modularity >0.8 are shown, as these taxa are considered to show strongest connections to other taxa in the network as well as strongest correlation to nutrient and oxygen concentrations. Colors and node sizes in the heat trees indicate abundance of respective microbial taxa. Unclassified taxa are abbreviated with 'u', only the most abundant taxa are labelled.

In this study we gathered 24 environmental parameters (Figure 3A, Table S11) and classified 14 watermasses from which sponges were sampled (Figure S3, Table S10). Sponge-associated microbiomes showed in almost all cases a significant difference in microbial community composition between watermasses (Figure 3B, Table S14), while microbial alpha diversity stayed constant across watermasses for all sponge types (Figure S5, Table S15). Depth-related parameters, temperature related parameters, salinity, as well as nutrient and oxygen concentrations turned out to be the main environmental drivers of microbial variability in

deep-sea sponges (Figure 3C). Correlations between microbial community compositions (weighted UniFrac distances) and each specific environmental parameter behind these four categories (euclidean distances) were statistically significant (Table S18). In accordance to our observations, temperature, salinity, and depth have previously been determined as the most relevant drivers of host-associated and free-living microbial communities (Lozupone and Knight, 2007; Sunagawa et al., 2015; Thompson et al., 2017; Lurgi et al., 2019), although these studies did not evaluate a comparably high number of parameters as we did for this study.

Interestingly we observed no taxonomic consistency in those microbial taxa which correlated strongest with variations in nutrient and oxygen concentrations (Figure 3D). Different microbial taxa responded strongest to environmental gradients for different sponge types (although the two LMA sponge types were more similar to each other). Those microbial taxa which responded strongest to environmental gradients were generally also those taxa which were the most prominent

members of the microbial community of respective sponge types. As a similar taxonomic pattern was not only observed for those microbial modules which were correlated with nutrient and oxygen concentrations, but also for all other main environmental drivers, we conclude that modularity of deep-sea sponge microbiomes may serve as a mean of taxonomic stability within sponge types.

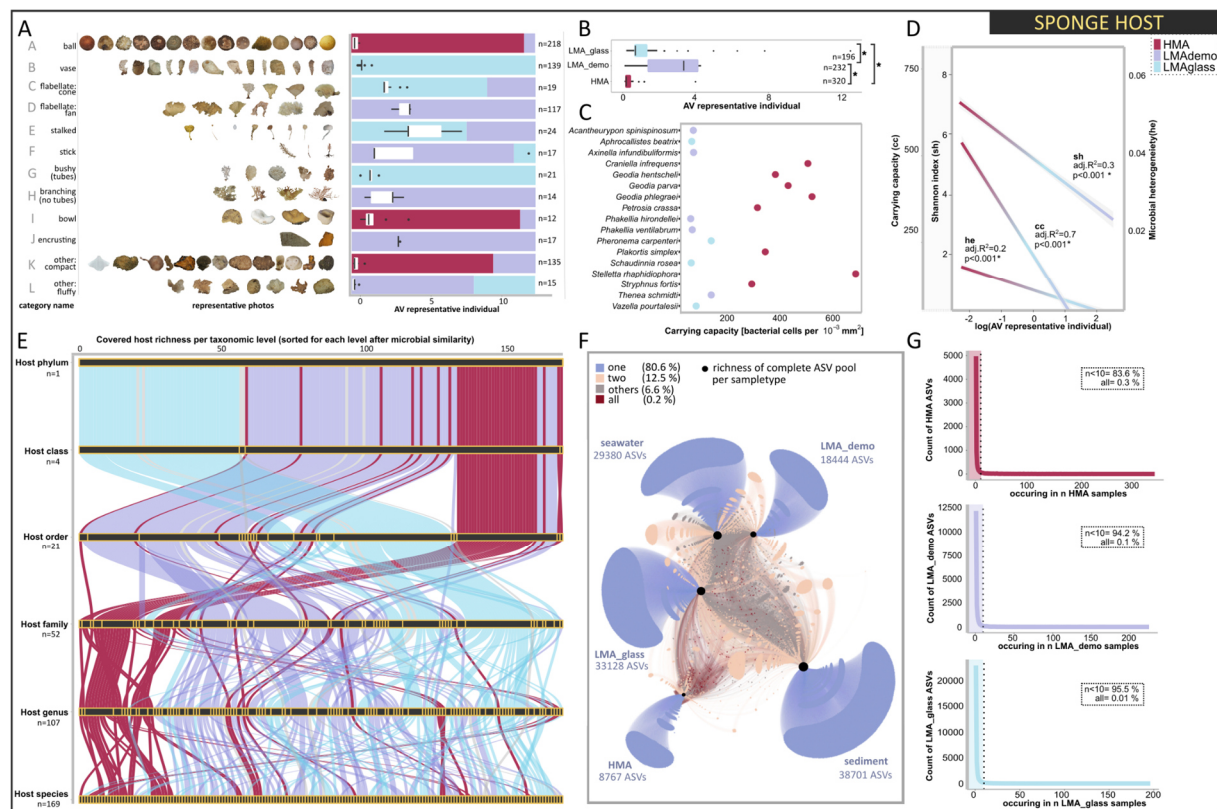


FIGURE 4 | A) Host morphological categories (A-L) and representative photos of sponges belonging to each category. Boxplots right next to the photos indicate area-volume (AV) ratios of each category. Stacked bar charts in the back of the boxplots show percentages how many samples per category belong to HMA, LMA_demo, or LMA_glass sponges (no axis shown). Sample numbers are indicated (n). **B)** Boxplot visualising differences in the area-volume ratios between different sponge types. Black asterisks indicate if output of statistical pairwise tests (Dunn's tests) was significant. **C)** Determined carrying capacity for the 17 microscopically analysed sponge species. Consistent color coding was applied for HMA sponges, LMA_demo sponges, and LMA_glass sponges. **D)** Regressions between carrying capacity (cc), richness (Shannon; sh), microbial heterogeneity (he), and (log-transformed) area-volume ratios. Key outputs of regressions are indicated next to each line. Color gradient refers to dominance of sponge type data points at the respective area of the lines. **E)** Alluvial diagram showing clustering based on microbial community similarity (weighted UniFrac distances) at different host taxonomic levels. Yellow lines mark boundaries of sponges belonging to the same taxonomic group. Colors are added according to sponge type. **F)** Bipartite network between microbial taxa and sample types. Total numbers of ASVs occurring in each sample type are indicated (blue text). Edge and node colors are referring to prevalence of ASVs. Blue marks ASVs occurring in only one sample type, beige marks ASVs shared between two sample types, red marks ASVs occurring in all sample types. **G)** Counts of ASVs indicating in how many samples a specific number of ASVs occurs for each sponge type. This plot can be seen as a visualisation of microbiome individuality.

By applying a trait-based approach, we established 12 morphological sponge categories (**Figure 4A**) and developed computational models to calculate the area-volume ratios for each category (**Table S8**, **Table S9**). Interestingly, HMA sponges, LMA_demo sponges, and LMA_glass sponges each showed preferences for specific morphological categories. As morphological categories differed in their area-volume ratio, the area-volume differed significantly between HMA sponges, LMA_demo sponges, and LMA_glass sponges (**Figure 4B**, **Table S4**). We observed a consistent difference in the carrying capacity between HMA sponges and LMA sponges (**Figure 4C**), and a higher richness and higher microbial heterogeneity for low area-volume ratios than for large area-volume ratios (**Figure 4D**). We conclude overall that sponge morphological traits correspond to the harboured microbial abundance. Furthermore, we observed statistically significant differences in the microbial community composition between morphological categories for each sponge type, but not in their microbial richness (**Figure S5**, **Table S12**, **Table S13**). In addition to the just described trait-based approach, we assessed how sponge taxonomic relatedness corresponds to similarities in microbial community composition. We observed that sponge taxonomy explained microbial betadiversity only to a limited degree (i.e. not well below host order level; **Figure 4E**). At the same time we observed that the pool of those ASVs, which are specific to only one samplotype, was extremely high (> 80 % of total number of ASVs; **Figure 4F**), while only a minor fraction of ASVs was shared between samplotypes (**Table S7**). Along similar lines, the pool of ASVs which only occurred in less than ten samples of the same sponge type was extremely large (>80 % up to 96 % of total number of ASVs; **Figure 4G**). The same holds true for seawater and sediment microbiomes (**Figure S4**). We conclude that a high individuality in microbial community

composition exists for different samplotypes and for different sponge types.

Overall, our results suggest that variations in microbial community composition can be explained to a minor part by variations in host parameters, on the sponge population level (i.e. sponge grounds) or by the environment. As the largest part of the variation is attributed to the individuality of each sample, every single microbiome should be considered a special edition. This is striking, and challenges the commonly applied scientific approach, in which the researcher reaches for a sample number as high as possible to deduce general patterns and minimise effects of individuals. In our case individuality has still the most pronounced imprint on microbial beta diversity, despite working on a global scale (but not on alpha diversity which shows constant patterns). A high individuality leads to the circumstance that the microbial core community (i.e. the predictable part of the community) gets smaller, the more samples are included in the analysis. Patterns of microbial community composition derived from each study may therefore strongly depend on the chosen scale. As general perception of individuality decreases with decreasing size of the organism, bacteria are most commonly referred to as populations or communities in ecology, rather than as individual cells. Their individuality vanishes in the mass. Our results suggest however that their individuality is even detectable on the global scale and should be considered in assessments of biodiversity.

METHODS

When compiling and processing this dataset, we aimed to reduce the extent of technical variation to a bare minimum. For doing so we established strict standard operating procedures. The wet-lab standard operating procedure has been archived at protocols.io. The computational script of our core bioinformatic pipeline (incl. visualisations of performed technical validations) can be found

on github. The applied analysis workflow is illustrated in detail in **Method S1**.

Ship cruises, 16S amplicon sequencing, and statistical analyses (described in brief)

The presented data originated from 20 deep-sea expeditions (which are listed in **Table S1**). Sampling gear is listed in **Table S2** and more detailed descriptions of sampling and lab work can be extracted from protocols.io. In brief, sponge samples originated from between 6 m and 4833 m depth. The average sampling depth across all samples was 795 m. We sampled 1077 sponge individuals, 355 seawater samples and 114 sediment samples. As not all samples passed our quality standards during processing, we ended up with a dataset spanning 1394 samples in total (931 sponges, 355 seawater samples, and 108 sediment samples). Sampling occurred at 52 sponge ground locations during 271 sampling events. Sample metadata were deposited in the PANGAEA database.

For 16S amplicon sequencing, in brief, the V3-V4 variable regions of the 16S rRNA gene were amplified with the primer pair 341F-806R and sequenced on a MiSeq platform (MiSeqFGx, Illumina, San Diego, CA, United States). Raw reads were archived in NCBI. Reads were processed within the QIIME2 environment (version 2018.11; Bolyen et al., 2018). Amplicon Sequence Variants (ASVs) were generated using the DADA2 algorithm (Callahan et al., 2016). Removal of singletons and chimeric sequences was performed concomitantly, and phylogenetic trees were calculated (FastTree2 plugin). For taxonomic classification of representative ASVs, a primer-specific trained Naïve Bayes taxonomic classifier, based on the Silva 132 99% OTUs 16S database (Quast et al., 2013), was used. Mitochondria, chloroplasts and unassigned sequences were removed during taxonomic filtering steps. To standardise the number of reads across samples, a sampling depth of 5000 was applied. Rarefaction curves showed overall

a saturation at this depth. We ended up with 27,815,393 reads in total, which is 77 % of all input reads. Visualisations were created using R (version 3.0.2; R Development Core Team, 2008; Wickham, 2016), Inkscape (version 0.92.4; Harrington and the Inkscape Development Team, 2005), python (version 3.7.3; van Rossum, 1995), anvio (version 6.2; Eren et al., 2015), QGIS (version 3.4.4; QGIS Development Team, 2017), and gephi (version 0.9.2; Bastian et al., 2009).

Method S1 contains a detailed overview about all statistical analyses conducted. In brief, we worked with four different alpha diversity metrics (Shannon index, Faith's phylogenetic diversity, Pielou's evenness, and number of ASVs). Due to an overall consistency between these metrics, we focussed only on the Shannon index for statistical testing (Dunn's tests). In terms of beta diversity, we also analysed four different metrics (weighted UniFrac distances, unweighted UniFrac distances, Bray-Curtis distances, and Jaccard distances). For statistical analyses (i.e. pairwise PERMANOVAs, sample clustering dendrograms, Mantel tests, PERMDISP2s) we focussed on weighted UniFrac distances. For the establishment of a similarity network between sponge grounds we used Jaccard (dis-)similarities. ASV abundance tables were standardised by either using relative abundances or presence-absences. Presence-absence data was used for bipartite networks between ASVs and locations, or between ASVs and sampletypes. Relative abundance tables were used for redundancy analyses, in variation partitioning models, and for weighted gene correlation networks. Relative abundance tables combined with taxonomic annotation (on the microbial phylum level) were used in Linear Discriminative Analyses (LefSe). In addition, relative abundance tables on the microbial phylum- and class-level were used for the applied machine learning approach, using the Random Forests algorithm (Moitinho-Silva et al., 2017).

Sponge taxonomy

Taxonomic classification of all sponge specimens was done by leading taxonomic experts. A mixture of barcoding (18S, COI sequencing) of representative individuals was performed together with classical morphological analyses of the sponge skeleton (spicules). Preliminary taxonomic identifications were assigned directly on board ship, often in combination with *in situ* photographs, and validated at a later stage. In total we sampled all today known four sponge classes (Demospongiae, Hexactinellidae, Calcarea, and Homoscleromorpha). We further sampled 21 sponge orders, 52 sponge families, 107 sponge genera, and 169 sponge species. We standardised all taxonomic identifications with the help of the World Register of Marine Species (WoRMS; WoRMS Editorial Board, 2020), by using AphiaIDs.

Imaging

Photographs of sponges were taken in the field together with a size standard. For image processing, one representative picture per sponge species was selected and the background removed manually using GIMP (version 2.10.14; The GIMP Development Team, 2019), replacing the size standard with a scale bar. Using ImageJ (version 1.53a; Schneider et al., 2012), the area of each sponge was determined by (i) splitting color channels, (ii) adjusting the threshold, using the red color channel, (iii) using the *analyse particles* function. Parameters needed to determine the volume of each sponge specimens were also derived with the help of ImageJ, using the *measure* function. Area-volume ratios were calculated from this data with the help of the equations presented in **Table S8**. Statistical differences in the area-volume ratios between sponge types were assessed with the help of Dunn's tests. Linear regressions were performed between log-transformed area-volume ratios and carrying capacity, microbial heterogeneity, and richness.

Micrographs of microbes were generated by sectioning of sponge tissue via two approaches (i) ultra-thin sections, and (ii) semi-thin sections. For both approaches 17 sponge species were processed in 3 replicates each. Immediately after sampling, tissue samples were fixed onboard ship in 2.5 % glutaraldehyde in 0.1 M natriumcacodylate buffer (pH 7.4; Science Services GmbH). Back on land, samples were rinsed with buffer for three times at 4 °C. Afterwards, they were post-fixed for 2 h in 2 % osmiumtetroxide (Carl Roth) and washed again with buffer (3x15 min at 4 °C). A dehydration series was started with an ascending ethanol (ROTIPURAN® Carl Roth) series (2x15 min 30 % EtOH, 1x15 min 50 % EtOH, up to 70 % EtOH). After overnight storage at 4 °C, desilicification with 4 % suprapure hydrofluoric acid (Merck) was performed with an incubation time of approximately five hours. Thorough washing of samples followed (8x15 min in 70 % EtOH) with an overnight storage at 4 °C in between washing steps. Using a graded ethanol series (1x15 min 90 % EtOH and 2x15 min 100 % EtOH), dehydration of samples was continued, followed by gradual infiltration with LR-White resin (AgarScientific). Infiltration was performed at room temperature with the following steps: 1x1 h 2:1 Ethanol:LR-White; 1x1 h 1:1 Ethanol:LR-White; 1x1 h 1:2 Ethanol:LR-White; 2x2 h pure LR-White. Overnight incubation of samples in pure LR-White was performed at 4 °C before transferring them into fresh resin inside embedding capsules. These capsules were polymerised at 57 °C for 2 days. After manual trimming, sections were cut (in technical replicates) with an ultramicrotome (Reichert-Jung ULTRACUT E), equipped with a diamond knife (DIATOME, Switzerland). Ultra-thin sections were cut at 70 nm thickness and mounted onto pioloform coated copper grids (75 mesh; Plano). They were contrasted with

uranyl acetate (Science Services; 20 min incubation with subsequent washing steps) and Reynold's lead citrate (ingredients from Carl Roth; 3 min incubation with subsequent washing steps). Visualisation of ultra-thin sections was conducted on a Tecnai G2 Spirit BioTwin transmission electron microscope (FEI Company), using an acceleration voltage of 80 kV. Semi-thin sections were cut at 0.5 μm thickness, stained with Richardson solution (ingredients from Carl Roth; prepared as described in Mulisch and Welsch, 2015), and visualised with an Axio Observer.Z1 microscope (Zeiss, Germany). With the help of both sections the HMA-LMA status of the sponges was qualitatively determined. Based on micrographs derived from semi-thin sections, carrying capacities were determined using ImageJ's *picture thresholding* and *analyse particles* functions.

Contextual data

66 full water conductivity-temperature-depth (CTD) profiles were conducted in different ocean regions. Profiles were trimmed to a starting depth of 20 m below the ocean surface and reached down to ~ 5 m above the ocean floor. Based on the resulting temperature-salinity profiles, prevailing watermasses were classified manually based on literature (Ridgway, 1969; Loeng, 1991; Fahrbach et al., 1992; Freiwald et al., 2002; Lavin et al., 2004; Hu et al., 2007; Aksenov et al., 2010; Rüggeberg et al., 2011; Buhl-Mortensen et al., 2012; Chiswell et al., 2015; Dever et al., 2016; Storesund et al., 2017; Roberts et al., 2018). In total 24 environmental parameters were gathered in this study. **Table S11** provides a detailed overview which parameters were

included and by which method they were retrieved. Those parameters which were not measured *in situ*, but derived from climatologies, originate from three sources: (i) the World Ocean Atlas (WOA; version WOA18; Garcia et al., 2018a, 2018b; Locarnini et al., 2018; Zweng et al., 2018), (ii) the Global Ocean Data Analysis Project (Glodap; v2 2020; Olsen et al., 2019, 2020), and (iii) satellite data (MODIS; NASA Goddard Space Flight Center, 2018a, 2018b, 2018c). For the downloaded WOA and GLODAP datasets we always extracted the deepest depth layer of each grid location. Based on the exact sampling coordinates we then extracted the datapoints of the closest positions present in the WOA and GLODAP bottom depth layer. The mixed layer depth data used in this study was derived from (Monterey and Levitus, 1997), and the bathymetry data was based on ETOPO1. Location ids of sponge grounds were standardised using Marine Region Gazetteers. Correlations between the 24 environmental parameters were visualized with the help of a principle component analysis.

Availability of supporting data

The supplementary material of this manuscript is included in the Appendix of this thesis.

ACKNOWLEDGEMENTS

This research has been supported by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement no. 679849 (the SponGES project). We thank all colleagues, as well as the crews and scientific parties of all research expeditions for their valuable support.

REFERENCES

- Aksenov, Y., Bacon, S., Coward, A.C., and Holliday, N.P. (2010). Polar outflow from the Arctic Ocean: A high resolution model study. *J Mar Syst* 83: 14–37.
- Bastian, M., Heymann, S., and Jacomy, M. (2009). Gephi: an open source software for exploring and manipulating networks. In *International AAAI Conference on Weblogs and Social Media*.
- Bolyen, E., Rideout, J. R., Dillon, M. R., Bokulich, N. A., Abnet, C. C., Al-Ghalith, G. A., et al. (2019). Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* 37: 852–857.
- Bosch, T.C.G. and McFall-Ngai, M.J. (2011). Metaorganisms as the new frontier. *Zoology* 114: 185–190.
- Buhl-Mortensen, L., Bøe, R., Dolan, M.F.J., Buhl-Mortensen, P., Thorsnes, T., Elvenes, S., and Hodnesdal, H. (2012). Banks, troughs, and canyons on the continental margin off Lofoten, Vesterålen, and Troms, Norway. *Seafloor Geomorphology as benthic habitat*. doi:10.1016/B978-0-12-385140-6.00051-7.
- Busch, K., Endres, S., Iversen, M.H., Michels, J., Nöthig, E.-M., and Engel, A. (2017). Bacterial colonization and vertical distribution of marine gel particles (TEP and CSP) in the Arctic Fram Strait. *Front Mar Sci* 4: 1–14.

- Callahan, B.J., McMurdie, P.J., Rosen, M.J., Han, A.W., Johnson, A.J.A., and Holmes, S.P. (2016). DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* 13: 581–583.
- Chiswell, S.M., Bostock, H.C., Sutton, P.J.H., and Williams, M.J. (2015). Physical oceanography of the deep seas around New Zealand: A review. *New Zeal J Mar Freshw Res* 49: 286–317.
- Corinaldesi, C. (2015). New perspectives in benthic deep-sea microbial ecology. *Front Mar Sci* 2: 1–12.
- Danovaro, R. (2009). Methods for the study of deep-sea sediments, their functioning and biodiversity, Boca Raton: CRC Press.
- Danovaro, R., Snelgrove, P.V.R., and Tyler, P. (2014). Challenging the paradigms of deep-sea ecology. *Trends Ecol Evol* 29: 465–475.
- Dever, M., Hebert, D., Greenan, B.J.W., Sheng, J., and Smith, P.C. (2016). Hydrography and Coastal Circulation along the Halifax Line and the Connections with the Gulf of St. Lawrence. *Atmos - Ocean* 54: 199–217.
- Eren, A.M., Esen, Ö.C., Quince, C., Vineis, J.H., Morrison, H.G., Sogin, M.L., and Delmont, T.O. (2015). Anvi'o: an advanced analysis and visualization platform for 'omics data. *PeerJ* 3: e1319.
- Fahrbach, E., Rohardt, G., and Krause, G. (1992). The Antarctic coastal current in the southeastern Weddell Sea. *Polar Biol* 12: 171–182.
- Freiwald, A., Hühnerbach, V., Lindberg, B., Wilson, J.B., and Campbell, J. (2002). The Sula Reef Complex, Norwegian shelf. *Facies* 179–200.
- Garcia, H.E., Weathers, K., Paver, C.R., Smolyar, I., Boyer, T.P., Locarnini, R.A., et al. (2018a). World Ocean Atlas 2018, Volume 3: Dissolved Oxygen, Apparent Oxygen Utilization, and Oxygen Saturation. *NOAA Atlas NESDIS 83*.
- Garcia, H.E., Weathers, K., Paver, C.R., Smolyar, I., Boyer, T.P., Locarnini, R.A., et al. (2018b). World Ocean Atlas 2018, Volume 4: Dissolved Inorganic Nutrients (phosphate, nitrate and nitrate+nitrite, silicate). *NOAA Atlas NESDIS 84*.
- Harrington, B. and the Development Team (2005). Inkscape. Available online at: www.inkscape.org.
- Hentschel, U., Fieseler, L., Wehrl, M., Gernert, C., Steinert, M., Hacker, J., and Horn, M. (2003). Microbial diversity of marine sponges. In: Müller, W.E.G. (ed), *Molecular Marine Biology of Sponges*. Heidelberg: Springer, 60–88.
- Hu, H., Liu, Q., Lin, X., and Liu, W. (2007). The South Pacific subtropical mode water in the Tasman Sea. *J Ocean Univ China* 6: 107–116.
- Karner, M.B., Delong, E.F., and Karl, D.M. (2001). Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* 409: 507–510.
- Klitgaard, A.B. and Tendal, O.S. (2004). Distribution and species composition of mass occurrences of large-sized sponges in the northeast Atlantic. *Prog Oceanogr* 61: 57–98.
- Lavin, L., Valdes, L., Sanchez, F., Abaunza, P., Forest, A., Boucher, J., et al. (2004) The Bay of Biscay: the encountering of the ocean and the shelf. In: Robinson, A. and Brink, K. (eds), *The Sea*. Harvard University Press.
- Locarnini, R.A., Mishonov, A. V., Baranova, O.K., Boyer, T.P., Zweng, M.M., Garcia, H.E., et al. (2018). World Ocean Atlas 2018, Volume 1: Temperature. *NOAA Atlas NESDIS 81*.
- Loeng, H. (1991). Features of the physical oceanographic conditions of the Barents Sea. *Polar Res* 10: 5–18.
- Lozupone, C.A. and Knight, R. (2007). Global patterns in bacterial diversity. *Proc Natl Acad Sci* 104: 11436–11440.
- Lurgi, M., Thomas, T., Wemheuer, B., Webster, N.S., and Montoya, J.M. (2019). Modularity and predicted functions of the global sponge-microbiome network. *Nat Commun* 10: 1–12.
- Maldonado, M., Aguilar, R., Bannister, R.J., Bell, J.J., Conway, K.W., Dayton, P.K., et al. (2016). Sponge grounds as key marine habitats: A synthetic review of types, structure, functional roles, and conservation concerns. In: Rossi, S., Bramanti, L., Gori, A., and Orejas, C. (eds), *Marine animal forests*. Switzerland: Springer International Publishing Switzerland.
- McFall-Ngai, M., Hadfield, M.G., Bosch, T.C.G., Carey, H. V., Domazet-Lošo, T., Douglas, A.E., et al. (2013). Animals in a bacterial world, a new imperative for the life sciences. *Proc Natl Acad Sci* 110: 3229–3236.
- Moitinho-Silva, L., Steinert, G., Nielsen, S., Hardoim, C.C.P., Wu, Y.C., McCormack, G.P., et al. (2017). Predicting the HMA-LMA status in marine sponges by machine learning. *Front Microbiol* 8: 1–14.
- Monterey, G. and Levitus, S. (1997). Seasonal Variability of Mixed Layer Depth for the World Ocean. *NOAA Atlas NESDIS 14*.
- Mulisch, M. and Welsch, U. (2015) Romeis Mikroskopische Technik Aufl. 19.
- NASA Goddard Space Flight Center (2018a). Moderate-resolution imaging spectroradiometer (MODIS) Aqua chlorophyll data. *Ocean Ecology Laboratory, Ocean Biology Processing Group*. doi: data/10.5067/AQUA/MODIS/L3M/CHL/2018. Accessed on 06/23/2020.
- NASA Goddard Space Flight Center (2018b). Moderate-resolution imaging spectroradiometer (MODIS) Aqua particulate inorganic carbon data. *Ocean Ecology Laboratory, Ocean Biology Processing Group*. doi: data/10.5067/AQUA/MODIS/L3M/PIC/2018. Accessed on 06/23/2020.
- NASA Goddard Space Flight Center (2018c). Moderate-resolution imaging spectroradiometer (MODIS) Aqua particulate organic carbon Data; *Ocean Ecology Laboratory, Ocean Biology Processing Group*. doi: data/10.5067/AQUA/MODIS/L3M/POC/2018. Accessed on 06/23/2020.
- Olsen, A., Lange, N., Key, R.M., Tanhua, T., Álvarez, M., Becker, S., et al. (2019). GLODAPv2.2019 – an update of GLODAPv2. *Earth Syst Sci Data Discuss* 1–39.
- Olsen, A., Lange, N., Key, R.M., Tanhua, T., Bittig, H.C., Kozyr, A., et al. (2020). GLODAPv2.2020 - the second update of GLODAPv2. *Earth Syst Sci Data Discuss* 1–41.
- Pita, L., Rix, L., Slaby, B.M., Franke, A., and Hentschel, U. (2018). The sponge holobiont in a changing ocean: from microbes to ecosystems. *Microbiome* 6: 46.
- QGIS Development Team (2017). Geographic Information System. Open Source Geospatial Foundation Project. Available online at: www.qgis.osgeo.org.
- Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., et al. (2013). The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res* 41: 590–596.
- R Development Core Team (2008). R: A language and environment for statistical computing. Available online at: www.r-project.org
- Ramirez-Llodra, E., Brandt, A., Danovaro, R., De Mol, B., Escobar, E., German, C.R., et al. (2010). Deep, diverse and definitely different: Unique attributes of the world's largest ecosystem. *Biogeosciences* 7: 2851–2899.
- Ramirez-Llodra, E., Clark, M.R., Smith, C.R., Tyler, P.A., Rowden, A.A., Bergstad, O.A., et al. (2011). Man and the last great wilderness: Human impact on the deep sea. *PLoS One* 6: e22588.
- Ridgway, N.M. (1969). Temperature and salinity of sea water at the ocean floor in the New Zealand region. *New Zeal J Mar Freshw Res* 3: 57–72.
- Roberts, E.M., Mienis, F., Rapp, H.T., Hanz, U., Meyer, H.K., and Davies, A.J. (2018). Oceanographic setting and short-timescale environmental variability at an Arctic seamount sponge ground. *Deep Sea Res Part I Oceanogr Res Pap* 98–113.
- van Rossum, G. (1995). Python tutorial, Technical Report CS-R9526, Amsterdam: *Centrum voor Wiskunde en Informatica* (CWI).
- Rüggeberg, A., Flögel, S., Dullo, W.C., Hissmann, K., and Freiwald, A. (2011). Water mass characteristics and sill dynamics in a sub-polar cold-water coral reef setting at Stjernsund, northern Norway. *Mar Geol* 282: 5–12.
- Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat Methods* 9: 671–675.
- Snelgrove, P.V.R. (1999). Getting to the bottom of marine biodiversity: Sedimentary habitats: Ocean bottoms are the most widespread habitat on earth and support high biodiversity and key ecosystem services. *Bioscience* 49: 129–130.
- Storesund, J.E., Sandaa, R.A., Thingstad, T.F., Asplin, L., Albretsen, J., and Erga, S.R. (2017). Linking bacterial community structure to advection and environmental impact along a coast-fjord gradient of the Sognefjord, western Norway. *Prog Oceanogr* 159: 13–30.

- Sunagawa, S., Coelho, L.P., Chaffron, S., Kultima, J.R., Labadie, K., Salazar, G., et al. (2015). Structure and function of the global ocean microbiome. *Science* 348: 1261359–1261359.
- The GIMP Development Team (2019) GIMP. Available online at: www.gimp.org.
- Thomas, T., Moitinho-Silva, L., Lurgi, M., Björk, J.R., Easson, C., Astudillo-García, C., et al. (2016). Diversity, structure and convergent evolution of the global sponge microbiome. *Nat Commun* 7: 11870.
- Thompson, L.R., Sanders, J.G., McDonald, D., Amir, A., Ladau, J., Locey, K.J., et al. (2017). A communal catalogue reveals Earth’s multiscale microbial diversity. *Nature* 551: 457–463.
- Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*, New York: *Springer-Verlag*.
- WoRMS Editorial Board (2020). World Register of Marine Species. *VIZ*. www.marinespecies.org.
- Xavier, J.R., Cárdenas, P., Cristobo, J., Van Soest, R., and Rapp, H.T. (2015). Systematics and biodiversity of deep-sea sponges of the Atlanto-Mediterranean region. *J Mar Biol Assoc United Kingdom* 95: 1285–1286.
- Zweng, M.M., Reagan, J.R., Seidov, D., Boyer, T.P., Locarnini, R.A., Garcia, H.E., et al. (2018). World Ocean Atlas 2018, Volume 2: Salinity. *NOAA Atlas NESDIS 82*.

Chapter 8

SVAmPEx – An interactive and interdisciplinary visualisation platform of the global Deep-sea Sponge Microbiome Project database

SVAmPEx – An interactive and interdisciplinary visualisation platform of the global Deep-sea Sponge Microbiome Project database

Kathrin Busch^{1*}, Lara Schmittmann¹, Ute Hentschel^{1,2}

¹ GEOMAR Helmholtz Centre for Ocean Research Kiel, Düsternbrooker Weg 20, 24105 Kiel, Germany, ² Christian-Albrechts University of Kiel, Christian-Albrechts-Platz 4, 24118 Kiel, Germany. *Correspondence author.

The *SVAmPEx (Spongy Variability Amplicon Explorer)* software was developed to get an interactive view on a large and complex next generation sequencing dataset by facilitating data exploration and pattern recognition through customised novel visualisation designs. It was created to facilitate (i) interactive data mining (ii) data analysis (iii) interdisciplinary system-orientated approaches and (iv) data recycling. Users can temporarily upload novel *16S* amplicon data to interact with their own data, put them into larger context and access related curated resources. In its current initial configuration, the tool supports a comprehensive perspective on deep-sea sponge microbiomes in the oceanographic context and represents an applied approach to conservation-relevant baseline data.

Keywords: *16S rRNA amplicon sequencing, database, metadata, FAIR, deep-sea, sponge taxonomy, software, ontologies, biodiversity, Ocean Health, conservation*

BACKGROUND

Sponge-dominated communities can be found in diverse deep-sea environmental settings such as seamounts, mid-ocean ridges, canyons, slopes, shelves and fjords. These communities (termed *sponge grounds*) constitute important ecosystems in the deep ocean, which can be considered as hotspots of biodiversity and ecosystem services (Klitgaard and Tendal, 2004; Xavier et al., 2015; Maldonado et al., 2016). In contrast to other diverse deep-sea habitats, such as cold-water coral reefs or hydrothermal vent and cold seep systems, sponge grounds remain largely unexplored.

The Deep-sea Sponge Microbiome Project (**Chapter 7**) is a large-scale study, integrating *16S* amplicon sequencing data of seawater, sediment, and sponges, with a large set of ecological and physical metadata. The aim of this project was to evaluate the biodiversity of deep-sea sponge-associated microbiomes in the ecosystem context. In total 1546 samples were included, for each of which 143 metadata

entries are indicated in the accompanying database. These entries are standardised by the use of ontologies and terms, which are defined in an associated database catalogue (**Figure 1**). The database itself is set up as a relational database and will be made available via a public archive. Database indices are for example of biological, ecological, physical, methodological, temporal, and geographical nature.

Our main aim was to establish a global baseline of deep-sea sponge microbial community composition. When we started this coordinated effort in 2016, all previous studies on deep-sea sponge microbiomes were neither replicated, nor standardised, which prevented comparisons between studies. By setting up a large-scale sampling effort, we aimed to provide a comprehensive reference catalogue and to establish a baseline of natural variability in deep-sea sponge microbiomes. As human impacts on the

ocean are steadily increasing (Duarte, 2014; Visbeck, 2018; Jouffray et al., 2020), reference databases are urgently needed to assess ongoing changes and to put sampling snapshots into

a larger perspective. Along these lines, we envision that our dataset may be used for example in providing conservation advice on the protection of deep-sea sponge ground ecosystems.

Catalogue

Index + Definition (click on cells!)	Entry format	Entry format: standardisation according to	Entry options, ranges & examples	Index context	Index used in database sheet
av_representative_individual	(numeric)	-	value between 0.1 and 22.5; precision:0.1	Biological	taxonomy_WORMS
biogeochemical_province	(term)	internal		Ecological	oceanography
chlorophyll_surface_mgperm3_MODIS	(numeric)	International System of Units (SI)	ProvinceA ProvinceB ProvinceC ProvinceD ProvinceE	Ecological	oceanography
conductivity_insitu_mSpercm	(numeric)	International System of Units (SI)		Physical	oceanography
conductivity_mSpercm_GSW	(numeric)	International System of Units (SI)		Physical	oceanography
control_PCR	(term)	internal		Methodological	samples_MIMARKS
cruise_chief_scientist	chlorophyll_surface_mgperm3_MODIS	-		Acknowledgements	cruises_PANGAEA
cruise_chief_scientist_affiliation	Surface chlorophyll concentration at the sampling location as derived from satellite images.	Official Name		Acknowledgements	cruises_PANGAEA
cruise_contact_person	(term)	Official Name		Acknowledgements	cruises_PANGAEA
cruise_contact_person_affiliation	(term)	Official Name		Acknowledgements	cruises_PANGAEA
cruise_departure_date	(year-month-day)	International Organization for Standardization (ISO8601)	e.g. 2017-01-27	Temporal	cruises_PANGAEA
cruise_departure_location	(city, (land))	Official Name		Geographical	cruises_PANGAEA
cruise_leg	(term)	Official Name		Naming	cruises_PANGAEA
cruise_organising_institution	(term)	Official Name		Acknowledgements	cruises_PANGAEA
cruise_platform	(term)	Official Name		Acknowledgements	cruises_PANGAEA
cruise_research_area	(term)	-		Geographical	cruises_PANGAEA
cruise_return_date	(year-month-day)	International Organization for Standardization (ISO8601)	e.g. 2017-02-26	Temporal	cruises_PANGAEA
cruise_return_location	(city, (land))	Official Name		Geographical	cruises_PANGAEA
CTD_profile	seawater:(numeric)	internal	e.g. seawater1	Naming	samples_PANGAEA
density_sigma_insitu_kgperm3	(numeric)	International System of Units (SI)	value between 27.35 and 49.73; precision:0.01	Physical	oceanography
density_sigma_kgperm3_GSW	(numeric)	International System of Units (SI)	value between 27.71506 and 50.01373; precision:0.00001	Physical	oceanography
depth_level_WOA	(numeric)	International System of Units (SI)	value between 100 and 4900; precision:1	Physical	oceanography
depth_m	(numeric)	International System of Units (SI)	value between 6 and 4833; precision:0	Geographical	samples_PANGAEA
distance_to_MLD_m	(numeric)	International System of Units (SI)	value between -552 and 4748; precision:1	Physical	oceanography
dna_extract_concentration_ng/ul	(numeric)	International System of Units (SI)	value between 0 and 430.6; precision:0.1	Methodological	samples_MIMARKS
dna_extract_lab.date	(year-month-day)	International Organization for Standardization (ISO8601)	e.g. 2017-04-27	Methodological	samples_MIMARKS
dna_extract_ratio_260/230	(numeric)	-	value between 0 and 368.24; precision:0.01	Methodological	samples_MIMARKS
dna_extract_ratio_260/280	(numeric)	-	value between 0 and 45.57; precision:0.01	Methodological	samples_MIMARKS
dna_extract_volume_ul	(numeric)	International System of Units (SI)	value between 45 and 100; precision: 1	Methodological	samples_MIMARKS
dna_store_temp_degrees_c	(numeric)	International System of Units (SI)	-80	Methodological	samples_MIMARKS
doc_umolperkg_GLIDAP	(numeric)	International System of Units (SI)	value between 35.7 and 82.0; precision: 0.1	Ecological	oceanography

FIGURE 1 | Interactive database catalogue, containing definitions, examples, format specifications and context of every single database index (143 indices in total).

Promoting open science and data re-usability

To minimise technical variation and to foster transparency and reproducibility, we have developed standard operation procedures (SOPs) for the complete research pipeline, from sampling, to sample processing and the bioinformatic core analyses. These SOPs are going to be made publicly available and archived permanently by creation of digital object identifiers. To promote compliance of our dataset to the FAIR data principles (which stand for *findability, accessibility, interoperability, and reusability*; Wilkinson et al., 2016), and make our data available interactively we have designed and programmed an accompanying software tool, the *Spongy Variability Amplicon Exploration (SVAmPEX)* tool.

Overall purpose of the SVAmPEX-software

Large datasets derived from next generation

sequencing approaches have increased massively within the last decade (Margulies et al., 2005; Shendure and Ji, 2008; Glenn, 2011). With data amount and complexity increasing tremendously, there is growing need that researchers develop skills and tools preventing them to 'get lost in the mass of data'. From a conceptual point of view, SVAmPEX was developed to support a comprehensive perspective on deep-sea sponges and their associated microbiota in the oceanographic context on a global scale. Technically, it enables an interactive view on the large 16S rRNA gene amplicon dataset and the accompanying metadata, as well as pattern recognition through customized novel visualisation designs. The user can mine for data and is referred to the archived raw data upon need (deposited in existing public repositories such as PANGAEA and NCBI). Users can further temporarily upload novel 16S amplicon data to visualise their own data, put them into larger context, and access

related curated resources. One long-standing question in ecology is which biotic and abiotic factors drive biological diversity. Traditionally, researchers conduct ecological research on different spatial scales to answer their study questions, with large-scale studies typically conducted to derive general concepts, and small-scale studies being conducted for more specific questions. The developed tool brings together both approaches, by providing a large-scale reference catalogue, which can be interactively

queried for very specific targets. In this line, the developed software aims to facilitate detection of correlations among data pieces, e.g. to select aspects for mechanistic follow-up studies and design of targeted experiments. We envision that our database in combination with the SVampEx tool, may be particularly suited for interdisciplinary, system-orientated approaches by the marine scientific community and ocean managers.

Implemented functionalities and specific target audiences

A deployment of the software will be made publicly accessible and its code is open-source. In addition, SVampEx comes along with a complete bioinformatic data analysis pipeline for V3-V4 region 16S amplicon data, generating amplicon single nucleotide variants (ASVs). This pipeline can be used in an isolated manner for statistical analysis of 16S data, but notably it also produces ready to use output files which can be uploaded manually into the developed software interface. By providing an automated Standard Operating Procedure (SOP) from scratch, comparability between different stu-

dies is promoted. The developed software draws upon the strengths of visualisations, believing that novel visualisation techniques provide new insights into data. By setting up a ‘playful’ approach through interactivity, the developed tool aims at stimulating creativity and promoting thinking outside the box. Furthermore, it fosters a discovery-driven research approach. An overview on our general data generation and curation workflow, as well as on the main functionalities of SVampEx can be found in **Figure 2**. In the following subsections we will provide more details about each module.

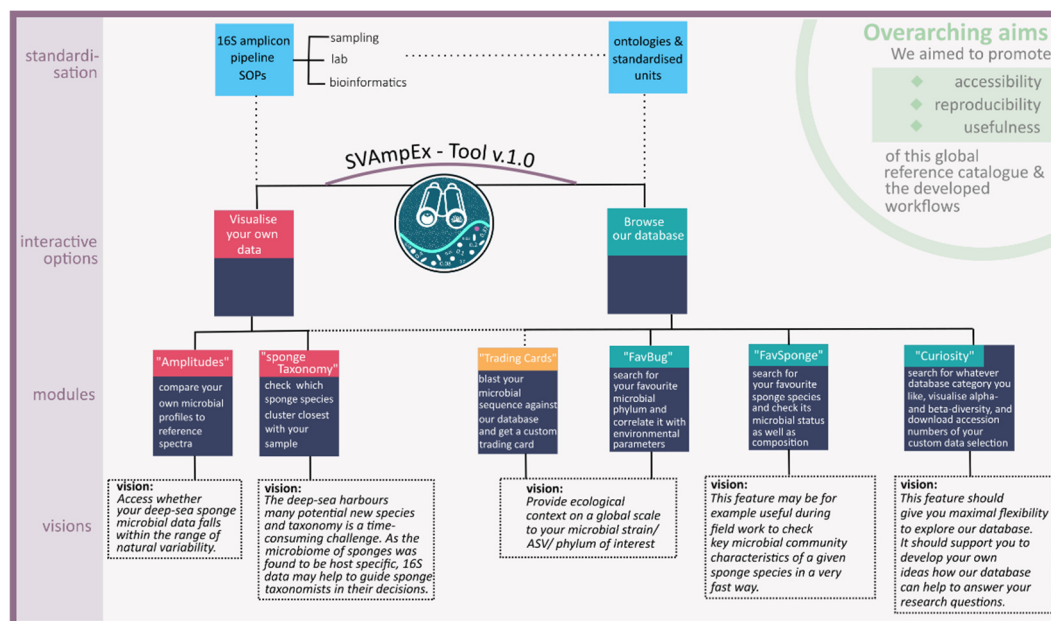


FIGURE 2 | General data generation and curation workflow as well as main modules developed and accessible within the SvAmpEx-software.

The ‘Amplitudes’ module: An option for Ocean Health assessments

One key feature within SVAmPE are 16S amplicon reference spectra, which were calculated based on a large number of deep-sea sponge microbiomes (Figure 3). Based on these reference spectra, the user can assess in an interactive manner if the microbial fingerprints of his/her own samples fall inside or outside the range of natural variability. An index (the *Svamp-i*) is calculated automatically when data

is uploaded in order to facilitate a more neutral interpretation of the presented graphic results and the user can scroll interactively through all of his/her samples. The mentioned index can be used for example in the context of Ocean Health assessments, as sick sponge hosts tend to have a microbial community composition which differs significantly from healthy reference samples (Webster et al., 2008; Gao et al., 2015; Blanquer et al., 2016; Luter et al., 2017; Deignan et al., 2018; Pita et al., 2018; Slaby et al., 2019).

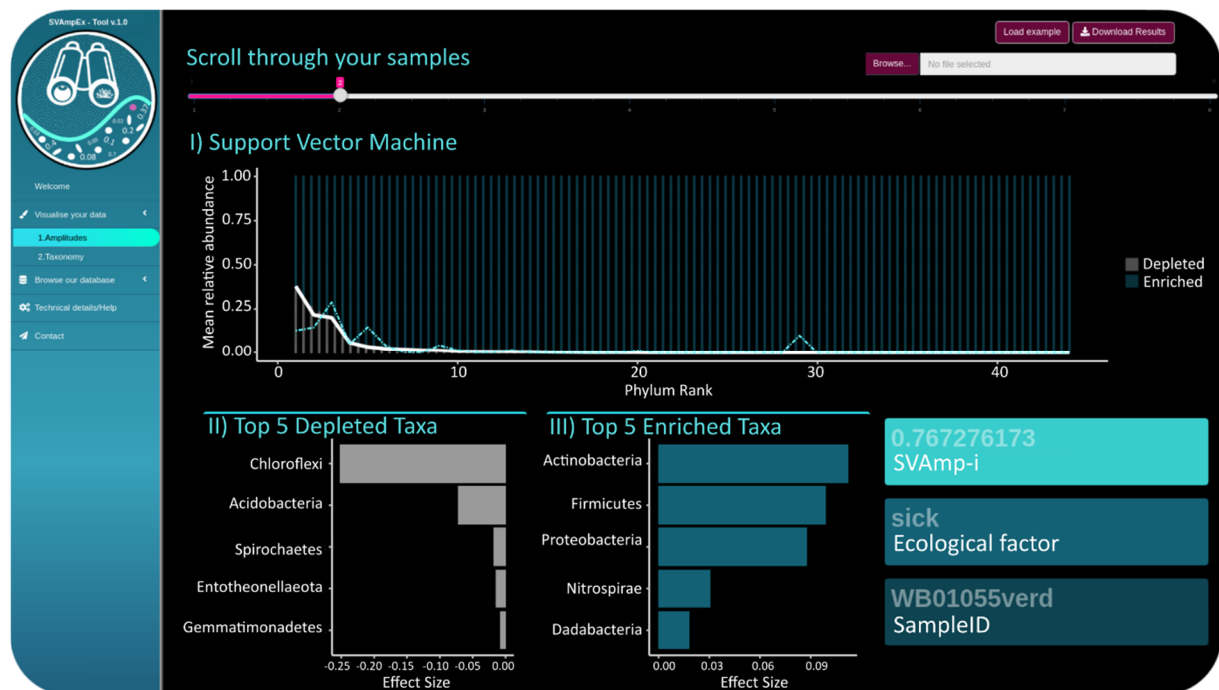


FIGURE 3 | A glimpse into the ‘Amplitudes’ module of SVAmPE. Depletions and enrichments of microbial taxa present in a newly uploaded dataset are shown in comparison to the database reference spectra. Further sample information is depicted, including the calculated SVAmPE-i index of that particular sample.

The ‘Sponge Taxonomy’ and ‘FavSponge’ modules: Discover biodiversity in the deep-sea

A scientific challenge is the taxonomic description of novel species from the deep-sea. Due to the fact that descriptions are time-consuming, researchers are slower at reporting new species than biodiversity is going extinct (Ramirez-Llodra et al., 2010; Snelgrove, 2016). With the help of SVAmPE, sponge species identification

and taxonomic description can be sped up by taking advantage of the host-specificity of sponge-associated microbial communities. The developed software provides the option to upload results from 16S rRNA amplicon sequencing and have data clustered with the backend database. The output is presented in a clustering dendrogram, in which the position and closest match of the uploaded

sample is indicated (Figure 4). This can give valuable hints for taxonomic identification and thus potentially accelerate species descriptions. Another option is to search for a specific sponge host and visualise its microbial community composition. During field-work, the software can help to promote sustainable and efficient sampling of deep-sea organisms onboard a research vessel. It allows fast detection of samples available in the long-term reference

collection and directly supports interactive visual exploration of the respective microbial biodiversity data. The programmed interactive graphical user interface further helps to make next generation sequencing data accessible to users who are interested in biodiversity, but are lacking the bioinformatic background and/or logistic capacities to process large, computing-intensive amounts of raw sequences.

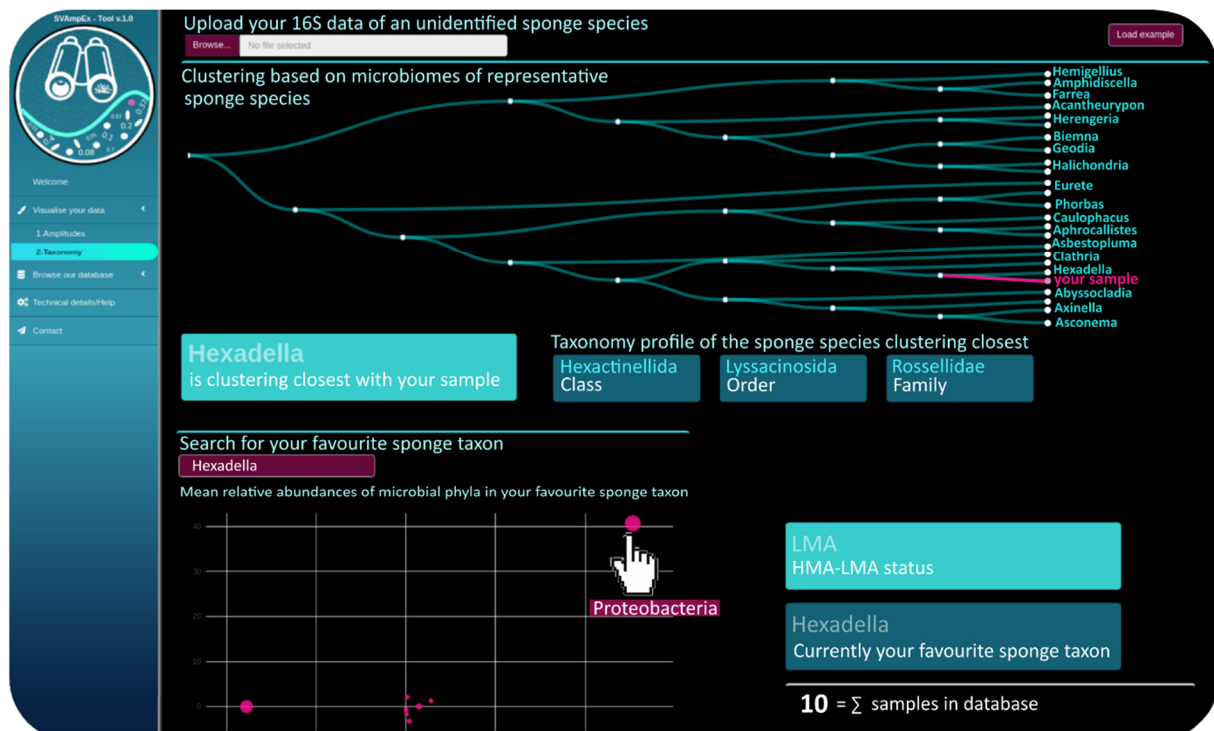


FIGURE 4 | A glimpse into the ‘Sponge Taxonomy’ and ‘FavSponge’ modules of SVampEx. An exemplary clustering dendrogram is shown, indicating the similarity of a newly uploaded dataset (‘your sample’) to samples in the reference database. In addition, key microbial community characteristics of a chosen sponge taxon are illustrated at the lower part of the figure. The microbial community composition can be explored by interactive hovering with the cursor.

The ‘FavBug’ and ‘Trading Cards’ modules: Assess the ecological context and distribution of microbial taxa

While numerous well-established databases with sophisticated interfaces exist in ocean research (e.g. OBIS (OBIS, 2020), WORMS (WoRMS Editorial Board, 2020), and the World Porifera Database (Van Soest et al., 2020)), these databases are traditionally not designed

to capture the *metaorganism*-level (i.e. both, the animal host and its associated symbionts). SVampEx can step in, as it represents an interdisciplinary biodiversity assessment platform for researchers from different biological fields (e.g. zoologists and microbiologists). This is achieved by visualising not only sponge diversity and microbial

diversity in an isolated manner, but by putting a focus on the biotic interplay between host diversity (sponges) and symbiont diversity (bacteria). Importantly, the developed software also accounts for the interplay of microbial diversity with abiotic factors (e.g. global measurements of O₂, CO₂, nutrients temperature, and many more parameters which are included in the database). Microbiologists, can actively search for microbial phyla of interest and visualise their relative abundances in correlation with biogeochemical parameters and host traits. The user can query/blast his/her raw sequences against the large dataset behind the scenes and visualise the prevalence of respective sequences in relation to the ecological context (Figure 5).

This function may be relevant for researchers working within traditional microbiology and biotechnology, as it can potentially point out directions in the cultivation and isolation of novel microbial strains. On a different note, SVAmPEX might also be a relevant resource for natural product chemists searching for novel pharmaceuticals, as sponge holobionts are known to be active producers of secondary metabolites (Blunt et al., 2006; Taylor et al., 2007). The created software may therefore help to detect niches, in which biotechnologically relevant microbial strains, such as for example known antibiotic producers (e.g. *Entotheonellaota*; Piel et al., 2004; Freeman et al., 2016; Bhushan et al., 2017; Mori et al., 2018), occur.

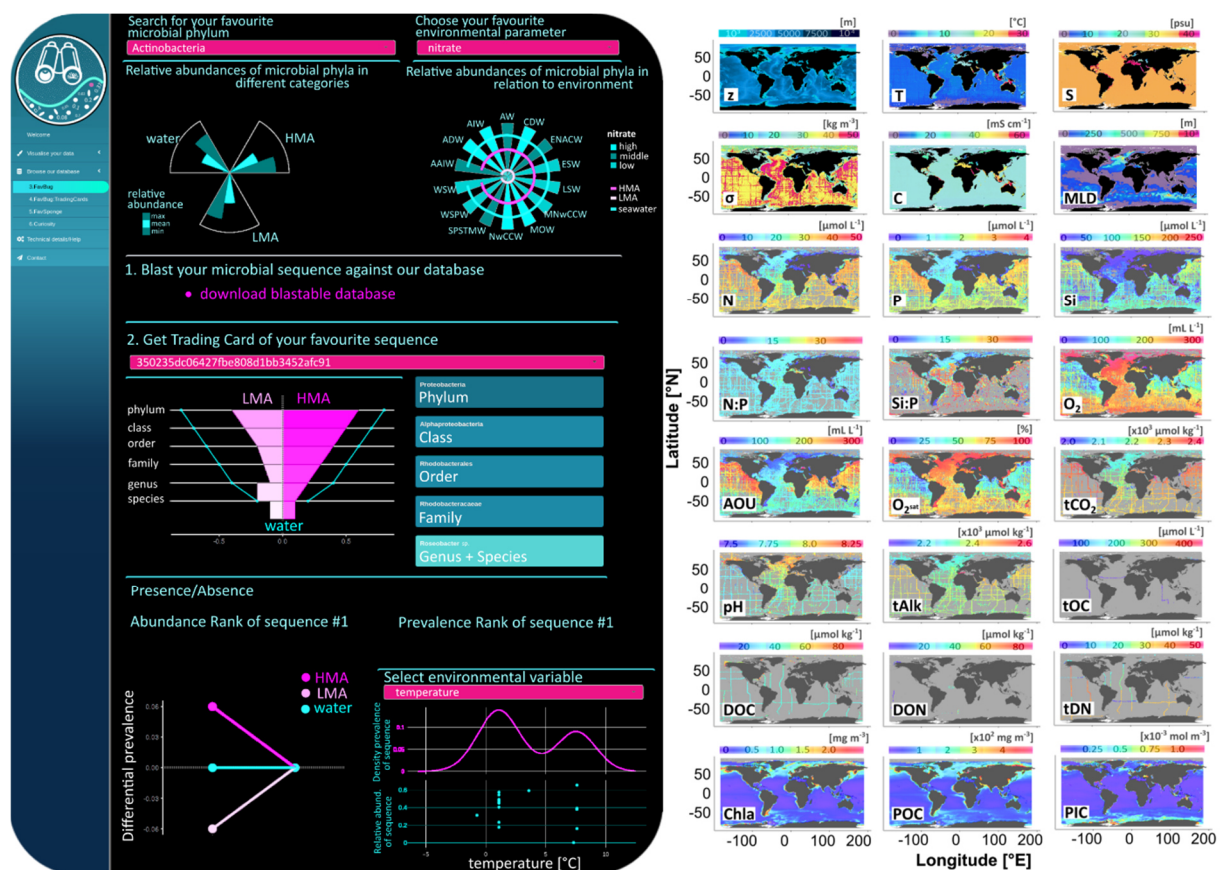


FIGURE 5 | A glimpse into the 'FavBug' and 'Trading Cards' modules of SVAmPEX. Prevalence and abundance of a chosen microbial taxa are shown in relation to samplotype and environmental gradients.

The 'Curiosity' module: Tailor the SVAmPEx database to your specific interests

This module is about enhancing the efficiency in data recycling. Importantly, SVAmPEx and the underlying database are not meant to replace of an official data archive. In fact, it was built to facilitate integration with public databases (such as PANGAEA and NCBI) by providing downloadable overviews of data links to respective repositories. This is useful, as for interdisciplinary studies (e.g. at the interface between molecular ecology, biogeochemistry, and physics) raw data are commonly stored in separate public databases. For example, sequence data are commonly stored in sequence archives such as NCBI, while the corresponding metadata are stored in

environmental databases such as PANGAEA. SVAmPEx includes a Geographic Information System (GIS) which allows to directly visualise key biodiversity parameters (microbial alpha-diversity indices) of selected samples on a map (**Figure 6**). This can help to determine spatial regions, and therefore data subsets, of interest. Querying is possible within various biologic, geographic as well as oceanographic categories and the resulting sample subsets are directly marked by colors inside ordination plots produced for the complete dataset. This approach facilitates the process of finding relevant data from the large dataset and therefore enhances the efficiency in data recycling. Additionally, it allows to identify relevant regions for sponge ground-associated microbial diversity and its conservation.



FIGURE 6 | A glimpse into the 'Curiosity' module of SVAmPEx. Here, the database can be subsetted to any category of interest. Accession numbers and links of the respective subset can be downloaded and basic characteristics (i.e. microbial alpha- and beta-diversity, as well as sampling locations) of the data subset are visualised automatically.

Documentation status & availability

The source code of SVAmPEx was written in R (shinyR) and will be made available via open-access. It can be executed on any operating system online via binder, without the need to conduct any software installations. In case of unstable or restricted internet connection (e.g.

during a research expedition), a desktop version can be downloaded. The structure of the graphical user interface includes a start page and a help page to guide users through the capabilities of the developed tool. In addition, the GUI includes 'load example'-buttons and 'help'-boxes on many tabs.

Future development and transfer potential

As the source-code is made publicly available and extensions of the existing code are straightforward, researchers can tailor SVampEx to their specific own needs and create their own, customised version. Potential future development could be for example integration of machine learning techniques. These approaches could help to continuously expand the database based on uploaded user data. On a different note, the developed code of SVampEx may be transferrable. The structure and the system could be easily transferred to shallow water sponges, or even go beyond this phylum and serve as an innovative example for comparative microbiome research (e.g. human gut microbiome). Generally speaking, the developed code allows to set up an interactive

visual data catalogue for large, congruent biodiversity datasets. These data catalogues can be used as add-ons to scientific publications. We suspect that this may be particularly relevant for large datasets where meta-analyses or re-analyses of several smaller studies are required. In these cases, the used data can be stored in fractionated subsets at various storage locations. SVampEx (or its adjusted source code) can in these cases serve as a coordination platform and single portal to the multiple data resources.

ACKNOWLEDGEMENTS

This research has been supported by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement no. 679849 (the SponGES project).

REFERENCES

- Bhushan, A., Peters, E.E., and Piel, J. (2017). *Entotheonella* bacteria as source of sponge-derived natural products: opportunities for biotechnological production. In Müller, W., Schröder, H., and Wang, X. (eds), *Blue Biotechnology. Progress in Molecular and Subcellular Biology*. Springer, Cham.
- Blanquer, A., Uriz, M.J., Cebrian, E., and Galand, P.E. (2016). Snapshot of a bacterial microbiome shift during the early symptoms of a massive sponge die-off in the western Mediterranean. *Front Microbiol* 7: 1–10.
- Blunt, J.W., Copp, B.R., Munro, M.H.G., Northcote, P.T., and Prinsep, M.R. (2006). Marine natural products. *Nat Prod Rep* 23: 26–78.
- Deignan, L.K., Pawlik, J.R., and Erwin, P.M. (2018). Agelas wasting syndrome alters prokaryotic symbiont communities of the Caribbean brown tube sponge, *Agelas tubulata*. *Microb Ecol* 76: 459–466.
- Duarte, C.M. (2014). Global change and the future ocean: A grand challenge for marine sciences. *Front Mar Sci* 1: 1–16.
- Freeman, M.F., Vagstad, A.L., and Piel, J. (2016). Polytheonamide biosynthesis showcasing the metabolic potential of sponge-associated uncultivated “*Entotheonella*” bacteria. *Curr Opin Chem Biol* 31: 8–14.
- Gao, Z.M., Wang, Y., Tian, R.M., Lee, O.O., Wong, Y.H., Batang, Z.B., Al-Suwailem, A., Lafi, F.F., Bajic, V.B., Qian, P.-Y. (2015). Pyrosequencing revealed shifts of prokaryotic communities between healthy and disease-like tissues of the Red Sea sponge *Crella cyathophora*. *PeerJ* 2015: 1–13.
- Glenn, T.C. (2011). Field guide to next-generation DNA sequencers. *Mol Ecol Resour* 11: 759–769.
- Jouffray, J.-B., Blasiak, R., Norström, A. V., Österblom, H., and Nyström, M. (2020). The blue acceleration: the trajectory of human expansion into the ocean. *One Earth* 2: 43–54.
- Klitgaard, A.B. and Tendal, O.S. (2004). Distribution and species composition of mass occurrences of large-sized sponges in the northeast Atlantic. *Prog Oceanogr* 61: 57–98.
- Luter, H.M., Bannister, R.J., Whalan, S., Kutti, T., Pineda, M.-C., and Webster, N.S. (2017). Microbiome analysis of a disease affecting the deep-sea sponge *Geodia barretti*. *FEMS Microbiol Ecol* 93: 1–6.
- Maldonado, M., Aguilar, R., Bannister, R.J., Bell, J.J., Conway, K.W., Dayton, P.K., et al. (2016). Sponge grounds as key marine habitats: a synthetic review of types, structure, functional roles, and conservation concerns. In Rossi, S., Bramanti, L., Gori, A., and Orejas, C. (eds), *Marine animal forests*. Switzerland: Springer International Publishing Switzerland.
- Margulies, M., Egholm, M., Altman, W.E., Attiya, S., Bader, J.S., Bemben, L.A., et al. (2005). Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437: 376–380.
- Mori, T., Cahn, J.K.B., Wilson, M.C., Meoded, R.A., Wiebach, V., Martinez, A.F.C., Helfrich, E.J.N., Albersmeier, A., Wibberg, D., Dätwyler, S., Keren, R., Lavy, A., Rückert, C., Ilan, M., Kalinowski, J., Matsunaga, S., Takeyama, H., Piel, J. (2018). Single-bacterial genomics validates rich and varied specialized metabolism of uncultivated *Entotheonella* sponge symbionts. *Proc Natl Acad Sci* 115: 1718–1723.
- OBIS (2020). Ocean Biodiversity Information System. *Intergov Oceanogr Comm UNESCO*. www.iobis.org.
- Piel, J., Hui, D., Wen, G., Butzke, D., Platzer, M., Fusetani, N., and Matsunaga, S. (2004). Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge *Theonella swinhoei*. *Proc Natl Acad Sci* 101: 16222–16227.
- Pita, L., Rix, L., Slaby, B.M., Franke, A., and Hentschel, U. (2018). The sponge holobiont in a changing ocean: from microbes to ecosystems. *Microbiome* 6: 46.
- Ramirez-Llodra, E., Brandt, A., Danovaro, R., De Mol, B., Escobar, E., German, C.R., et al. (2010). Deep, diverse and definitely different: Unique attributes of the world’s largest ecosystem. *Bio-geosciences* 7: 2851–2899.
- Shendure, J. and Ji, H. (2008). Next-generation DNA sequencing. *Nat Biotechnol* 26: 1135–1145.
- Slaby, B., Franke, A., Rix, L., Pita, L., Bayer, K., Jahn, M., and Hentschel, U. (2019). Marine sponge holobionts in health and disease. In Li, Z. (ed), *Symbiotic microbiomes of coral reefs sponges and corals*. Dordrecht, Springer Netherlands, 81–104.
- Snelgrove, P.V.R. (2016). An Ocean of Discovery: Biodiversity beyond the Census of Marine Life. *Planta Med* 82: 790–799.

- Van Soest, R.W.M., Boury-Esnault, N., Hooper, J.N.A., Rützler, K., de Voogd, N.J., Alvarez, B., Hajdu, E., Pisera, A.B., Manconi, R., Schönberg, C., Klautau, M., Kelly, M., Vacelet, J., Dohrmann, M., Díaz, M.-C., Cárdenas, P., Carballo, J.L., Ríos, P., Downey, R., Morrow, C.C. (2020). World Porifera Database. www.marinespecies.org/porifera.
- Taylor, M.W., Radax, R., Steger, D., and Wagner, M. (2007). Sponge-associated microorganisms: Evolution, ecology, and biotechnological potential. *Microbiol Mol Biol Rev* 71: 295–347.
- Visbeck, M. (2018). Ocean science research is key for a sustainable future. *Nat Commun* 9: 1–4.
- Webster, N.S., Xavier, J.R., Freckelton, M., Motti, C.A., and Cobb, R. (2008). Shifts in microbial and chemical patterns within the marine sponge *Aplysina aerophoba* during a disease outbreak. *Environ Microbiol* 10: 3366–3376.
- Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.W., da Silva Santos, L.B., Bourne, P.E., Bouwman, J., Brookes, A.J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C.T., Finkers, R., Gonzalez-Beltran, A., Gray, A.J.G., Groth, P., Goble, C., Grethe, J.S., Heringa, J., 't Hoen, P.A.C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S.J., Martone, M.E., Mons, A., Packer, A.L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M.A., Thompson, M., Van Der Lei, J., Van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data* 3: 1–9.
- WoRMS Editorial Board (2020). World Register of Marine Species. VLIZ. www.marinespecies.org.
- Xavier, J.R., Cárdenas, P., Cristobo, J., Van Soest, R., and Rapp, H.T. (2015). Systematics and biodiversity of deep-sea sponges of the Atlanto-Mediterranean region. *J Mar Biol Assoc United Kingdom* 95: 1285–1286.

One aspect which has fascinated me since the beginning of my PhD work is how individual organisational units are connected across scales. In the present thesis, I explored the microbial community composition of deep-sea sponges across different scales, from the ecosystem- and biogeographical-level, to individual sponge species and to the microbial taxon level. By looking at sponge microbiomes on different levels of integration and by using a nested sampling design, I was able to identify the overarching factors, that drive microbiome composition in a statistically proven manner. Generalisation is possible as > 1000 sponge microbiomes were explored in this PhD thesis. In the following, I will discuss the main findings of my thesis. Microbiome composition data was generated for all four existing sponge classes. Particularly glass sponge microbiomes have not been well studied at the onset of this thesis, but they constitute a large sample fraction in our study (n=243 sponges belonging to 56 species). Establishment of a reference baseline is key in order to evaluate VME integrity and resilience. Along these lines my findings contribute to improved conservation and management strategies of VMEs in the long run.

Microbial diversity in deep-sea sponges

Comparison between deep and shallow water sponge microbiomes

Most of the knowledge in sponge microbiology has been gained from studies on shallow water sponges while only a few studies have explored deep-sea sponge microbiomes (Kennedy *et al.*, 2014; Reveillaud *et al.*, 2014; Tian *et al.*, 2016; Borchert *et al.*, 2017). This thesis includes quantitative and qualitative analyses of deep-sea sponge microbiomes. To place findings into perspective, I compiled a global comparison between shallow and deep water microbiomes. This was conducted based on the two complementary large-scale datasets which are now available on sponge-associated microbial communities: the shallow water sponge microbiome project (SMP; Thomas *et al.*, 2016; Moitinho-Silva *et al.*, 2017c), and the deep-sea sponge microbiome project (DSMP; i.e. this PhD thesis **Chapter 7** and **Chapter 8**) (**Fig. 12**). While the average sampling depth of the SMP was 10 m, the average sampling depth of the DSMP was 657 m. Interestingly, the covered sponge species did largely not overlap (**Fig. 12b**). As for this PhD thesis every available sponge species was sampled at the respective collection sites, I conclude that the majority of sponge species is adapted to either, shallow waters or to deep-waters.

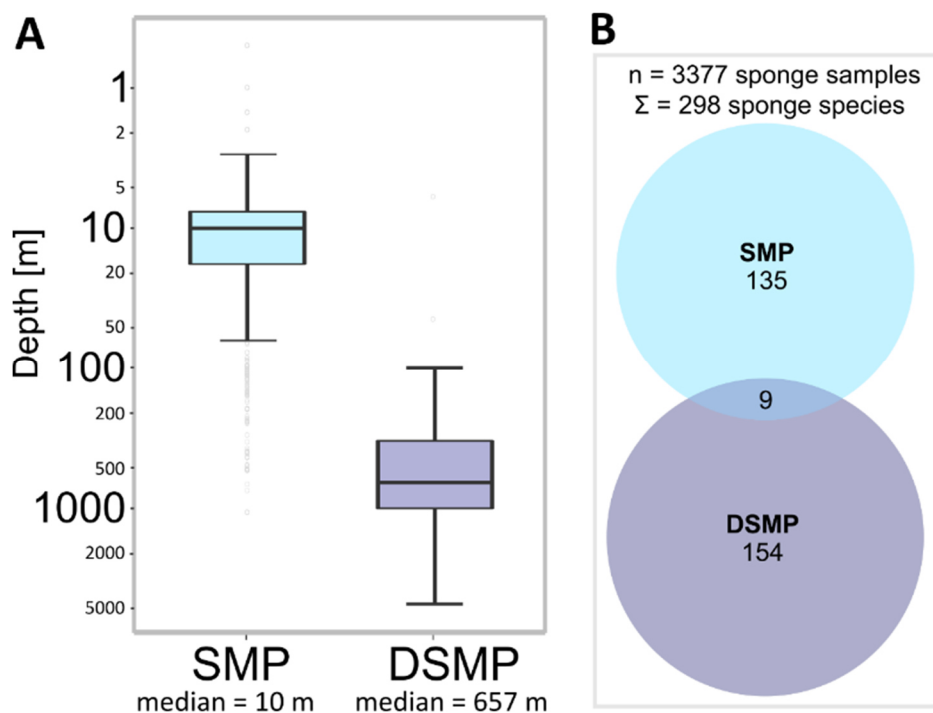


Fig.12 Overview of the two complementary large datasets on sponge microbiomes. SMP=Thomas *et al.*, 2016; DSMP= this thesis. **A)** Difference in main sampling depth. **B)** Overlap of covered sponge species.

In our publication (Steinert *et al.*, 2020) we showed that bacterial gene copy numbers were within the same range as known from shallow-water sponges. However, archaeal gene copy numbers were up to three orders of magnitude higher in deep-sea sponges than in shallow water sponges. We concluded that this highlights the importance of archaeal symbionts for deep-sea sponges.

Alpha diversity (richness) was overall lower in deep-sea sponges (22 to 1238 ASVs per host; **Chapter 7**) than in shallow-water sponges (50-3820 OTUs per host; Thomas *et al.*, 2016). A less rich deep-sea sponge microbial community is consistent with expectations because pelagic microbial community complexity also decreases with depth. Due to their overall low cell numbers, encounter rates of microbes in the deep pelagic ocean are rare. Based on our quantitative and qualitative results we conclude that sponges should be considered major hotspots of microbial interactions in the deep ocean. A direct comparison of beta diversity was not feasible between the SMP and the DSMP datasets because different primers had been used. Instead I provide further comparisons between shallow water and deep-sea sponge microbiomes on the microbial phylum-level, on the HMA-LMA-level, and on the community structure-level (**see below**).

The key microbial players: Old and new friends

We identified the *Proteobacteria* (47.6 %), *Chloroflexi* (15.8 %), *Nitrospirota* (3.2 %), *Nitrospinota* (1.1 %), *Entotheonellaeota* (0.5 %), *Poribacteria* (0.3 %), and *Cyanobacteria* (<0.1 %)

as the major phyla in deep-sea sponges. Here, the indicated percentage values represent average relative abundances across all analysed sponges. Similar to shallow water habitats (Thomas *et al.*, 2016), *Proteobacteria* were dominant in most examined sponges species. Notably, cyanobacterial relative abundances were much lower in deep-sea sponges (<0.1% relative abundance) compared to shallow-water sponges (up to ~10 % relative abundance; Thomas *et al.*, 2016). On the contrary, *Chloroflexi* were more enriched in deep-sea sponges (15.8 % relative abundance; compared to up to ~10 % relative abundance; Thomas *et al.*, 2016). The observed decrease in (photoautotrophic) cyanobacterial abundances with decreasing light availability in the deep-sea was expected. *Chloroflexi* were previously shown to play a relevant role in (recalcitrant) organic matter recycling (Landry *et al.*, 2017; Bayer *et al.*, 2018). Their observed higher relative abundances in deep-sea sponges may point into a similar direction: As particulate organic matter concentrations generally decrease with depth, a nutritional mode which is more tailored towards a recycling of dissolved organic compounds may be beneficial. *Chloroflexi* may therefore be equipped with a useful enzymatic repertoire, which enables them to flourish in the deep-sea environment. It is noteworthy however, that different primer sets were used for the shallow water dataset (Thomas *et al.*, 2016) and the DSMP dataset. In accordance with previous observations from shallow-water sponges, we observed a dominance of *Thaumarchaeota* in the archaeal community (on average 99.9 % relative abundance of the total archaeal community). We detected both, *Candidatus Nitrosopumilus* and *Candidatus Cenarchaeum* in the analysed deep-sea sponges and speculate about their potential role in ammonia-oxidation. I conclude that the phylum-level richness is high in the bacterial community with the typical phyla present, but with relatively more *Chloroflexi* and relatively less *Cyanobacteria*. On the contrary, richness is very low in the archaeal community with a pronounced dominance of the phylum *Thaumarchaeota*.

Taxonomic novelty in deep-sea sponge microbiomes

We detected a total of 71 microbial phyla associated with deep-sea sponges (n=931; seawater=66 phyla; sediment=73 phyla). In comparison, the previous large-scale survey of shallow water sponge microbiomes detected 42 sponge-associated microbial phyla (Thomas *et al.*, 2016). This suggests that microbial novelty at high taxonomic ranks is surprisingly high in deep-sea sponges, and that the diversity of sponge-associated microbes may be higher in the deep-sea than in shallow waters. More than 41 % percent of all microbial phyla had previously not been detected in sponges. We showed furthermore that the number of *unknown* microbial taxa was very high at lower taxonomic levels. For example, 31.6 % of all bacterial taxa remained unclassified on class level in the glass sponge *Vazella pourtalesii*. I conclude that the hidden taxonomic novelty in deep-sea sponge microbiomes is extensive.

Sponge-specificity and host individual-specificity of deep-sea sponge microbiomes

The phyla *Asgardarchaeota*, *Deferribacterota*, and *Hydrothermarchaeota* were exclusively found in sponges albeit at low abundances below 0.1 % (average across all 931 sponges; **Chapter 7**). Interestingly, *Asgardarchaeota* have been highlighted in previous studies due to their proposed important role in the early stages of eukaryotic cell origin (Spang *et al.*, 2015, 2017;

Eme *et al.*, 2017; Zaremba-Niedzwiedzka *et al.*, 2017). We further observed *Nitrosopumilaceae* to be sponge- but not sponge species-specific. Statistical analyses (LEfSe) revealed *Chloroflexi*, *Proteobacteria*, *Nitrospirota*, *Dadabacteria*, *Spirochaetota*, *Entotheonellaeota*, *Anck6*, and *PAUC34f* as sponge-enriched phyla. *Poribacteria* and *Entotheonalla* were detected in sponges as well as seawater and sediment reference samples. Although not statistically significantly enriched in sponges, *Poribacteria* still had on average higher relative abundances in sponges compared to environmental reference samples (sponges=0.3 %, seawater=<0.1 %, sediment=0.1 %). In all our studies sponges clustered clearly apart from environmental reference samples, which points out that a sponge-specific microbiome indeed exists in the deep-sea. Glass sponges seem to harbor a distinct microbiome, while carnivorous sponges showed no surprisingly distinct signature. In total we detected three distinct sponge groups, whose microbiomes clustered clearly apart from each other. I classified them as LMA demosponges (*LMA_demo*), LMA glass sponges (*LMA_glass*), and HMA sponges (*HMA*).

The relation between microbial community composition and deep-sea sponge phylogeny is discussed in more detail below. Besides imprints of sponge phylogeny, our results imply that the deep-sea sponge microbiome is host individual-specific on the ASV-level. A high variability between individuals is in line with observations by Jahn (2019), but in contrast to results by Webster *et al.* (2012). Between 83.6 % (for HMA sponges) and 95.5 % (for LMA sponges) of all bacterial ASVs occurred in less than 10 samples. Specifically, on average 65.5 % occurred in only one sample (**Fig.13**). We observed a high individuality in both, the bacterial community, but also in the archaeal community. Concerning potential drivers of the observed individuality, both, niche theory and neutral theory may apply. The following papers provide background readings for these concepts: Wiens, 2011 for niche theory; Caswell, 1976 and Sieber *et al.*, 2019 for neutral theory; and Wennekes *et al.*, 2012 for an integrated perspective. I would expect that the nestedness of each sponge microbiome into an individual ecosystem context and evolutionary history, together with a stochastic component, ultimately results in highly specific assemblages. Disentangling the individual drivers might however be a tremendous, if not impossible, task.

We generally observed highest alpha diversity in sediment samples, intermediate diversity in seawater samples, and lowest alpha diversity in sponges. The number of observed ASVs ranged between 22 to 1238 ASVs per sponge individual. Alpha diversity patterns were strictly correlated with sample type and remained constant within each sample type across various environmental categories (e.g. ecological realm, world oceans, geologic setting, watermass, biogeochemical province, and seawater cluster). Beta diversity on the contrary differed significantly across the just mentioned categories. I suppose that the extensive degree of microbiome individuality may also play a role in this respect.

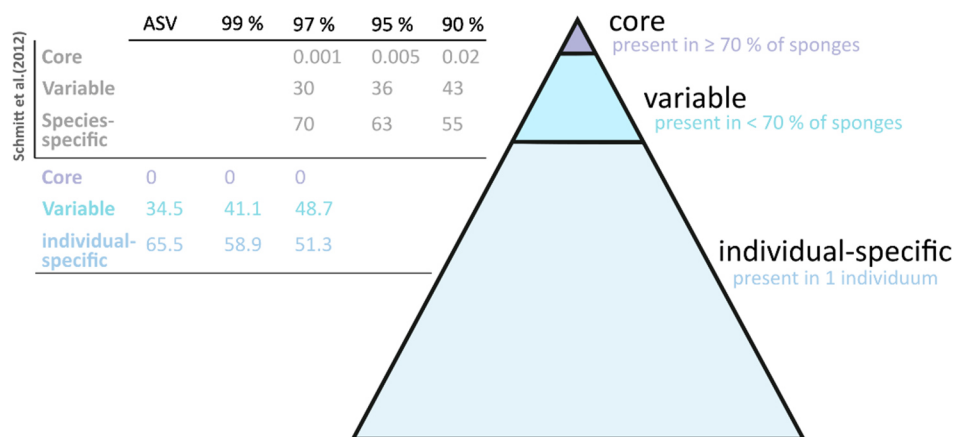


Fig.13 Schematic representation of the sponge microbiome after Schmitt *et al.*, 2012 expanded for the deep-sea sponge microbiomes covered in this PhD thesis (colorful values in table). Note that the term species-specific was replaced by the term individual-specific.

Is the core community concept a reasonable approach?

The core community concept only considers those microbes that are persistent members of a community (Shade and Handelsman, 2012; Astudillo-Garcia *et al.*, 2017). Core community members can be determined through time, among different environments, different hosts or different geographic locations (Astudillo-Garcia *et al.*, 2017) and are highly context dependent. The actual core community membership may thus depend strongly on the definition of core (but see Astudillo-Garcia *et al.*, 2017 who suggest a surprising insensitivity of results to changing core definitions). In our global dataset on the ASV-level, we observed that indeed, 0 microbial ASVs fulfill the definition of core community members based on Schmitt *et al.*, 2012 (**Fig.13**). I conclude that the natural variability of deep-sea sponge microbiomes is very high and that a classical core community does not exist. It has to be noted that core community depends on the size and heterogeneity of the respective dataset. I would expect a negative correlation between the numbers of core community members and the heterogeneity of the dataset, e.g. if many different sponge species are covered. For our large and diverse dataset, the core community concept (applied as in Schmitt *et al.*, 2012) was not applicable. This observation of a lacking core community taken together with the observed high individuality of the microbiome, leads to the philosophical question about what we, the scientific community, regard as the sponge microbiome. Are we applying a strict definition (core concept) and neglect the majority of resident microbial taxa? Or are we considering the full residing diversity and risk the capacity to separate signal from noise? From my point of view trait-based concepts may be very suitable for such types of studies. Future studies should be directed towards finding novel ways of grouping individual microbial taxa together based on ecological traits, instead of sequence similarity.

A deep-sea sponge specific microbiome

In a previous study by Kennedy *et al.* (2014) evidence of a deep-sea specific microbial community in sponges has been presented based on the analysis of four sponges. Although we used different approaches than Kennedy *et al.* (2014) and a much larger dataset (931 sponges), I also conclude that a deep-sea sponge specific bacterial microbiome exists, but due to different

reasons. I derive my conclusions from the following observations: (i) glass sponges showed a specific microbial community composition and almost exclusively occur in the deep-sea; (ii) sponges showed a host individual-specific microbial community composition, hence deep-sea sponge individuals are per se different to shallow water individuals; (iii) sponge community composition differs between shallow and deep-water, therefore the species-specific fraction of the microbiome will differ between shallow and deep-sea sponges; (iv) depth turned out as one of the main drivers of microbiome variability, suggesting that also for those taxa which are shared between shallow and deep-waters, relative abundances may differ; (v) deep-sea sponges contained 29 more microbial phyla than found in shallow waters. Therefore multiple taxa may thrive exclusively in deep-sea sponges; (vi) the observation of a lower bacterial richness per host in deep-sea sponges compared to shallow water sponges, suggests in combination with the other just mentioned aspects, that deep-sea sponge microbiomes not just harbor some additional microbes, but rather that their composition is quite substantially different from shallow water counterparts.

Microbial lifestyle within a deep-sea glass sponge

In our publication (Bayer, Busch *et al.*, 2020), we studied the microbial lifestyle within the glass sponge *Vazella pourtalesii* by metagenomics and microscopy. We observed symbionts to have particularly small genomes and low GC content. Microscopy revealed particularly small microbial cell sizes (Fig. 14).

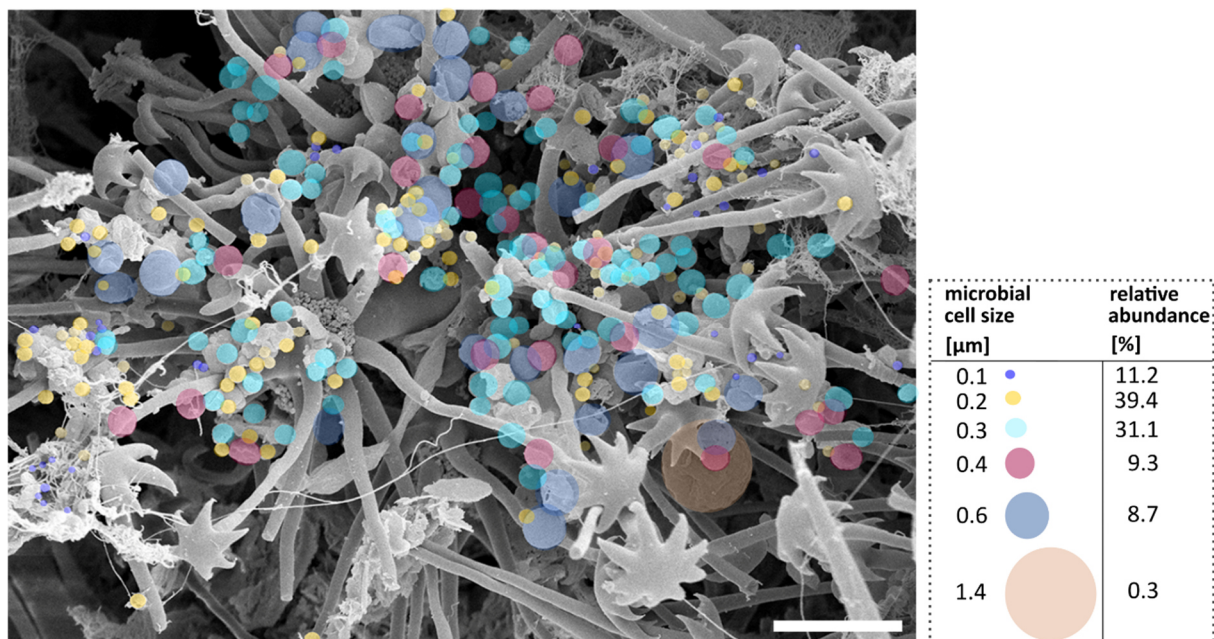


Fig. 14 Scanning electron microscopy of the syncytial tissue of the glass sponge *Vazella pourtalesii*. Microbial cells are colored according to their cell size. Scale bar = 2 μm. Picture by Kathrin Busch & Jan Michels.

These observations of genome reduction, minimal cell sizes and complexity point towards the phenomenon of *streamlining* in the LMA glass sponge *V. pourtalesii*. We suspect a particular

intimate host-microbe association as an adaptation to life in the deep-sea. We further examined the interactions between individual microbial members of the *V. pourtalesii* microbiome. In particular we checked on SAR324 bacteria, *Crenarchaeota*, *Patescibacteria*, and *Nanoarchaeota*, because they appeared to be particularly abundant. Based on functional analyses of metagenome-assembled genomes (MAGs), we identified two major metabolic strategies. We termed one strategy the “givers” and the other one the “takers”. *Crenarchaeota* and SAR324 were givers, meaning that they were heterotrophs and facultative anaerobes, which have the capacity to produce and partly secrete all required amino acids and vitamins. *Nanoarchaeota* and *Patescibacteria* were takers, meaning that they were anaerobes with reduced genomes which rely on the services of other microbial community members for resource provision. Our results suggest that the microbiome of the deep-sea glass sponge *V. pourtalesii* is distinct in terms of its metabolic intertwinedness. The microscopic view into *V. pourtalesii* revealed that microbes and host cells were located in dense biomass patches, which were presumably aggregated syncytial tissue. We further observed smaller cells attached to larger ones via pilus-like structures. Together with the metagenomic data, these observations point out that spatial distribution on the microscale may be important for biotic interactions.

Sponge host related factors driving microbiome composition

Host phylogeny

Host phylogeny was identified as one determinant of the microbial community composition. This is in line with records from shallow water sponges (Easson and Thacker, 2014; Moitinho-Silva *et al.*, 2017b). We observed a particular strong relation between the microbial community composition and sponge host taxonomy on the host phylum, class and order level. Further, we also found microbiomes to be overall host-species specific, despite the large degree of observed individual-specific variation. Sponge taxonomy is challenging and can frequently not be properly resolved despite considerable efforts, using classical taxonomy (spicules) and molecular (*18S*, *COI*, *28S*) markers. In these cases, where taxonomy could not be resolved, we identified microbiome composition as a helpful tool to disentangle host taxonomy. This observation may be valuable as the number of novel sponge taxa still waiting to be discovered is particularly high in the deep ocean. We have therefore developed a feature within the SVampEx software which should provide valuable hints for taxonomic identification based on *16S* data, and thus potentially accelerate species descriptions of deep-sea sponges.

HMA-LMA dichotomy

Whether the HMA-LMA dichotomy would also exist in other places than the shallow ocean, was not clear at the time when this PhD thesis was started. Our results reveal that it does indeed also exist in the deep ocean. For this evaluation I conducted microscopic analyses of 17 sponge species. Similar to what is known from shallow water sponge microbiomes, microbes were mainly located extracellularly in the mesohyl of the visually examined deep-sea sponges. One exception to this pattern was the sponge *Petrosia crassa*, which harboured microbial cells intracellularly, inside bacteriocytes. Three representative pictures are shown in **Fig.15**.

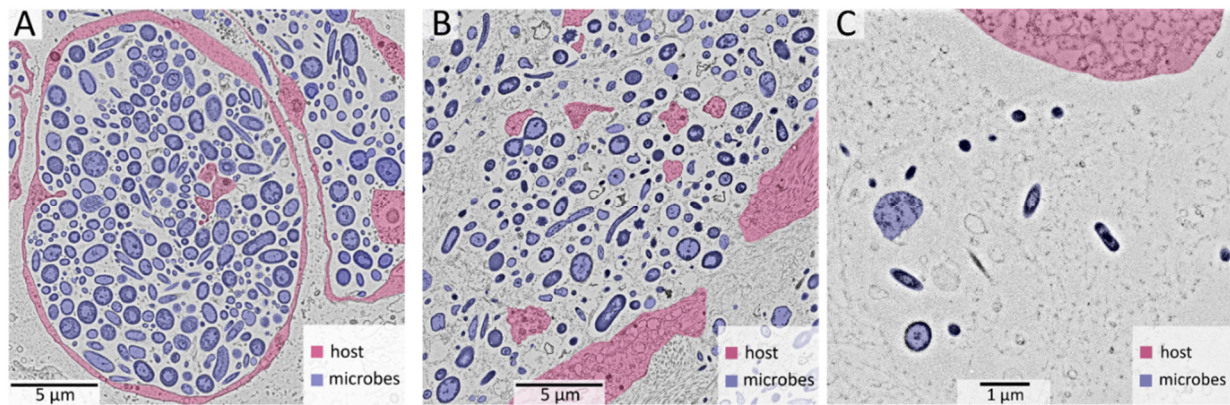


Fig.15 Representative microbial images of deep-sea sponges analysed in the context of this PhD thesis. **A)** Intracellular symbionts within bacteriocytes of a HMA sponge (*Petrosia crassa*). **B)** Extracellular symbionts in a HMA sponge. **C)** Extracellular symbionts in a LMA sponge.

Similar to the shallow water environment (e.g. Blanquer *et al.*, 2013), the HMA-LMA dichotomy seems to be the main driver of the microbial community structure in deep-sea sponges. Based on the machine learning approach which has been designed for shallow water sponge microbiomes (Moitinho-Silva *et al.*, 2017b), we were able to predict the HMA-LMA status. In terms of enrichment, *Chloroflexi*, *Acidobacteriota*, *Dadabacteria*, *Gemmatimonadota*, *Myxococcota*, *Entotheonellaeota*, *Spirochaetota*, *Poribacteria* were the eight most enriched taxa in HMA over LMA sponges. On the other side, *Proteobacteria*, *Bacteroidota*, *SAR324 clade*, *Planctomycetota*, *Verrucomicrobiota*, *Nitrospirota*, *Patescibacteria*, and *Marinimicrobia* were the eight most enriched taxa in LMA over HMA sponges. Half of these phyla (*Spirochaetota*, *Poribacteria*, *Gemmatimonadota*, *Chloroflexi*, *Acidobacteriota*, *Proteobacteria*, *Planctomycetota*, *Bacteroidota*) overlap with the HMA-LMA indicator phyla known from shallow waters. Notably, *Nitrospirota* were classified as HMA indicator phyla for shallow waters, while they were classified as LMA indicator taxa for deep-sea sponges in this thesis. We observed an overall dominance of LMA sponges in our dataset (i.e. 131 LMA species vs 38 HMA species). Our results revealed a lower community similarity between LMA sponges, than between HMA sponges. This observation may indicate that LMA sponges host more specialists, while HMA sponges host more generalist (Erwin *et al.*, 2011). Similar to what is known from shallow waters (Moitinho-Silva *et al.*, 2017b), we find the HMA-LMA dichotomy to not be strictly related with host phylogeny in deep-sea sponges.

While it has been noted that HMA and LMA sponges commonly thrive at the same shallow water location, a systematic analysis about their spatial distribution has not been previously published. With the data of this PhD thesis at hand, the global distribution patterns of HMA and LMA sponges were evaluated. For this evaluation I joined the data of the SMP with our newly generated dataset from the deep-sea. Based on the combined data set, I calculated the fraction of HMA and LMA sponges for defined depth ranges (**Fig.16a**). Results suggest that the proportion of HMA and LMA is relatively constant for water depth above 800 m. Below this depth, LMA sponges constitute a higher proportion. For a second analysis, I extracted the data on species occurrence for those sponge species which have already been classified as HMA or LMA by either Moitinho-Silva *et al.* (2017b,c), or myself (**Chapter 7**) from the OBIS

database. I could match 152 sponge species in total, providing me with occurrence data of 32717 sponge individuals in total (5908 HMA sponges and 26809 LMA sponges). I then determined geographic areas where only HMA or only LMA sponges or both together occurred (**Fig.16b**). This approach revealed that indeed large areas exist where only one single sponge type had been recorded, which is mainly true for LMA sponges. The observation of areas where seemingly only one sponge type persists is in line with the previous detection of monospecific sponge grounds (but note that the picture may potentially be biased by sampling efforts skewed towards a specific sponge type or species).

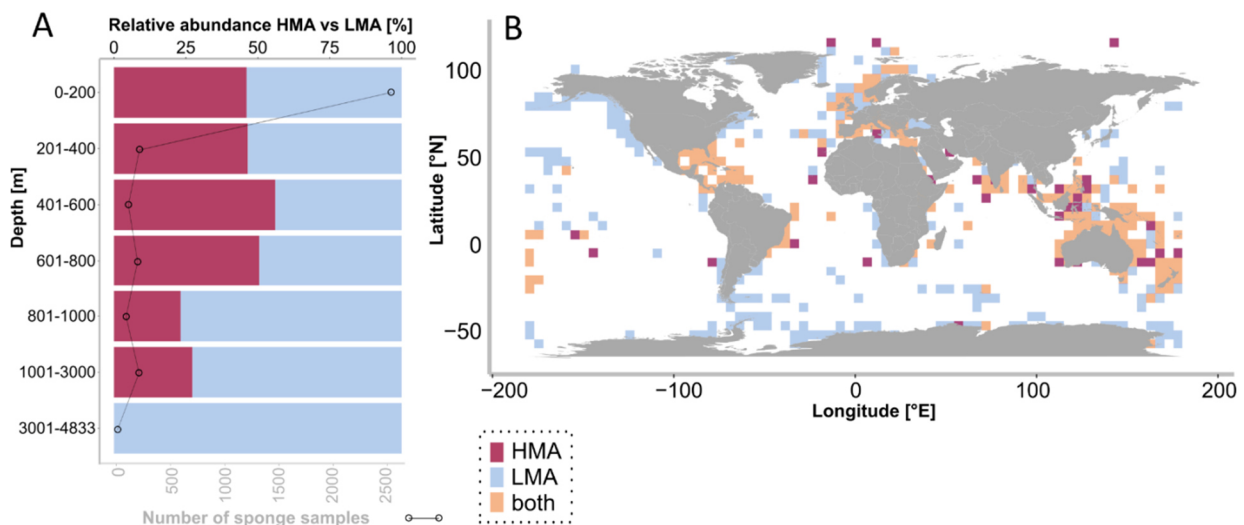


Fig.16 Spatial distribution of the HMA-LMA dichotomy. **A)** Change in HMA-LMA proportions over depth, based on a joint dataset build on the SMP and the DSMP. **B)** Geographic distribution of the HMA-LMA status. 32717 sponges of 152 species are covered (5908 HMA, 26809 LMA). Occurrence data was derived from OBIS.

I observed a more wide-spread occurrence of LMA exclusive areas than HMA exclusive areas. Generally, LMA hotspots seem to coincide with high nutrient concentrations and low pH regions (but not with primary productivity). Interestingly, in the Southern Ocean a LMA exclusivity was observed, while in the Arctic three HMA exclusive areas were seen. Except from these three areas in the Arctic Ocean, HMA sponge hotspots in the global ocean seem to be distributed in a wide band along the equator. Their occurrence thus coincides with rather oligotrophic areas. I expect that observed patterns of HMA-LMA spatial distribution may be related to the specific physiological capacities of each group. As Whittaker (1972) formulated “Given a resource gradient [...] in a community, species evolve to use different parts of this gradient; competition between them is thereby reduced.” Comparing the ecological roles (**Table 2**) with the prevailing environmental conditions (**Fig.17**), it seems that the statement by Whittaker (1972) may apply to HMA and LMA sponges. Morganti *et al.* (2017) have shown nicely how HMA and LMA sponges may perform a trophic niche separation which facilitates co-existence in shallow waters. Mixed sponge grounds may therefore represent areas in which limits of competition vs cooperation are tested. Notably, mixed sponge grounds seem to be particularly present in shallow waters. The deep-sea is often regarded an extreme place, but

in this case I would assume that the shallow ocean is more extreme in terms of its competition and cooperation degree which may drive species distributions in space.

Host morphology

We could show statistically that microbiome composition is strongly related to sponge morphology. Several previous studies have pointed out that sponge morphology may be a relevant trait (Abelson *et al.*, 1993; Kötter *et al.*, 2002; DeGoeij *et al.*, 2017; Morganti *et al.*, 2019; Bart *et al.*, 2020b). In the frame of this thesis we classified the sampled sponges into 12 morphological categories. These were ball, vase, flabellate (cone or fan), stalked, stick, bushy (tubes), branching (no tubes), bowl, other (compact or fluffy). Ball-shaped, bowl-shaped, and compact sponges belonged almost exclusively to the HMA category. Vase-shape and bushy (tubes) sponges consisted only of LMA glass sponges. Finally, flabellate (fan), branching (no tubes), and encrusting sponges belonged exclusively to the LMA demosponges. The four remaining categories (flabellate (cone), stalked, stick, and other (fluffy)) were composed of a mixture of LMA glass and LMA demosponges. I suggest that the microbial community complexity and basic structure may be governed by how self-contained the respective sponge is. A larger area-to-volume ratio could suggest a stronger “openness” towards the seawater environment, while a smaller area-to-volume ratio may suggest a more “enclosed” system. A further support may be that several key parameters (i.e. carrying capacity, microbial heterogeneity, and alpha diversity) were negatively correlated with the AV-ratio. Sponge abundance data in the (deep) ocean are mainly generated based on video surveys (Van Soest *et al.*, 2007; Roberts *et al.*, 2009). By making use of our newly detected link between sponge morphology and the HMA-LMA status, it may be possible to (i) predict the microbial community composition based on ROV footage and to expand more into the direction of non-invasive biodiversity assessments, (ii) establish a trait-based approach based on sponge morphology, because it has been linked to distinct physiological functions of the holobiont (at least in shallow waters; **Table 2**) (iii) upscale sponge-associated microbial community composition data via (ROV) image-based machine learning techniques.

The importance of biological traits in the description and evaluation of ecosystem functioning is increasingly getting more attention in marine research (Kiørboe *et al.*, 2018). It is based on the idea to characterise organisms by essential quantitative traits which capture key aspects of diversity, instead of considering species taxonomically (Barton *et al.*, 2015). Based on the results of my PhD thesis, I conclude that trait-based approaches may be very powerful for an evaluation of deep-sea sponge associated microbial diversity in the ecosystem context. I can easily envision at least three traits to be considered for trait-based approaches in the sponge-microbiome context: (i) the HMA-LMA dichotomy, (ii) sponge morphology (iii) microbial phenotypes. While the trait-based approach is quite well established in the field of ecology, it has not really reached the field of microbiology, yet (but note e.g. the work of Martiny *et al.*, 2015 on microbial phenotypic traits). My thesis makes an attempt to move forward into this direction.

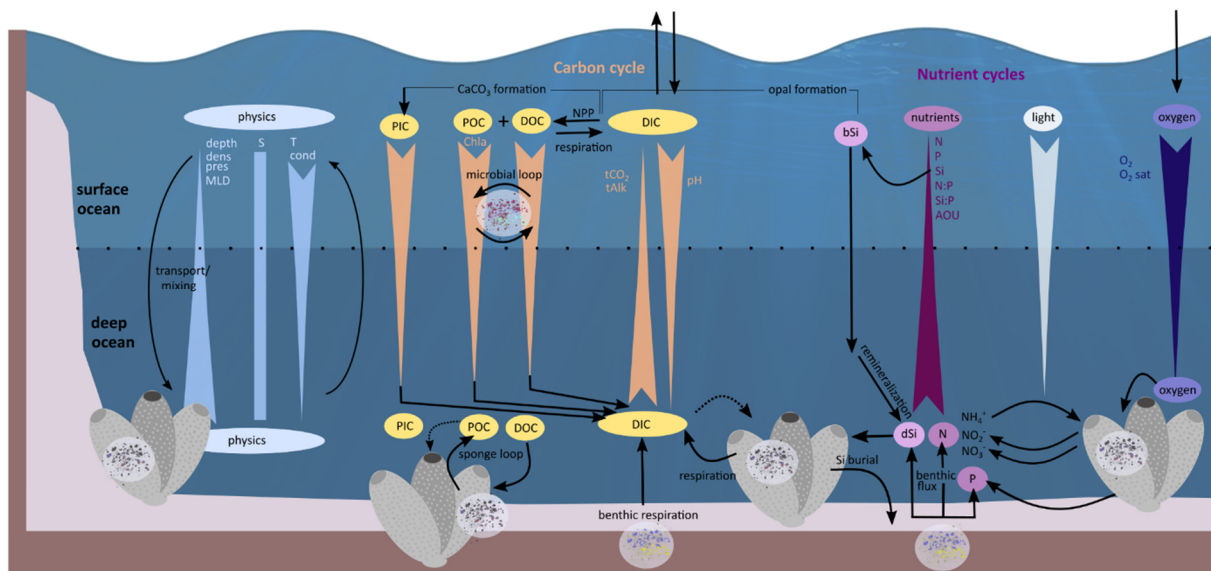
Vertical microbial transmission

We showed that entire microbial consortia can be vertically transmitted also in deep-sea sponge species (Busch *et al.*, 2020a). Interestingly, *Chloroflexi* were shown to be the dominant phylum across all life stages in the two analysed *Craniella* species, and we hypothesized that those ASVs which occur in all sponge life stages may fulfill important functions. In particular ASVs of the *SAR202 clade* may be particular good candidates for microbial key taxa. Future studies may check on their potential (key) functional role (e.g. in the cycling of recalcitrant organic matter). Overall, the symbiotic transmission model of a “leaky vertical transmission”, which is known from shallow water sponges, seems also plausible for deep-sea sponges, as we observed distinct microbial taxa in sponge embryos, which were not shared with the parent sponges. Generally, I would expect vertical transmission to play a larger role in deep-sea sponges compared to their shallow water counterparts, as abundances of pelagic microbes (i.e. the pool for horizontal acquisition) is considerably lower in the deep-sea. However, this hypothesis remains to be tested in future studies.

Environmental factors driving microbiome composition

Biogeochemical setting

One main aim of this thesis was to determine the main environmental drivers of microbial variability. **Fig.17** provides an overview of covered environmental parameters and their gradients throughout the water column and geographic distribution at the seafloor.



----- continues on next page -----

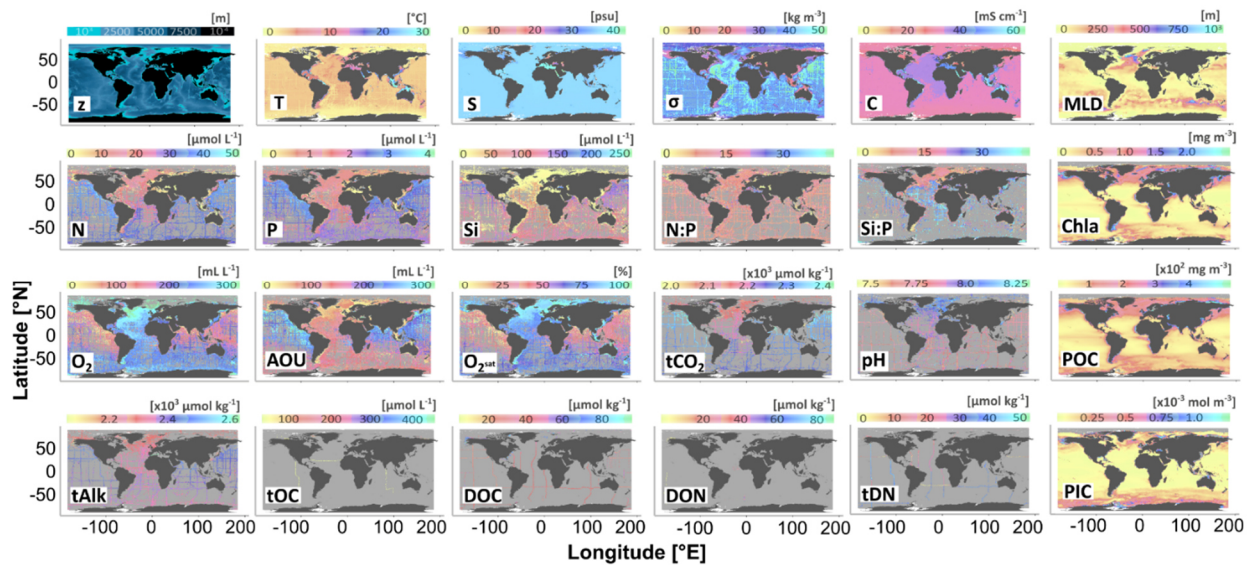


Fig.17 Resource gradients. **Previous page:** Vertical distribution of physical and biogeochemical parameters over depth, and major biogeochemical cycles of sponge holobionts and in the watercolumn. **This page:** Characteristics of physical and biogeochemical parameters at the seafloor on a global scale.

Our results presented in Busch *et al.*, (2020b) and **Chapter 7** both point into a similar direction: the four main identified environmental driving forces are temperature, salinity, depth, and nutrients/oxygen. These parameters explained around 25.3 % of the variability in HMA sponges, 14.2 % for LMA demosponges, and 16.4 % for LMA demosponges. The result of a higher percentage in HMA sponges may seem contradictory first because a closer coupling of LMA sponges to seawater has previously been observed (Moitinho-Silva *et al.*, 2014). However, I suspect that it is due to the fact that in HMA sponges more microbial phyla were occurring across multiple samples than in LMA sponges. One difficulty in determining the main environmental drivers, is the co-variation between the individual parameters. For example nutrients and oxygen are negatively correlated with each other. We therefore did not detect one single driver, but rather a set of environmental drivers. Targeted following-up experiments may help to tease those sets of environmental drivers further apart (e.g. the *nutrient* driver set, which consist of nitrate, silicate, and phosphate). Complementary analyses on the microbial functional level may be valuable in this respect.

We observed that co-occurrence patterns for the three analysed sponge species (*Geodia hentscheli*, *Lissodendoryx complicata*, and *Schaudinnia rosea*) were distinct from each other. We concluded that each host species shows an individual response to biogeochemical parameters. We further state that biological interactions might have masking effects. In particular we suspected that primary responders to abiotic parameters may have cascading effects on other microbes which are not directly affected by water biogeochemistry. In **Chapter 7** we worked with weighted gene correlation networks and determined microbial modules which are related to specific environmental parameters. We observed that the taxonomic composition of each module was fairly similar at high taxonomic ranks, no matter with which environmental parameter it is correlated. We conclude that the modular architecture

of the sponge microbiome, which had also been noted by (Lurgi *et al.*, 2019), provides stability. Although we characterised the interaction strength for both approaches, the co-occurrence networks and the weighted gene correlation networks, we cannot say whether interaction strength is overall stronger or weaker in the deep-sea. We observed that interactions between microbial taxa (determined based on co-occurrences) depend on the environmental context. For future studies it will be interesting to check whether the interactions between microbial taxa are dependent on the environmental context based on metagenomic analyses.

Population connectivity

With respect to population connectivity, we examined how well connected deep-sea sponge populations may be on a regional scale (< 100 km distance) (Busch *et al.*, 2020c). Our results revealed subtle sponge population genetic and pronounced microbial differences between a canyon and a bank. This observation may indicate that, although sponge and microbes form a metaorganism, the two partners are still individual evolving units. I suspect that microbes may be one step ahead of their host in differentiation processes. The presence of biophysical barriers (i.e., geological features or water masses) plays a crucial role in oceanographic connectivity. Our performed Lagrangian modelling illustrated a variable inter-annual connectivity between the three studied locations via ocean currents. Disrupted connectivity by biophysical barriers or reduced dispersal capacity can lead to an increase in richness. Along similar lines, I suspect that the enclosed morphology of HMA sponges may also promote an increased microbial richness inside their body. From my point of view the paradigm *everything is everywhere, but the environment selects* is an oversimplification which neglects the role of dispersal in structuring microbial communities. For deep-sea sponge associate microbes dispersal might happen either by the microbes themselves, or passively via dispersing sponge larvae. Due to the observed heterogeneity of the deep-sea observed in this thesis, I expect dispersal limitation to play a relevant role.

Seamounts as case study sites

In our publication (Busch *et al.*, 2020b) we discussed how a seamount can influence the microbial landscape in seawater and deep-sea sponges. We detected a seamount effect on the mid-water microbial community, which extended at least up to 200 m into the water column. Further, we observed distinct seawater microbial communities across the seamount which could not be explained by water masses alone. We suspect that the seamount ecosystem has an imprint on the near-bed water. Lastly, we observed that sponge microbiomes reflected water biogeochemistry. Our main take home message from this chapter is that a seamount can have an imprint on the microbial community composition of seawater and sponges. I suspect that the observed patterns originate from dispersal and environmental selection processes. In fact, in my opinion seamounts represent ideal sites to study the interdependency and intertwinedness of biotic and abiotic factors. Considering the widespread abundance of

seamounts in the global ocean (**Fig.18**), the results presented in Busch *et al.* (2020b) shed a new light on seamount microbiology.

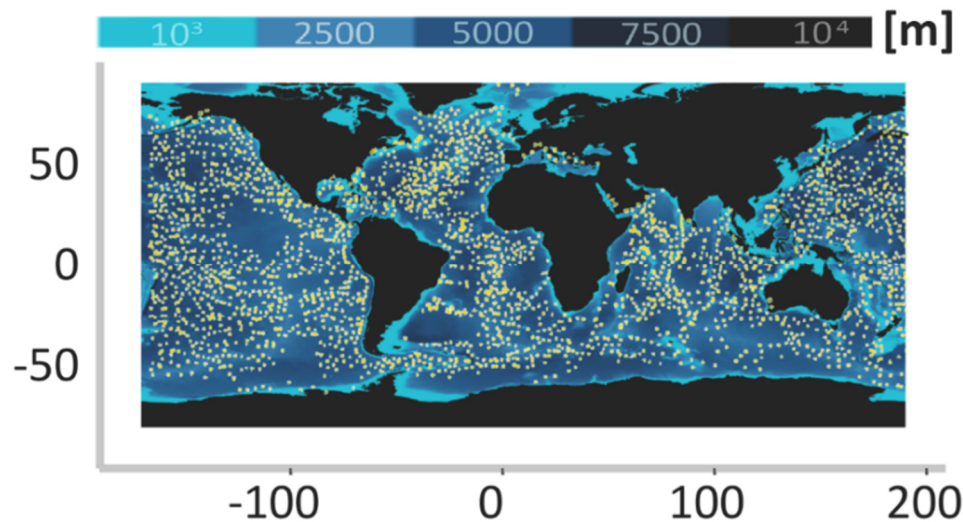


Fig.18 Global seamount distribution, replotted based on Danovaro *et al.*, 2014.

Conservation

Based on the determined main environmental drivers, trawling and plastic introduction effects on deep-sea sponge microbiomes have been studied within the frame of this thesis. In our publication Busch *et al.* (2020d) we discussed that biofilms on artificial substrates may ‘prime’ sponge-associated microbial communities when small sponges settle on such substrates. Results from this chapter further highlight that establishment of adequate marine conservation areas is a crucial task in order to also protect the microbial diversity thriving within sponges (along similar lines as Webster and Hill, 2014; Webster and Reusch, 2017; Webster *et al.*, 2018; Cavicchioli *et al.*, 2019). Within this thesis, deep-sea sponges were identified as microbial biodiversity hotspots in the deep ocean. Our network analyses in **Chapter 7** revealed that particularly the Schulz Bank seamount has a high within-module degree, as well as a high between-module degree, making it a good candidate for protection efforts. Within the context of this thesis we established a baseline and reference catalogue (**Chapter 7, Chapter 8**). We observed a high natural variability of deep-sea sponge microbiomes and provide *microbial reference spectra* as a baseline, with which future (more localised) studies can assess if their observed microbial variability falls within the natural range of variability or if it is extraordinary. The SVAmPEX tool (**Chapter 8**) may also help to expand further on trait-based approaches, as it offers the opportunity to link individual taxa to environmental conditions and thereby predict phenotypic microbial traits. On a different note, SVAmPEX helps to detect diseased sponges early on. I can also imagine that the platform may help in the design of a suitable microbial conservation approach, as it can help to answer questions such as the following ones: Where/ which sponges are biodiversity hotspots? Where/ in which sponges can key microbes/rare microbes be found? What would be a mindful biodiversity conservation approach: Does it make sense to protect certain areas / species ?

Conclusions

In this thesis I provide a first comprehensive overview on the diversity, evolution, biogeography, and ecology of deep-sea sponge microbiomes. A trait-based approach could be established to predict microbial biodiversity based on sponge morphology. In addition, a particular aim was to evaluate a potential coupling between biogeochemical as well as physical parameters and microbial community data. We could show that deep-sea sponge microbiomes indeed reflect water biogeochemistry and oceanographic patterns. Based on 20 deep-sea expeditions a large-scale reference collection was compiled with which applicability of theoretical concepts established for shallow waters were tested in the deep-sea environment. On the methodological side several novel resources were developed, such as a relational database covering an extensive set of metadata, a frozen sample collection, a high-throughput amplicon pipeline, and a software tool to foster operability of the created dataset. As this created dataset is very large it can be used in a variety of scopes and easily tailored towards the specific needs of future follow-up studies (**Fig.19**).

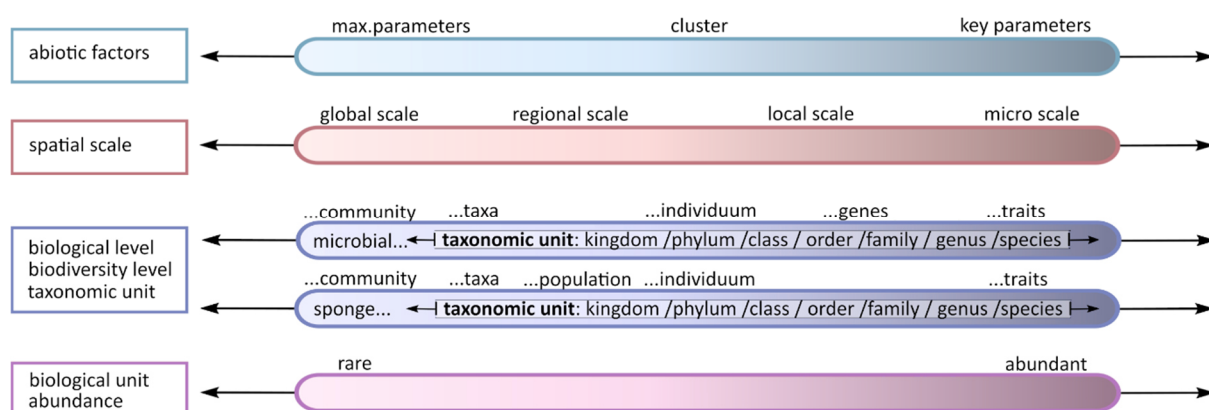


Fig.19 I had approached my thesis with the ambition to analyse deep-sea sponge microbiomes at different organisational scales. This figure illustrates the different hierarchical scales that were explored in the context of my thesis. Results of my thesis reveal that it is particularly useful to look at traits.

A large effort was put into standardisation of the dataset and making it machine readable. This should support the rapidly developing field of ocean data science. In stark contrast to the increasing phenomenon of particular disciplines “drowning in data”, data from the deep ocean are still scarce and urgently needed. Tools like the developed SVAmPEX software may help to make use of existing data more efficient. Today the deep-sea is the largest and still most understudied habitat on Earth. To derive at a holistic understanding and to capture ecosystem dynamics collaborative, interdisciplinary studies such as this thesis project are needed. Results of my thesis suggest a strong nestedness of deep-sea sponge microbiomes within the ecological context. In this regard the “closest layer”, the host, exerts the strongest impact.

I suspect that a high degree of interactivity and coupling between single ecological units promotes the observed high individuality of deep-sea sponge-associated microbiomes. A high variability in the lowest ecological unit however comes along with a trade-off in the detection of general patterns. Trait-based approaches are thus powerful to accomplish both, to consider individuality and enable upscaling. To this end, the generated data and information will provide the scientific basis for conservation of deep-sea sponge holobionts and deep-sea sponge ground ecosystems.

My PhD thesis arrives at the following main conclusions:

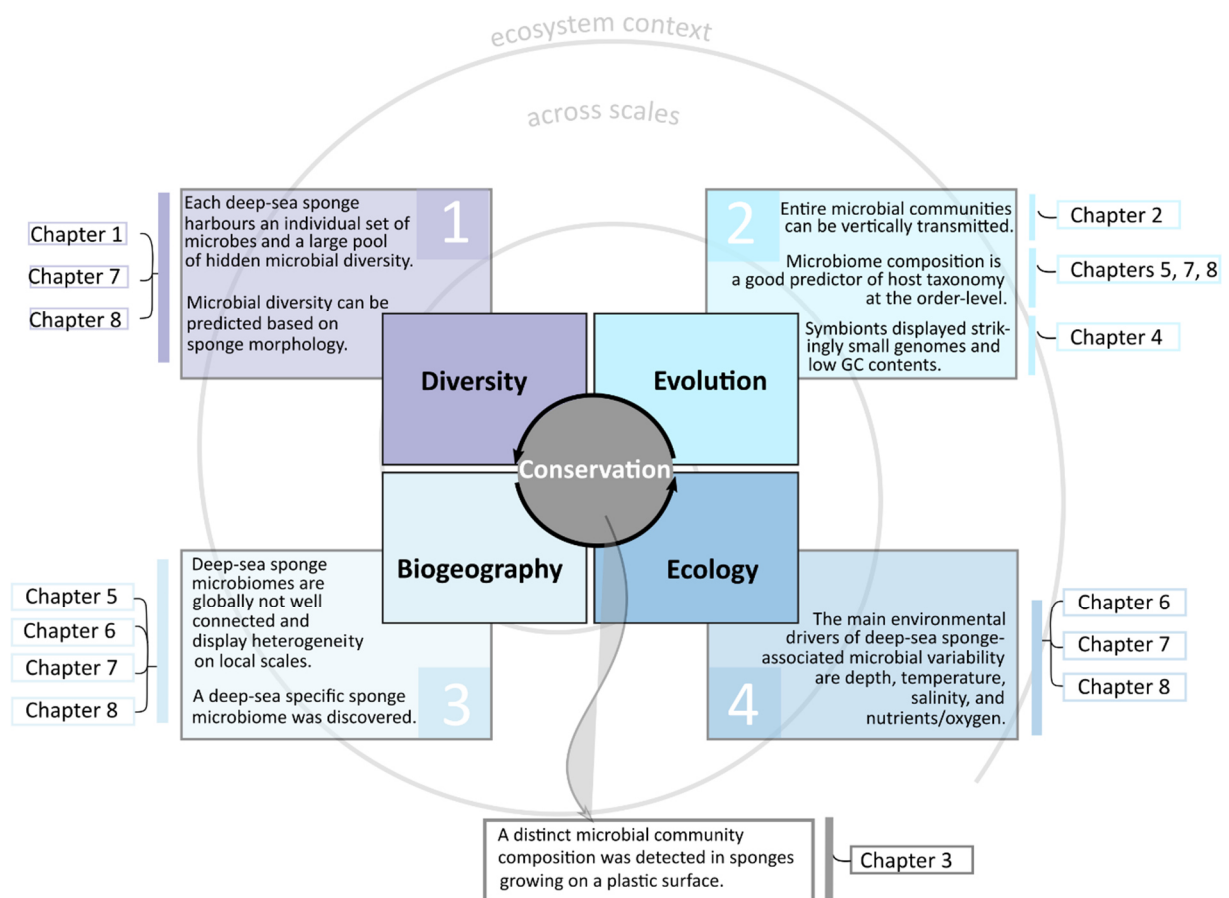


Fig.20 Major findings obtained on the diversity, evolution, biogeography, and ecology of deep-sea sponges and their microbiomes.

Acknowledgements

Danke **Ute Hentschel**, dass du meine geistige Beweglichkeit in den letzten Jahren gefördert hast. Danke für dein offenes Ohr und deinen offenen Geist, sowie für deine weisende oder manchmal auch einfach zum High Five bereit stehende Hand. Ich möchte dir für die vielseitigen Möglichkeiten danken und dafür, dass du mir den Raum gegeben hast eigene Ideen weiterzuentwickeln und während der Dissertation zu wachsen.

Arne Biastoch, vielen Dank für die Begleitung in den letzten Jahren. Für deine zielführenden, stets wohlwollenden und präzisen Anmerkungen sowie Einschätzungen sei herzlich gedankt. Ich bin dankbar, dass du meine Doktorarbeit mit nötigem Abstand beobachtet und unterstützt hast.

Hans Tore Rapp was one of the key players for making this work possible, a dear friend and mentor. His passing, just at a time when the manuscripts of this thesis finally started to shape up, was unexpected and devastating. I will treasure the good and valuable memories from the ROV back seat, when we flew together through the deep ocean.

Beate Slaby, aka die SponGES-Komplizin am GEOMAR. Vielen Dank für dein vielseitiges administratives Geschick und dein scharfes Auge auf alle SponGES bezogenen Termine und Deadlines. Danke auch für die Hygiene des *jazz-Servers* und die gemeinsame Zeit bei den SponGES-GAMs in London, Porto, und Wageningen.

Ina Clefsen und **Andrea Hethke**, herzlichen Dank für all die DNA-Extraktionen, eure Gewissenhaftigkeit und Zuverlässigkeit. Ich bin dankbar dass wir gemeinsam die Amplicon-Pipeline in Schwung bringen konnten und dass ihr alle meine Projekte im Labor unterstützt habt. **Corinna Bang** sei es hoch angerechnet, dass sie den Fehlerteufel damals bei den Hörnern gepackt hat.

Lucía Pita, you supported, advised, and enriched me and my work constantly throughout the years. When looking back, I realise now that (similar to a magician) you always provided exactly what was urgently needed at the respective time of the PhD. Often even without having been explicitly asked for help. Thank you for everything.

Danke **Lara Schmittmann**, für die bunte Tischgemeinschaft. Danke für all die spontanen von-Tisch-zu-Tisch Gespräche, die beflügelten Ideen, und dass wir gemeinsam die Schönheit in der Wissenschaft feiern konnten.

Erik Wurz, Danke für die regelmäßigen Gespräche und all die gemeinsamen lustigen, abenteuerlichen, und in letzter Zeit leider teilweise auch traurigen arbeitsbezogenen Aktionen. Ich würde dich jederzeit wieder an einer Autobahnraststätte zu geheimen Übergaben treffen.

Thank you to all members of the **RUMS working group**. How important and valuable a nice group of colleagues and a quick chat on the hallway is became so clear during the last 8 months, when working under physical distance due to the COVID-19 pandemic. I am thankful to have spend time with such a diverse group of intelligent people.

Eine globale Tiefseestudie aufzusetzen erfordert zunächst einmal viel Logistik. **Astrid Francke** sei gedankt für ihre kontinuierliche Unterstützung an der Poststelle und beherztes Handeln als alles schief lief. (Jeglichen Logistikunternehmen sei ausdrücklich nicht gedankt).

I would like to thank all **coauthors** for their contributions to the individual chapters of this thesis and for their support. Additional acknowledgements go to:

Furu Mienis for inspirational scientific discussions and an exciting time in San Diego.

Jasper de Goeij for lively and critical discussions about science and research, and for hospitality in Amsterdam.

Benjamin Mueller für die zahlreichen Gespräche, die entscheidende Unterstützung bei der Umsetzung von Chapter 6, und für die rettende Jacke in der Arktis als mein Koffer verloren ging.

Ana Riesgo and **Sergi Taboada** for keeping up the good spirit while working on a hard nut.

Javier Cristobo and **Pilar Ríos** for their hospitality in Gijón, their kindness and enthusiasm.

Peter Schupp für die spannende und ergebnisreiche Zeit auf der SONNE in und um Neuseeland.

Ellen Kenchington, **Lindsay Beazley**, and the **Canadian team** for bringing *Vazella* into our life, and for being so generous with knowledge and support that we ended up

writing three manuscripts on a sponge species which we have never met ourselves.

Joana Xavier for keeping the SponGES project going and for her editorial support with Chapter 2.

Paco Cárdenas for his critical eye, judgement, and feedback on Chapter 2 and Chapter 5.

I am thankful for methodological support by **Georg Steinert** (amplicon sequencing), **Christina Schmidt** (Lagrangian modelling), **Lucas Moitinho-Silva** (machine-learning), **Yu-Chen Wu** and **Cay Kruse** (TEM), **Álvaro Villalobos** and **Jan Michels** (SEM), **Martin Jahn** and **Kristina Bayer** (FISH).

For this thesis I spend in total almost half a year at sea.

Jackson Cahn, **Tessa Clemens**, **Julius Degenhardt**, **Asimena Gavriilidou**, **Helge-Ansgar Giebel**, **Vasiliki Koutsouveli**, **Sadie Mills**, **Venicio Pita**, **Marion Pohlner**, **Pedro Ribeiro**, **Emyr Martyn Roberts**, **Sven Rohde**, and **Detmer Sipkema** are acknowledged for making long working nights onboard ship go by fast, and for supporting my sampling activities.

Thank you for sending samples **Melanie Bergmann**, **Marie Creemers**, **Luisa Federwisch**, **Ruth Flerus**, and **Frederick Whoriskey**.

Our interaction strength in the context of this thesis was not too high, but I think still positive. Thanks for that **Martijn Bart**, **Francisca Carvalho**, **Ana Colaço**, **Ulrike Hanz**, **Anak Agung Gede Indraningrat**, **Anna de Kluijver**, **María López-Acosta**, **Manuel Maldonado**, **Heidi Kristina Meyer**, **Teresa Morganti**, **Anna Patova**, **Christopher Pham**, **Elena Prado**, **Francisco Sánchez**, **Karin Steffen**, and **Tanja Stratmann**.

This work was carried out at **GEOMAR, Helmholtz Centre for Ocean Research, Kiel** between August 2016 and December 2020. The research has been supported by the European Union's Horizon 2020 Research and Innovation Programme under grant agreement no. 679849 (the **SponGES project**).

Ein herzliches Dankeschön geht an meine Familie und an meine Freunde.
Ohne euch geht nichts.

References

8

- Abelson, A., Miloh, T., and Loya, Y. (1993) Flow patterns induced by substrata and body morphologies of benthic organisms, and their roles in determining availability of food particles. *Limnol Oceanogr* **38**: 1116–1124.
- Alexander, B.E., Liebrand, K., Osinga, R., Van Der Geest, H.G., Admiraal, W., Cleutjens, J.P.M., et al. (2014) Cell turnover and detritus production in marine sponges from tropical and temperate benthic ecosystems. *PLoS One* **9**: e109486.
- Alexander, B.E., Achlatis, M., Osinga, R., Van der Geest, H.G., Cleutjens, J.P.M., Schutte, B., and De Goeij, J.M. (2015) Cell kinetics during regeneration in the sponge *Halisarca caerulea*: How local is the response to tissue damage? *PeerJ* **2015**: 1–19.
- Anderson, S.A., Northcote, P.T., and Page, M.J. (2010) Spatial and temporal variability of the bacterial community in different chemotypes of the New Zealand marine sponge *Mycale hentscheli*. *FEMS Microbiol Ecol* **72**: 328–342.
- Anson, R., Romano, S., Sayavedra, L., Porras, M.Á.G., Kupczok, A., Tegetmeyer, H.E., et al. (2019) Functional diversity enables multiple symbiont strains to coexist in deep-sea mussels. *Nat Microbiol* **4**: 2487–2497.
- Appeltans, W., Ah Yong, S.T., Anderson, G., Angel, M. V., Artois, T., Bailly, N., et al. (2012) The magnitude of global marine species diversity. *Curr Biol* **22**: 2189–2202.
- Astudillo-García, C., Bell, J.J., Webster, N.S., Glasl, B., Jompa, J., Montoya, J.M., and Taylor, M.W. (2017) Evaluating the core microbiota in complex communities: a systematic investigation. *Environ Microbiol* **19**: 1450–1462.
- Azam, F., Fenchel, T., Field, J., Gray, J., Meyer-Reil, L., and Thingstad, F. (1983) The Ecological Role of Water-Column Microbes in the Sea. *Mar Ecol Prog Ser* **10**: 257–263.
- Baas Becking, L.G.M. (1934) *Geobiologie of Inleiding tot de Milieukunde*, Van Stockum & Zoon, The Hague.
- Bang, C., Dagan, T., Deines, P., Dubilier, N., Duschl, W.J., Fraune, S., et al. (2018) Metaorganisms in extreme environments: do microbes play a role in organismal adaptation? *Zoology* **127**: 1–19.
- Barker Jørgensen, C., Møhlenberg, F., and Sten-Knudsen, O. (1986) Nature of relation between ventilation and oxygen consumption in filter feeders. *Mar Ecol Prog Ser* **29**: 73–88.
- Bart, M.C., de Kluijver, A., Hoetjes, S., Absalah, S., Mueller, B., Kenchington, E., et al. (2020a) Differential processing of dissolved and particulate organic matter by deep-sea sponges and their microbial symbionts. *Sci Rep* **10**: 1–13.
- Bart, M.C., Mueller, B., Rombouts, T., van de Ven, C., Tompkins, G.J., Osinga, R., et al. (2020b) Dissolved organic carbon (DOC) is essential to balance the metabolic demands of North-Atlantic deep-sea sponges. *bioRxiv Prepr.*
- Barton, A.D., Dutkiewicz, S., Andersen, K.H., Fiksen, Ø.Ø.F., Follows, M.J., Mouw, C.B., et al. (2015) Report on the “Trait-based approaches to ocean life” scoping workshop October 5-8, 2015. doi:10.1575/1912/8017.
- Bayer, K., Moitinho-Silva, L., Brümmer, F., Cannistraci, C. V., Ravasi, T., and Hentschel, U. (2014a) GeoChip-based insights into the microbial functional gene repertoire of marine sponges (high microbial abundance, low microbial abundance) and seawater. *FEMS Microbiol Ecol* **90**: 832–843.
- Bayer, K., Kamke, J., and Hentschel, U. (2014b) Quantification of bacterial and archaeal symbionts in high and low microbial abundance sponges using real-time PCR. *FEMS Microbiol Ecol* **89**: 679–690.
- Bayer, K., Jahn, M.T., Slaby, B.M., Moitinho-Silva, L., and Hentschel, U. (2018) Marine sponges as *Chloroflexi* hot spots: Genomic insights and high-resolution visualization of an abundant and diverse symbiotic clade. *mSystems* **3**: 1–19.
- Bayer, K., Busch, K., Kenchington, E., Beazley, L., Franzenburg, S., Michels, J., et al. (2020) Microbial strategies for survival in the glass sponge *Vazella pourtalesii*. *mSystems* **5**: 1–20.
- Bell, J.J. (2007) Contrasting patterns of species and functional composition of coral reef sponge assemblages. *Mar Ecol Prog Ser* **339**: 73–81.
- Bennett, H.M., Altenrath, C., Woods, L., Davy, S.K., Webster, N.S., and Bell, J.J. (2017) Interactive effects of temperature and pCO₂ on sponges: from the cradle to the grave. *Glob Chang Biol* **23**: 2031–2046.
- Bhushan, A., Peters, E.E., and Piel, J. (2017) *Entotheonella* bacteria as source of sponge-derived natural products: opportunities for biotechnological production. In Müller, W., Schröder, H., and Wang, X. (eds), *Blue Biotechnology. Progress in Molecular and Subcellular Biology*. Springer, Cham.
- Björk, J.R., Díez-Vives, C., Coma, R., Ribes, M., and Montoya, J. (2013) Specificity and temporal dynamics of complex bacteria-sponge symbiotic interactions. *PLoS One* **8**: e72791.
- Björk, J., O’Hara, R., Ribes, M., Coma, R., and Montoya, J. (2017) The dynamic core microbiome: Structure, dynamics and stability. *bioRxiv* 137885.
- Björk, J.R., Díez-Vives, C., Astudillo-García, C., Archie, E.A., and Montoya, J.M. (2019) Vertical transmission of sponge microbiota is inconsistent and unfaithful. *Nat Ecol Evol* **3**: 1172–1183.
- Blanquer, A., Uriz, M.J., and Galand, P.E. (2013) Removing environmental sources of variation to gain insight on symbionts vs. transient microbes in high and low microbial abundance sponges. *Environ Microbiol* **15**: 3008–3019.
- Blunt, J.W., Copp, B.R., Munro, M.H.G., Northcote, P.T., and Prinsep, M.R. (2006) Marine natural products. *Nat Prod Rep* **23**: 26–78.
- Borchert, E., Selvin, J., Kiran, S.G., Jackson, S.A., O’Gara, F., and Dobson, A.D.W. (2017) A novel cold active esterase from a deep sea sponge *Stelletta normani* metagenomic library. *Front Mar Sci* **4**: 1–13.
- Botté, E.S., Nielsen, S., Abdul Wahab, M.A., Webster, J., Robbins, S., Thomas, T., and Webster, N.S. (2019) Changes in the metabolic potential of the sponge microbiome under ocean acidification. *Nat Commun* **10**: 1–10.
- Boury-Esnault, N., de Vos, L., Donadey, C., and Vacelet, J. (1990) Ultrastructure of choanosome and sponge classification. In Rützler, K. (ed), *New Perspectives in Sponge Biology*. Smithsonian Books, 237–244.
- Breusing, C., Biastoch, A., Drews, A., Metaxas, A., Jollivet, D., Vrijenhoek, R.C., et al. (2016) Biophysical and population genetic models predict the presence of “phantom” stepping stones connecting Mid-Atlantic Ridge vent ecosystems. *Curr Biol* **26**: 2257–2267.
- Burgsdorf, I., Erwin, P.M., López-Legentil, S., Cerrano, C., Haber, M., Frenk, S., and Steindler, L. (2014) Biogeography rather than association with *Cyanobacteria* structures symbiotic microbial communities in the marine sponge *Petrosia ficiformis*. *Front Microbiol* **5**: 1–11.
- Burgsdorf, I., Handley, K.M., Bar-Shalom, R., Erwin, P.M., and Steindler, L. (2019) Life at home and on the roam: Genomic adaptations reflect the dual lifestyle of an intracellular, facultative symbiont. *mSystems* **4**: 1–19.
- Busch, K., Endres, S., Iversen, M.H., Michels, J., Nöthig, E.-M., and Engel, A. (2017) Bacterial Colonization and Vertical Distribution of Marine Gel Particles (TEP and CSP) in the Arctic Fram Strait. *Front Mar Sci* **4**: 1–14.
- Busch, K., Wurz, E., Rapp, H.T., Bayer, K., Franke, A., and Hentschel, U. (2020a) *Chloroflexi* dominate the deep-sea golf ball sponges *Craniella zetlandica* and *Craniella infrequens* throughout different life stages. *Front Mar Sci* **7**: 1–12.
- Busch, K., Hanz, U., Mienis, F., Mueller, B., Franke, A., Martyn Roberts, E., et al. (2020b) On giant shoulders: How a seamount affects the microbial community composition of seawater and sponges. *Biogeosciences* **17**: 3471–3486.

- Busch, K., Taboada, S., Riesgo, A., Koutsouveli, V., Ríos, P., Cristobo, J., Franke, A., Getzlaff, K., Schmidt, C., Biastoch, A., Hentschel, U. (2020c) Population connectivity of fan-shaped sponge holobionts in the deep Cantabrian Sea. *Deep Sea Res. Part I Oceanogr.* doi:10.1016/j.dsr.2020.10342.
- Busch, K., Beazley, L., Kenchington, E., Whoriskey, F., Slaby, B.M., and Hentschel, U. (2020d) Microbial diversity of the glass sponge *Vazella pourtalesii* in response to anthropogenic activities. *Conserv Genet* **21**: 1001–1010.
- Buttigieg, P.L., Morrison, N., Smith, B., Mungall, C.J., and Lewis, S.E. (2013) The environment ontology: Contextualising biological and biomedical entities. *J Biomed Semantics* **4**: 1–9.
- Callahan, B.J., McMurdie, P.J., and Holmes, S.P. (2017) Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME J* **11**: 2639–2643.
- Campbell, N.A. and Reece, J.B. (2009) *Biologie*, 8th ed. Kratochwil, A., Scheibe, R., and Wiczorek, H. (eds) Pearson, Deutschland.
- Cattaneo-Vietti, R., Chiantore, M., Mistic, C., Povero, P., and Fabiano, M. (1999) The role of pelagic-benthic coupling in structuring littoral benthic communities at Terra Nova Bay (Ross Sea) and in the Straits of Magellan. *Sci Mar* **63**: 113–121.
- Cavicchioli, R., Ripple, W.J., Timmis, K.N., Azam, F., Bakken, L.R., Baylis, M., et al. (2019) Scientists' warning to humanity: microorganisms and climate change. *Nat Rev Microbiol* **17**: 569–586.
- Cebrian, E., Uriz, M.J., and Turon, X. (2007) Sponges as biomonitors of heavy metals in spatial and temporal surveys in northwestern Mediterranean: Multispecies comparison. *Environ Toxicol Chem* **26**: 2430–2439.
- Cleary, D.F.R., Becking, L.E., de Voogd, N.J., Pires, A.C.C., Polónia, A.R.M., Egas, C., and Gomes, N.C.M. (2013) Habitat- and host-related variation in sponge bacterial symbiont communities in Indonesian waters. *FEMS Microbiol Ecol* **85**: 465–482.
- Cleary, D.F.R., Swierts, T., Coelho, F.J.R.C., Polónia, A.R.M., Huang, Y.M., Ferreira, M.R.S., et al. (2019) The sponge microbiome within the greater coral reef microbial metacommunity. *Nat Commun* **10**: 1–12.
- Cleary, D.F.R., Ferreira, M.R.S., Bat, N.K., Polónia, A.R.M., Gomes, N.C.M., and de Voogd, N.J. (2020) Bacterial composition of sponges, sediment and seawater in enclosed and open marine lakes in Ha Long Bay Vietnam. *Mar Biol Res* **16**: 18–31.
- Corinaldesi, C. (2015) New perspectives in benthic deep-sea microbial ecology. *Front Mar Sci* **2**: 1–12.
- Corredor, J.E., Wilkinson, C.R., Vicente, V.P., Morell, J.M., and Otero, E. (1988) Nitrate release by Caribbean reef sponges. *Limnol Oceanogr* **33**: 114–120.
- Costello, M.J. and Chaudhary, C. (2017) Marine biodiversity, biogeography, deep-sea gradients, and conservation. *Curr Biol* **27**: R511–R527.
- Curtis, T.P., Sloan, W.T., and Scannell, J.W. (2002) Estimating prokaryotic diversity and its limits. *Proc Natl Acad Sci* **99**: 10494–10499.
- Danovaro, R. (2009) *Methods for the study of deep-sea sediments, their functioning and biodiversity*, CRC Press, Boca Raton.
- Danovaro, R., Snelgrove, P.V.R., and Tyler, P. (2014) Challenging the paradigms of deep-sea ecology. *Trends Ecol Evol* **29**: 465–475.
- Davis, A.R., de Mestre, C., Maher, W., Krikowa, F., and Broad, A. (2014) Sponges as sentinels: Metal accumulation using transplanted sponges across a metal gradient. *Environ Toxicol Chem* **33**: 2818–2825.
- DeLong, E.F. (1992) Archaea in coastal marine environments. *Proc Natl Acad Sci* **89**: 5685–5689.
- Diaz, M.C. and Ward, B.B. (1997) Sponge-mediated nitrification in tropical benthic communities. *Mar Ecol Prog Ser* **156**: 97–107.
- van Dover, C.L. (2000) *The ecology of deep-sea hydrothermal vents*, Princeton University Press, Princeton, New Jersey.
- Dubilier, N., Bergin, C., and Lott, C. (2008) Symbiotic diversity in marine animals: The art of harnessing chemosynthesis. *Nat Rev Microbiol* **6**: 725–740.
- van Duyl, F.C., Hegeman, J., Hoogstraten, A., and Maier, C. (2008) Dissolved carbon fixation by sponge-microbe consortia of deep water coral mounds in the northeastern Atlantic Ocean. *Mar Ecol Prog Ser* **358**: 137–150.
- van Duyl, F.C., Lengger, S.K., Schouten, S., Lundälv, T., van Oevelen, D., and Müller, C.E. (2020) Dark CO₂ fixation into phospholipid-derived fatty acids by the cold-water coral associated sponge *Hymedesmia (Stylopus) coriacea* (Tisler Reef, NE Skagerrak). *Mar Biol Res* **16**: 1–17.
- Easson, C.G. and Thacker, R.W. (2014) Phylogenetic signal in the community structure of host-specific microbiomes of tropical marine sponges. *Front Microbiol* **5**: 532.
- Eme, L., Spang, A., Lombard, J., Stairs, C.W., and Ettema, T.J.G. (2017) Archaea and the origin of eukaryotes. *Nat Rev Microbiol* **15**: 711–723.
- Erpenbeck, D., Breeuwer, J.A.J., Van der Velde, H.C., and Van Soest, R.W.M. (2002) Unravelling host and symbiont phylogenies of halichondrid sponges (Demospongiae, Porifera) using a mitochondrial marker. *Mar Biol* **141**: 377–386.
- Erwin, P.M. and Thacker, R.W. (2008) Cryptic diversity of the symbiotic cyanobacterium *Synechococcus spongiarum* among sponge hosts. *Mol Ecol* **17**: 2937–2947.
- Erwin, P.M., Olson, J.B., and Thacker, R.W. (2011) Phylogenetic diversity, host-specificity and community profiling of sponge-associated bacteria in the northern Gulf of Mexico. *PLoS One* **6**: e26806.
- Erwin, P.M., Pita, L., López-Legentil, S., and Turon, X. (2012) Stability of sponge-associated bacteria over large seasonal shifts in temperature and irradiance. *Appl Environ Microbiol* **78**: 7358–7368.
- Erwin, P.M., Coma, R., López-Sendino, P., Serrano, E., and Ribes, M. (2015) Stable symbionts across the HMA-LMA dichotomy: Low seasonal and interannual variation in sponge-associated bacteria from taxonomically diverse hosts. *FEMS Microbiol Ecol* **91**: 1–11.
- Esteves, A.I.S., Cullen, A., and Thomas, T. (2017) Competitive interactions between sponge-associated bacteria. *FEMS Microbiol Ecol* **93**: 1–8.
- Fan, L., Reynolds, D., Liu, M., Stark, M., Kjelleberg, S., Webster, N.S., and Thomas, T. (2012) Functional equivalence and evolutionary convergence in complex communities of microbial sponge symbionts. *Proc Natl Acad Sci* **109**: E1878–E1887.
- Fan, L., Liu, M., Simister, R., Webster, N.S., and Thomas, T. (2013) Marine microbial symbiosis heats up: the phylogenetic and functional response of a sponge holobiont to thermal stress. *ISME J* **7**: 991–1002.
- Fierer, N. (2008) Microbial biogeography: patterns in microbial diversity across space and time. In Zengler, K. (ed), *Accessing uncultivated microorganisms: from the environment to organisms and genomes and back*. ASM Press, Washington DC., 95–115.
- Fieseler, L., Horn, M., Wagner, M., and Hentschel, U. (2004) Discovery of the novel candidate phylum "Poribacteria" in marine sponges. *Appl Environ Microbiol* **70**: 3724–3732.
- Fiore, C.L., Jarett, J.K., and Lesser, M.P. (2013) Symbiotic prokaryotic communities from different populations of the giant barrel sponge, *Xestospongia muta*. *MicrobiologyOpen* **2**: 938–952.
- Fox, A.D., Henry, L.A., Corne, D.W., and Roberts, J.M. (2016) Sensitivity of marine protected area network connectivity to atmospheric variability. *R Soc Open Sci* **3**: 160494.
- Franke, A., Blenckner, T., Duarte, C.M., Ott, K., Fleming, L.E., Antia, A., et al. (2020) Operationalizing ocean health: Toward integrated research on ocean health and recovery to achieve ocean sustainability. *One Earth* **2**: 557–565.
- Freckelton, M.L., Luter, H.M., Andreakis, N., Webster, N.S., and Motti, C.A. (2012) Qualitative variation in colour morphotypes of *lanthella basta* (Porifera: Verongida). *Hydrobiologia* **687**: 191–203.
- Freeman, C.J. and Thacker, R.W. (2011) Complex interactions between marine sponges and their symbiotic microbial communities. *Limnol Oceanogr* **56**: 1577–1586.
- Freeman, C.J., Easson, C.G., and Baker, D.M. (2014) Metabolic diversity and niche structure in sponges from the Miskito Cays, Honduras. *PeerJ* **2014**: 1–20.
- Freeman, M.F., Vagstad, A.L., and Piel, J. (2016) Polytheonamide biosynthesis showcasing the metabolic potential of sponge-associated uncultivated "Entotheonella" bacteria. *Curr Opin Chem Biol* **31**: 8–14.

- Friedrich, A.B., Merkert, H., Fendert, T., Hacker, J., Proksch, P., and Hentschel, U. (1999) Microbial diversity in the marine sponge *Aplysina cavernicola* (formerly *Verongia cavernicola*) analyzed by fluorescence in situ hybridization (FISH). *Mar Biol* **134**: 461–470.
- Frost, T.M. (1976) Sponge feeding: A review with a discussion of some continuing research. In Harrison, F.W. and Cowden, R.R. (eds), *Aspects of Sponge Biology*. Academic Press, New York, 283–298.
- Frost, T.M. (1980) Selection in sponge feeding processes. In Smith, D.C. and Tiffon, Y. (eds), *Nutrition in the Lower Metazoa*. Pergamon Press, Oxford, 33–44.
- Fuhrman, J.A., McCallum, K., and Davis, A. (1992) Novel major archaeobacterial group from marine plankton. *Nature* **359**: 148–149.
- Gantt, S.E., McMurray, S.E., Stubler, A.D., Finelli, C.M., Pawlik, J.R., and Erwin, P.M. (2019) Testing the relationship between microbiome composition and flux of carbon and nutrients in Caribbean coral reef sponges. *Microbiome* **7**: 1–13.
- Giles, E.C., Kamke, J., Moitinho-Silva, L., Taylor, M.W., Hentschel, U., Ravasi, T., and Schmitt, S. (2012) Bacterial community profiles in low microbial abundance sponges. *FEMS Microbiol Ecol* **83**: 232–241.
- Gili, J.M. and Coma, R. (1998) Benthic suspension feeders: Their paramount role in littoral marine food webs. *Trends Ecol Evol* **13**: 316–321.
- Giovannoni, S.J. (2017) SAR11 Bacteria: The most abundant plankton in the oceans. *Ann Rev Mar Sci* **9**: 231–255.
- Girard, E.B., Fuchs, A., Kaliwoda, M., Lasut, M., Ploetz, E., Schmahl, W.W., and Wörheide, G. (2020) Sponges as bioindicators for microparticulate pollutants? *Environ Pollut* doi: 10.1016/j.envpol.2020.115851.
- Glasl, B., Webster, N.S., and Bourne, D.G. (2017) Microbial indicators as a diagnostic tool for assessing water quality and climate stress in coral reef ecosystems. *Mar Biol* **164**: 91.
- Glasl, B., Smith, C.E., Bourne, D.G., and Webster, N.S. (2018) Exploring the diversity-stability paradigm using sponge microbial communities. *Sci Rep* **8**: 1–9.
- Glasl, B., Bourne, D.G., Frade, P.R., Thomas, T., Schaffelke, B., and Webster, N.S. (2019) Microbial indicators of environmental perturbations in coral reef ecosystems. *Microbiome* **7**: 1–13.
- Gloeckner, V., Wehrl, M., Moitinho-silva, L., Schupp, P., Pawlik, J.R., Lindquist, N.L., et al. (2014) The HMA-LMA dichotomy revisited: an electron microscopical survey of 56 sponge species. *Biol Bull* **227**: 78–88.
- De Goeij, J.M., Van Den Berg, H., Van Oostveen, M.M., Epping, E.H.G., and Van Duyl, F.C. (2008a) Major bulk dissolved organic carbon (DOC) removal by encrusting coral reef cavity sponges. *Mar Ecol Prog Ser* **357**: 139–151.
- De Goeij, J.M., Moodley, L., Houtekamer, M., Carballeira, N.M., and Van Duyl, F.C. (2008b) Tracing ¹³C-enriched dissolved and particulate organic carbon in the bacteria-containing coral reef sponge *Halisarca caerulea*: Evidence for DOM feeding. *Limnol Oceanogr* **53**: 1376–1386.
- De Goeij, J.M., De Kluijver, A., Van Duyl, F.C., Vacelet, J., Wijffels, R.H., De Goeij, A.F.R.M., et al. (2009) Cell kinetics of the marine sponge *Halisarca caerulea* reveal rapid cell turnover and shedding. *J Exp Biol* **212**: 3892–3900.
- De Goeij, J.M., Van Oevelen, D., Vermeij, M.J.A., Osinga, R., Middelburg, J.J., De Goeij, A.F.P.M., and Admiraal, W. (2013) Surviving in a marine desert: The sponge loop retains resources within coral reefs. *Science* **342**: 108–110.
- De Goeij, J.M., Lesser, M.P., and Pawlik, J.R. (2017) Nutrient fluxes and ecological functions of coral reef sponges in a changing ocean. In Carballo, J.L. and Bell, J.J. (eds), *Climate Change, Ocean Acidification and Sponges: Impacts Across Multiple Levels of Organization*. Springer, Cham, 373–410.
- Greenspan, S.E., Migliorini, G.H., Lyra, M.L., Pontes, M.R., Carvalho, T., Ribeiro, L.P., et al. (2020) Warming drives ecological community changes linked to host-associated microbiome dysbiosis. *Nat Clim Chang* doi: 10.1038/s41558-020-0899-5.
- Gutleben, J., Loureiro, C., Ramirez Romero, L.A., Shetty, S., Wijffels, R.H., Smidt, H., and Sipkema, D. (2020) Cultivation of bacteria from *Aplysina aerophoba*: Effects of oxygen and nutrient gradients. *Front Microbiol* **11**: 1–19.
- Hadas, E., Marie, D., Shpigel, M., and Ilan, M. (2006) Virus predation by sponges is a new nutrient-flow pathway in coral reef food webs. *Limnol Oceanogr* **51**: 1548–1550.
- Hadas, E., Shpigel, M., and Ilan, M. (2009) Particulate organic matter as a food source for a coral reef sponge. *J Exp Biol* **212**: 3643–3650.
- Hallam, S.J., Konstantinidis, K.T., Putnam, N., Schleper, C., Watanabe, Y.I., Sugahara, J., et al. (2006) Genomic analysis of the uncultivated marine crenarchaeote *Cenarchaeum symbiosum*. *Proc Natl Acad Sci* **103**: 18296–18301.
- Han, B., Hong, L., Gu, B., Sun, Y., Wang, J., Liu, J.-T., and Lin, H.-W. (2019) Natural products from sponges. In *Symbiotic Microbiomes of Coral Reefs Sponges and Corals*. Springer Nature, 329–445.
- Hanson, C.A., Fuhrman, J.A., Horner-Devine, M.C., and Martiny, J.B.H. (2012) Beyond biogeographic patterns: Processes shaping the microbial landscape. *Nat Rev Microbiol* **10**: 497–506.
- Hanz, U., Wienberg, C., Hebbeln, D., Duineveld, G., Lavaleye, M., Juva, K., et al. (2019) Environmental factors influencing benthic communities in the oxygen minimum zones on the Angolan and Namibian margins. *Biogeosciences* **16**: 4337–4356.
- Hentschel, U., Hopke, J., Horn, M., Friedrich, A.B., Wagner, M., Hacker, J., Moore, B.S. (2002) Molecular evidence for a uniform microbial community in sponges from different oceans. *Appl Environ Microbiol* **68**: 4431–4440.
- Hentschel, U., Fieseler, L., Wehrl, M., Gernert, C., Steinert, M., Hacker, J., and Horn, M. (2003) Microbial diversity of marine sponges. In Müller, W.E.G. (ed), *Molecular Marine Biology of Sponges*. Springer, Heidelberg, 60–88.
- Hentschel, U., Usher, K.M., and Taylor, M.W. (2006) Marine sponges as microbial fermenters. *FEMS Microbiol Ecol* **55**: 167–177.
- Hentschel, U., Piel, J., Degnan, S.M., and Taylor, M.W. (2012a) Genomic insights into the marine sponge microbiome. *Nat Rev Microbiol* **10**: 641–654.
- Hentschel, U., Weis, V.M., and McFall-Ngai, M.J. (2012b) Biological bulletin virtual symposium: Discoveries in animal symbiosis in the “omics” age. *Biol Bull* **223**: 5–6.
- Hinzke, T., Kleiner, M., Breusing, C., Felbeck, H., Häsler, R., Sievert, S.M., et al. (2019) Host-microbe interactions in the chemosynthetic *Riftia pachyptila* symbiosis. *mBio* **10**: 1–20.
- Hochmuth, T., Niederkrüger, H., Gernert, C., Siegl, A., Taudien, S., Platzer, M., et al. (2010) Linking chemical and microbial diversity in marine sponges: Possible role for *Poribacteria* as producers of methyl-branched fatty acids. *ChemBioChem* **11**: 2572–2578.
- Hoer, D.R., Gibson, P.J., Tommerdahl, J.P., Lindquist, N.L., and Martens, C.S. (2017) Consumption of dissolved organic carbon by Caribbean reef sponges. *Limnol Oceanogr* **63**: 337–351.
- Hoer, D.R., Tommerdahl, J.P., Lindquist, N.L., and Martens, C.S. (2018) Dissolved inorganic nitrogen fluxes from common Florida Bay (U.S.A.) sponges. *Limnol Oceanogr* **63**: 2563–2578.
- Hoffmann, F., Radax, R., Woebken, D., Holtappels, M., Lavik, G., Rapp, H.T., et al. (2009) Complex nitrogen cycling in the sponge *Geodia barretti*. *Environ Microbiol* **11**: 2228–2243.
- Holmes, B. and Blanch, H. (2007) Genus-specific associations of marine sponges with *Group I crenarchaeotes*. *Mar Biol* **150**: 759–772.
- Jackson, S.A., Flemer, B., McCann, A., Kennedy, J., Morrissey, J.P., O’Gara, F., and Dobson, A.D.W. (2013) Archaea appear to dominate the microbiome of *Inflatella pellicula* deep sea sponges. *PLoS One* **8**: 1–8.
- Jahn, M.T., Markert, S.M., Ryu, T., Ravasi, T., Stigloher, C., Hentschel, U., and Moitinho-Silva, L. (2016) Shedding light on cell compartmentation in the candidate phylum *Poribacteria* by high resolution visualisation and transcriptional profiling. *Sci Rep* **6**: 35860.
- Jahn, M.T. (2019) Physiology, syntrophy and viral interplay in the marine sponge holobiont. Dissertation, Kiel University.
- Jiménez, E. and Ribes, M. (2007) Sponges as a source of dissolved inorganic nitrogen: Nitrification mediated by temperate sponges. *Limnol Oceanogr* **52**: 948–958.
- Jochum, K.P., Wang, X., Vennemann, T.W., Sinha, B., and Müller, W.E.G. (2012) Siliceous deep-sea sponge *Monorhaphis chuni*: A potential paleoclimate archive in ancient animals. *Chem Geol* **300–301**: 143–151.

- Johnson, S.B., Krylova, E.M., Audzijonyte, A., Sahling, H., and Vrijenhoek, R.C. (2017) Phylogeny and origins of chemosynthetic vesicomid clams. *Syst Biodivers* **15**: 346–360.
- Kahn, A.S., Yahel, G., Chu, J.W.F., Tunnicliffe, V., and Leys, S.P. (2015) Benthic grazing and carbon sequestration by deep-water glass sponge reefs. *Limnol Oceanogr* **60**: 78–88.
- Kahn, A.S. and Leys, S.P. (2017) Spicule and flagellated chamber formation in a growth zone of *Aphrocallistes vastus* (Porifera, Hexactinellida). *Invertebr Biol* **136**: 22–30.
- Kamke, J., Taylor, M.W., and Schmitt, S. (2010) Activity profiles for marine sponge-associated bacteria obtained by 16S rRNA vs 16S rRNA gene comparisons. *ISME J* **4**: 498–508.
- Kamke, J., Szczyrba, A., Ivanova, N., Schwientek, P., Rinke, C., Mavromatis, K., et al. (2013) Single-cell genomics reveals complex carbohydrate degradation patterns in poribacterial symbionts of marine sponges. *ISME J* **7**: 2287–2300.
- Kamke, J., Rinke, C., Schwientek, P., Mavromatis, K., Ivanova, N., Szczyrba, A., et al. (2014) The candidate phylum *Poribacteria* by single-cell genomics: New insights into phylogeny, cell-compartmentation, eukaryote-like repeat proteins, and other genomic features. *PLoS One* **9**: e87353.
- Kandler, N.M., Abdul Wahab, M.A., Noonan, S.H.C., Bell, J.J., Davy, S.K., Webster, N.S., and Luter, H.M. (2018) In situ responses of the sponge microbiome to ocean acidification. *FEMS Microbiol Ecol* **94**: fty205.
- Karimi, E., Ramos, M., Gonçalves, J.M.S., Xavier, J.R., Reis, M.P., and Costa, R. (2017) Comparative metagenomics reveals the distinctive adaptive features of the *Spongia officinalis* endosymbiotic consortium. *Front Microbiol* **8**: 1–16.
- Karimi, E., Keller-Costa, T., Slaby, B.M., Cox, C.J., da Rocha, U.N., Hentschel, U., and Costa, R. (2019) Genomic blueprints of sponge-prokaryote symbiosis are shared by low abundant and cultivatable *Alphaproteobacteria*. *Sci Rep* **9**: 1–15.
- Karner, M.B., Delong, E.F., and Karl, D.M. (2001) Archaeal dominance in the mesopelagic zone of the Pacific Ocean. *Nature* **409**: 507–510.
- Kennedy, J., Flemer, B., Jackson, S.A., Morrissey, J.P., O’Gara, F., and Dobson, A.D.W. (2014) Evidence of a putative deep sea specific microbiome in marine sponges. *PLoS One* **9**: 1–13.
- Kjørboe, T., Visser, A., Andersen, K.H., and Browman, H. (2018) A trait-based approach to ocean ecology. *ICES J Mar Sci* **75**: 1849–1863.
- Koch, E.J., Moriano-Gutierrez, S., Ruby, E.G., McFall-Ngai, M., and Liebecke, M. (2020) The impact of persistent colonization by *Vibrio fischeri* on the metabolome of the host squid *Euprymna scolopes*. *J Exp Biol* **223**: jeb212860.
- Könneke, M., Bernhard, A.E., De La Torre, J.R., Walker, C.B., Waterbury, J.B., and Stahl, D.A. (2005) Isolation of an autotrophic ammonia-oxidizing marine archaeon. *Nature* **437**: 543–546.
- Kötter, I. and Pernthaler, J. (2002) In situ feeding rates of obligate and facultative coelobite (cavity-dwelling) sponges in a Caribbean coral reef. In *Proc 9th Int Coral Reef Symp.*, 347–352.
- Kötter, I., Richter, C., Wunsch, M., and Marie, D. (2002) In situ uptake of ultraplankton by Red Sea cavity-dwelling and epi-reefal sponges. In *Feeding ecology of coral reef sponges*, 46–63.
- Laffy, P.W., Wood-Charlson, E.M., Turaev, D., Jutz, S., Pascelli, C., Botté, E.S., et al. (2018) Reef invertebrate viromics: diversity, host specificity and functional capacity. *Environ Microbiol* **20**: 2125–2141.
- Laffy, P.W., Botté, E.S., Wood-Charlson, E.M., Weynberg, K.D., Rattei, T., and Webster, N.S. (2019) Thermal stress modifies the marine sponge virome. *Environ Microbiol Rep* **11**: 690–698.
- Lafi, F.F., Fuerst, J.A., Fieseler, L., Engels, C., Goh, W.W.L., and Hentschel, U. (2009) Widespread distribution of poribacteria in demospongiae. *Appl Environ Microbiol* **75**: 5695–5699.
- Landry, Z., Swa, B.K., Herndl, G.J., Stepanauskas, R., and Giovannoni, S.J. (2017) SAR202 genomes from the dark ocean predict pathways for the oxidation of recalcitrant dissolved organic matter. *mBio* **8**: 1–19.
- Lavy, A., Keren, R., Haber, M., Schwartz, I., and Ilan, M. (2014) Implementing sponge physiological and genomic information to enhance the diversity of its culturable associated bacteria. *FEMS Microbiol Ecol* **87**: 486–502.
- Lee, O.O., Wong, Y.H., and Qian, P.Y. (2009) Inter- and intraspecific variations of bacterial communities associated with marine sponges from San Juan Island, Washington. *Appl Environ Microbiol* **75**: 3513–3521.
- Lee, O.O., Wang, Y., Yang, J., Lafi, F.F., Al-Suwailem, A., and Qian, P.Y. (2011) Pyrosequencing reveals highly diverse and species-specific microbial communities in sponges from the Red Sea. *ISME J* **5**: 650–664.
- Lejon, D.P.H., Kennedy, J., and Dobson, A.D.W. (2011) Identification of novel bioactive compounds from the metagenome of the marine sponge *Haliclona simulans*. In de Bruijn, F.J. (ed), *Handbook of Molecular Microbial Ecology, Volume II: Metagenomics in Different Habitats*, John Wiley & Sons, 553–562.
- Lesser, M.P. (2006) Benthic-pelagic coupling on coral reefs: Feeding and growth of Caribbean sponges. *J Exp Mar Bio Ecol* **328**: 277–288.
- Levin, L.A. and Bris, N.L. (2015) The deep ocean under climate change. *Science* **350**: 766–768.
- Leys, S.P., Mackie, G.O., and Reischwig, H.M. (2007) The biology of glass sponges. *Adv Mar Biol* **52**: 1–145.
- Leys, S.P., Yahel, G., Reidenbach, M.A., Tunnicliffe, V., Shavit, U., and Reischwig, H.M. (2011) The sponge pump: The role of current induced flow in the design of the sponge body plan. *PLoS One* **6**: e27787.
- Leys, S.P., Kahn, A.S., Fang, J.K.H., Kutti, T., and Bannister, R.J. (2018) Phagocytosis of microbial symbionts balances the carbon and nitrogen budget for the deep-water boreal sponge *Geodia barretti*. *Limnol Oceanogr* **63**: 187–202.
- Li, Z.Y., Wang, Y.Z., He, L.M., and Zheng, H.J. (2014) Metabolic profiles of prokaryotic and eukaryotic communities in deep-sea sponge *Lamellomorpha* sp. indicated by metagenomics. *Sci Rep* **4**: 1–12.
- Liu, M., Fan, L., Zhong, L., Kjelleberg, S., and Thomas, T. (2012) Metaproteomic analysis of a community of sponge symbionts. *ISME J* **6**: 1515–1525.
- López-Acosta, M., Leynaert, A., Chavaud, L., Amice, E., Bihannic, I., Le Bec, T., and Maldonado, M. (2019) In situ determination of Si, N, and P utilization by the demosponge *Tethya citrina*: A benthic-chamber approach. *PLoS One* **14**: 1–15.
- Lurgi, M., Thomas, T., Wemheuer, B., Webster, N.S., and Montoya, J.M. (2019) Modularity and predicted functions of the global sponge-microbiome network. *Nat Commun* **10**: 1–12.
- Luter, H.M., Whalan, S., and Webster, N.S. (2010a) Exploring the role of microorganisms in the disease-like syndrome affecting the sponge *Ianthella basta*. *Appl Environ Microbiol* **76**: 5736–5744.
- Luter, H.M., Whalan, S., and Webster, N.S. (2010b) Prevalence of tissue necrosis and brown spot lesions in a common marine sponge. *Mar Freshw Res* **61**: 484–489.
- Luter, H.M., Whalan, S., and Webster, N.S. (2012) Thermal and sedimentation stress are unlikely causes of brown spot syndrome in the Coral Reef sponge, *Ianthella basta*. *PLoS One* **7**: 1–9.
- Luter, H.M., Gibb, K., and Webster, N.S. (2014) Eutrophication has no short-term effect on the *Cymbastela stiptata* holobiont. *Front Microbiol* **5**: 1–11.
- Luter, H.M., Widder, S., Botté, E.S., Abdul Wahab, M., Whalan, S., Moitinho-Silva, L., et al. (2015) Biogeographic variation in the microbiome of the ecologically important sponge, *Carteriospongia foliascens*. *PeerJ* **3**: e1435.
- Luter, H.M., Bannister, R.J., Whalan, S., Kutti, T., Pineda, M.-C., and Webster, N.S. (2017) Microbiome analysis of a disease affecting the deep-sea sponge *Geodia barretti*. *FEMS Microbiol Ecol* **93**: 1–6.
- Luter, H.M. and Webster, N.S. (2017) Sponge disease and climate change. In Carballo, J.L. and Bell, J.J. (eds), *Climate Change, Ocean Acidification and Sponges*. Springer International Publishing, 411–428.
- Luter, H.M., Whalan, S., Andreakis, N., Abdul Wahab, M., Botté, E.S., Negri, A.P., and Webster, N.S. (2019) The effects of crude oil and dispersant on the larval sponge holobiont. *mSystems* **4**: 1–17.

- Maldonado, M., Carmona, M.C., Velásquez, Z., Puig, A., Cruzado, A., López, A., and Young, C.M. (2005) Siliceous sponges as a silicon sink: An overlooked aspect of benthopelagic coupling in the marine silicon cycle. *Limnol Oceanogr* **50**: 799–809.
- Maldonado, M., Zhang, X., Cao, X., Xue, L., Cao, H., and Zhang, W. (2010a) Selective feeding by sponges on pathogenic microbes: A reassessment of potential for abatement of microbial pollution. *Mar Ecol Prog Ser* **403**: 75–89.
- Maldonado, M., Sánchez-Tocino, L., and Navarro, C. (2010b) Recurrent disease outbreaks in corneous demosponges of the genus *Ircinia*: Epidemic incidence and defense mechanisms. *Mar Biol* **157**: 1577–1590.
- Maldonado, M., Ribes, M., and van Duyl, F.C. (2012) Nutrient fluxes through sponges. Biology, budgets, and ecological implications. *Adv Mar Biol* **62**: 113–182.
- Maldonado, M. (2016) Sponge waste that fuels marine oligotrophic food webs: a re-assessment of its origin and nature. *Mar Ecol* **37**: 477–491.
- Maldonado, M., Aguilar, R., Bannister, R.J., James, J., Conway, K.W., Dayton, P.K., et al. (2017) Sponge grounds as key marine habitats: A synthetic review of types, structure, functional roles, and conservation concerns. In Rossi, S., Bramanti, L., Gori, A., and Orejas, C. (eds), *Marine Animal Forests*. Springer, Cham.
- Maldonado, M., López-Acosta, M., Sitjà, C., García-Puig, M., Galobart, C., Ercilla, G., and Leynaert, A. (2019) Sponge skeletons as an important sink of silicon in the global oceans. *Nat Geosci* **12**: 815–822.
- De Mares, M.C., Sipkema, D., Huang, S., Bunk, B., Overmann, J., and van Elsas, J.D. (2017) Host specificity for bacterial, archaeal and fungal communities determined for high- and low-microbial abundance sponge species in two genera. *Front Microbiol* **8**: 1–13.
- Margulis, L. (1998) *Symbiotic planet: A new look at evolution*, Basic Books, New York.
- Mariani, S., Baillie, C., Colosimo, G., and Riesgo, A. (2019) Sponges as natural environmental DNA samplers. *Curr Biol* **29**: R401–R402.
- Márquez, L.M., Redman, R.S., Rodríguez, R.J., and Roossinck, M.J. (2007) A virus in a fungus in a plant: Three-way symbiosis required for thermal tolerance. *Science* **315**: 513–515.
- Marshall, N.B. (1971) *Explorations in the life of fishes*, Harvard University Press, Cambridge, MA.
- Martiny, J.B.H., Bohannon, B.J.M., Brown, J.H., Colwell, R.K., Fuhrman, J.A., Green, J.L., et al. (2006) Microbial biogeography: Putting microorganisms on the map. *Nat Rev Microbiol* **4**: 102–112.
- Martiny, J.B.H., Jones, S.E., Lennon, J.T., and Martiny, A.C. (2015) Microbiomes in light of traits: A phylogenetic perspective. *Science* **350**: aac9323.
- Massana, R., Delong, E.F., and Pedrós-Alió, C. (2000) A few cosmopolitan phylotypes dominate planktonic archaeal assemblages in widely different oceanic provinces. *Appl Environ Microbiol* **66**: 1777–1787.
- Massaro, A.J., Weisz, J.B., Hill, M.S., and Webster, N.S. (2012) Behavioral and morphological changes caused by thermal stress in the Great Barrier Reef sponge *Rhopaloeides odorabile*. *J Exp Mar Bio Ecol* **416–417**: 55–60.
- Maxwell, S.M., Gjerde, K.M., Connors, M.G., and Crowder, L.B. (2020) Mobile protected areas for biodiversity on the high seas. *Science* **367**: 252–254.
- McFall-Ngai, M., Hadfield, M.G., Bosch, T.C.G., Carey, H. V, Domazet-Lošo, T., Douglas, A.E., et al. (2013) Animals in a bacterial world, a new imperative for the life sciences. *Proc Natl Acad Sci* **110**: 3229–3236.
- Michin, E.A. (1900) The Porifera and Coelenterata. In *A treatise on zoology, part II*. Adam and Charles Black, London, UK, 1–78.
- Moitinho-Silva, L., Bayer, K., Cannistraci, C. V., Giles, E.C., Ryu, T., Seridi, L., et al. (2014) Specificity and transcriptional activity of microbiota associated with low and high microbial abundance sponges from the Red Sea. *Mol Ecol* **23**: 1348–1363.
- Moitinho-Silva, L., Díez-Vives, C., Batani, G., Esteves, A.I.S., Jahn, M.T., and Thomas, T. (2017a) Integrated metabolism in sponge-microbe symbiosis revealed by genome-centered metatranscriptomics. *ISME J* **11**: 1651–1666.
- Moitinho-Silva, L., Steinert, G., Nielsen, S., Hardoim, C.C.P., Wu, Y.C., McCormack, G.P., et al. (2017b) Predicting the HMA-LMA status in marine sponges by machine learning. *Front Microbiol* **8**: 1–14.
- Moitinho-Silva, L., Nielsen, S., Amir, A., Gonzalez, A., Ackermann, G.L., Cerrano, C., et al. (2017c) The sponge microbiome project. *Gigascience* **1**–13.
- Morganti, T., Coma, R., Yahel, G., and Ribes, M. (2017) Trophic niche separation that facilitates co-existence of high and low microbial abundance sponges is revealed by in situ study of carbon and nitrogen fluxes. *Limnol Oceanogr* **62**: 1963–1983.
- Morganti, T.M., Ribes, M., Yahel, G., and Coma, R. (2019) Size is the major determinant of pumping rates in marine sponges. *Front Physiol* **10**: 1474.
- Mori, T., Cahn, J.K.B., Wilson, M.C., Meoded, R.A., Wiebach, V., Martinez, A.F.C., et al. (2018) Single-bacterial genomics validates rich and varied specialized metabolism of uncultivated *Entotheonella* sponge symbionts. *Proc Natl Acad Sci* **115**: 1718–1723.
- Morrow, K.M., Bourne, D.G., Humphrey, C., Botté, E.S., Laffy, P., Zaneveld, J., et al. (2015) Natural volcanic CO₂ seeps reveal future trajectories for host-microbial associations in corals and sponges. *ISME J* **9**: 894–908.
- Mueller, B., De Goeij, J.M., Vermeij, M.J.A., Mulders, Y., Van Der Ent, E., Ribes, M., and Van Duyl, F.C. (2014a) Natural diet of coral-excavating sponges consists mainly of dissolved organic carbon (DOC). *PLoS One* **9**: 1–7.
- Mueller, B., Van Der Zande, R.M., Van Leent, P.J.M., Meesters, E.H., Vermeij, M.J.A., and Van Duyl, F.C. (2014b) Effect of light availability on dissolved organic carbon release by Caribbean reef algae and corals. *Bull Mar Sci* **90**: 875–893.
- Naim, M.A., Morillo, J.A., Sørensen, S.J., Waleed, A.A.S., Smidt, H., and Sipkema, D. (2014) Host-specific microbial communities in three sympatric North Sea sponges. *FEMS Microbiol Ecol* **90**: 390–403.
- Naim, M.A., Smidt, H., and Sipkema, D. (2017) Fungi found in Mediterranean and North Sea sponges: How specific are they? *PeerJ* **2017**: 1–19.
- Nyholm, S. V. and McFall-Ngai, M.J. (2004) The winnowing: Establishing the squid - *Vibrios* symbiosis. *Nat Rev Microbiol* **2**: 632–642.
- Van Oevelen, D., Duineveld, G., Lavaleye, M., Mienis, F., Soetaert, K., and Heip, C.H.R. (2009) The cold-water coral community as a hot spot for carbon cycling on continental margins: A food-web analysis from rockall bank (northeast atlantic). *Limnol Oceanogr* **54**: 1829–1844.
- Off, S., Alawi, M., and Spieck, E. (2010) Enrichment and physiological characterization of a novel nitrospira-like bacterium obtained from a marine sponge. *Appl Environ Microbiol* **76**: 4640–4646.
- Osinga, R., Tramper, J., and Wijffels, R.H. (1999) Cultivation of marine sponges. *Mar Biotechnol* **1**: 509–532.
- Pande, S. and Kost, C. (2017) Bacterial unculturability and the formation of intercellular metabolic networks. *Trends Microbiol* **25**: 349–361.
- Pantile, R. and Webster, N. (2011) Strict thermal threshold identified by quantitative PCR in the sponge *Rhopaloeides odorabile*. *Mar Ecol Prog Ser* **431**: 97–105.
- Pape, T., Hoffmann, F., Quéric, N.V., Von Juterzenka, K., Reitner, J., and Michaelis, W. (2006) Dense populations of Archaea associated with the demosponge *Tentorium semisuberites* Schmidt, 1870 from Arctic deep-waters. *Polar Biol* **29**: 662–667.
- Parks, D.H., Chuvochina, M., Waite, D.W., Rinke, C., Skarshewski, A., Chaumeil, P.-A., and Hugenholtz, P. (2018) A standardized bacterial taxonomy based on genome phylogeny substantially revises the tree of life. *Nat Biotechnol*. **36**: 996–1004.
- Partensky, F., Hess, W.R., and Vaulot, D. (1999) *Prochlorococcus*, a marine photosynthetic prokaryote of global significance. *Microbiol Mol Biol Rev* **63**: 106–127.
- Pedrós-Alió, C. (2012) The rare bacterial biosphere. *Ann Rev Mar Sci* **4**: 449–466.
- Perea-Blázquez, A., Davy, S.K., and Bell, J.J. (2012) Estimates of particulate organic carbon flowing from the pelagic environment to the benthos through sponge assemblages. *PLoS One* **7**: e29569.

- Pham, C.K., Murillo, F.J., Lirette, C., Maldonado, M., Colaço, A., Ottaviani, D., and Kenchington, E. (2019) Removal of deep-sea sponges by bottom trawling in the Flemish Cap area: conservation, ecology and economic assessment. *Sci Rep* **9**: 15843.
- Piel, J., Hui, D., Wen, G., Butzke, D., Platzer, M., Fusetani, N., and Matsunaga, S. (2004) Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge *Theonella swinhoei*. *Proc Natl Acad Sci* **101**: 16222–16227.
- Pile, A.J., Patterson, M.R., Savarese, M., Chernykh, V.I., and Fialkov, V.A. (1997) Trophic effects of sponge feeding within Lake Baikal's littoral zone. Sponge abundance, diet, feeding efficiency, and carbon flux. *Limnol Oceanogr* **42**: 178–184.
- Pile, A.J. and Young, C.M. (2006) The natural diet of a hexactinellid sponge: Benthic-pelagic coupling in a deep-sea microbial food web. *Deep Sea Res. Part I Oceanogr.* **53**: 1148–1156.
- Pineda, M.-C., Strehlow, B., Duckworth, A., Doyle, J., Jones, R., and Webster, N.S. (2016) Effects of light attenuation on the sponge holobiont- implications for dredging management. *Sci Rep* **6**: 39038.
- Pineda, M.-C., Strehlow, B., Kamp, J., Duckworth, A., Jones, R., and Webster, N.S. (2017a) Effects of combined dredging-related stressors on sponges: a laboratory approach using realistic scenarios. *Sci Rep* **7**: 5155.
- Pineda, M.C., Strehlow, B., Sternel, M., Duckworth, A., Haan, J. Den, Jones, R., and Webster, N.S. (2017b) Effects of sediment smothering on the sponge holobiont with implications for dredging management. *Sci Rep* **7**: 1–15.
- Pita, L., Turon, X., López-Legentil, S., and Erwin, P.M. (2013a) Host rules: Spatial stability of bacterial communities associated with marine sponges (*Ircinia* spp.) in the western Mediterranean sea. *FEMS Microbiol Ecol* **86**: 268–276.
- Pita, L., López-Legentil, S., and Erwin, P.M. (2013b) Biogeography and host fidelity of bacterial communities in *Ircinia* spp. from the Bahamas. *Microb Ecol* **66**: 437–447.
- Pita, L., Hoepfner, M.P., Ribes, M., and Hentschel, U. (2018a) Differential expression of immune receptors in two marine sponges upon exposure to microbial-associated molecular patterns. *Sci Rep* **8**: 1–15.
- Pita, L., Rix, L., Slaby, B.M., Franke, A., and Hentschel, U. (2018b) The sponge holobiont in a changing ocean: from microbes to ecosystems. *Microbiome* **6**: 46.
- Polónia, A.R.M., Cleary, D.F.R., Duarte, L.N., de Voogd, N.J., and Gomes, N.C.M. (2014) Composition of Archaea in seawater, sediment, and sponges in the Kepulauan Seribu reef system, Indonesia. *Microb Ecol* **67**: 553–567.
- Polónia, A.R.M., Cleary, D.F.R., Freitas, R., De Voogd, N.J., and Gomes, N.C.M. (2015) The putative functional ecology and distribution of archaeal communities in sponges, sediment and seawater in a coral reef environment. *Mol Ecol* **24**: 409–423.
- Polónia, A.R.M., Cleary, D.F.R., Coelho, F.J.R. da C., Becking, L.E., de Voogd, N.J., Toha, A.H.A., and Gomes, N.C.M. (2018) Compositional analysis of archaeal communities in high and low microbial abundance sponges in the Misool coral reef system, Indonesia. *Mar Biol Res* **14**: 537–550.
- Poppell, E., Weisz, J., Spicer, L., Massaro, A., Hill, A., and Hill, M. (2013) Sponge heterotrophic capacity and bacterial community structure in high- and low-microbial abundance sponges. *Mar Ecol* **35**: 414–424.
- Pourbaix, N. (1933) Mecanisme de la nutrition chez les Spongillidae. *Ann la Société R Zool Belgique* **64**: 11–20.
- Preston, C.M., Wu, K.Y., Molinski, T.F., and DeLong, E.F. (1996) A psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp. nov. *Proc Natl Acad Sci* **93**: 6241–6246.
- Püttner, A.E. (1909) Die Ernährung der Wassertiere und der Stoffhaushalt der Gewässer. *Fisher*, Jena.
- Püttner, A.E. (1914) Der Stoffwechsel der Kiesel Schwämme. *Zeit Allg Physiol* **16**: 65–114.
- Radax, R., Rattei, T., Lanzen, A., Bayer, C., Rapp, H.T., Urlich, T., and Schleper, C. (2012a) Metatranscriptomics of the marine sponge *Geodia barretti*: Tackling phylogeny and function of its microbial community. *Environ Microbiol* **14**: 1308–1324.
- Radax, R., Hoffmann, F., Rapp, H.T., Leininger, S., and Schleper, C. (2012b) Ammonia-oxidizing archaea as main drivers of nitrification in cold-water sponges. *Environ Microbiol* **14**: 909–923.
- Ramirez-Llodra, E., Brandt, A., Danovaro, R., De Mol, B., Escobar, E., German, C.R., et al. (2010) Deep, diverse and definitely different: Unique attributes of the world's largest ecosystem. *Biogeosciences* **7**: 2851–2899.
- Ramirez-Llodra, E., Clark, M.R., Smith, C.R., Tyler, P.A., Rowden, A.A., Bergstad, O.A., et al. (2011) Man and the last great wilderness: Human impact on the deep sea. *PLoS One* **6**: e22588.
- Ramsby, B.D., Hoogenboom, M.O., Whalan, S., and Webster, N.S. (2018) Elevated seawater temperature disrupts the microbiome of an ecologically important bioeroding sponge. *Mol Ecol* **27**: 2124–2137.
- Reiswig, H.M. (1971a) In situ pumping activities of tropical Demospongiae. *Mar Biol* **9**: 38–50.
- Reiswig, H.M. (1971b) Particle feeding in natural populations of three marine demosponges. *Biol Bull* **141**: 568–591.
- Reiswig, H.M. (1974) Water transport, respiration and energetics of three tropical marine sponges. *J Exp Mar Bio Ecol* **14**: 231–249.
- Reiswig, H. (1981) Partial carbon and energy budgets of the bacteriosponge *Verongia* in Barbados. *PSZNI Mar Ecol* **2**: 273–293.
- Reveillaud, J., Maignien, L., Eren, M.A., Huber, J.A., Apprill, A., Sogin, M.L., and Vanreusel, A. (2014) Host-specificity among abundant and rare taxa in the sponge microbiome. *ISME J* **8**: 1198–1209.
- Ribes, M., Coma, R., and Gili, J.M. (1999) Natural diet and grazing rate of the temperate sponge *Dysidea avara* (Demospongiae, Dendroceratida) throughout an annual cycle. *Mar Ecol Prog Ser* **176**: 179–190.
- Ribes, M., Coma, R., Atkinson, M.J., and Kinzie, R.A. (2003) Particle removal by coral reef communities: Picoplankton is a major source of nitrogen. *Mar Ecol Prog Ser* **257**: 13–23.
- Ribes, M., Coma, R., Atkinson, M.J., and Kinzie, R.A. (2005) Sponges and ascidians control removal of particulate organic nitrogen from coral reef water. *Limnol Oceanogr* **50**: 1480–1489.
- Ribes, M., Jiménez, E., Yahel, G., López-Sendino, P., Diez, B., Massana, R., et al. (2012) Functional convergence of microbes associated with temperate marine sponges. *Environ Microbiol* **14**: 1224–1239.
- Ribes, M., Dziallas, C., Coma, R., and Riemann, L. (2015) Microbial diversity and putative diazotrophy in high- and low-microbial-abundance mediterranean sponges. *Appl Environ Microbiol* **81**: 5683–5693.
- Richter, C., Wunsch, M., Rasheed, M., Kötter, I., and Badran, M.I. (2001) Endoscopic exploration of Red Sea coral reefs reveals dense populations of cavity-dwelling sponges. *Nature* **413**: 726–730.
- Riesgo, A., Peterson, K., Richardson, C., Heist, T., Strehlow, B., McCauley, M., et al. (2014) Transcriptomic analysis of differential host gene expression upon uptake of symbionts: A case study with *Symbiodinium* and the major bioeroding sponge *Cliona varians*. *BMC Genomics* **15**: 376.
- Rinke, C., Schwientek, P., Sczyrba, A., Ivanova, N.N., Anderson, I.J., Cheng, J.-F., et al. (2013) Insights into the phylogeny and coding potential of microbial dark matter. *Nature* **499**: 431–437.
- Rix, L., Goeij, J.M. De, Oevelen, D. Van, Struck, U., Al-horani, F.A., Wild, C., and Naumann, M.S. (2016) Differential recycling of coral and algal dissolved organic matter via the sponge loop. *Funct Ecol* **31**: 1–12.
- Rix, L., Ribes, M., Coma, R., Jahn, M.T., de Goeij, J.M., van Oevelen, D., et al. (2020) Heterotrophy in the earliest gut: a single-cell view of heterotrophic carbon and nitrogen assimilation in sponge-microbe symbioses. *ISME J* **14**: 2554–2567.
- Roberts, J.M., Davies, A.J., Henry, L.A., Dodds, L.A., Duineveld, G.C.A., Lavaleye, M.S.S., et al. (2009) Mingulay reef complex: An interdisciplinary study of cold-water coral habitat, hydrography and biodiversity. *Mar Ecol Prog Ser* **397**: 139–151.

- Rodrigues de Oliveira, B.F., Freitas-Silva, J., Sanchez-Robinet, C., and Silva Laport, M. (2020) Transmission of the sponge microbiome: moving towards a unified model. *Environ Microbiol Rep*. doi: 10.1111/1758-2229.12896.
- Rodríguez-Marconi, S., De La Iglesia, R., Díez, B., Fonseca, C.A., Hajdu, E., Trefault, N., and Webster, N. (2015) Characterization of bacterial, archaeal and eukaryote symbionts from antarctic sponges reveals a high diversity at a three-domain level and a particular signature for this ecosystem. *PLoS One* **10**: 1–19.
- Rooks, C., Kar-Hei Fang, J., Mørkved, P.T., Zhao, R., Tore Rapp, H., Xavier, J.J., and Hoffmann, F. (2020) Deep-sea sponge grounds as nutrient sinks: Denitrification is common in boreo-Arctic sponges. *Biogeosciences* **17**: 1231–1245.
- Ryu, T., Seridi, L., Moitinho-Silva, L., Oates, M., Liew, Y.J., Mavromatis, C., et al. (2016) Hologenome analysis of two marine sponges with different microbiomes. *BMC Genomics* **17**: 158.
- Sachs, J.L. (2004) The evolution of cooperation. *Q Rev Biol* **79**: 135–160.
- Sachs, J.L. and Simms, E.L. (2006) Pathways to mutualism breakdown. *Trends Ecol Evol* **21**: 585–592.
- Salem, H., Bauer, E., Kirsch, R., Berasategui, A., Cripps, M., Weiss, B., et al. (2017) Drastic genome reduction in an herbivore's pectinolytic symbiont. *Cell* **171**: 1520–1531.
- Santos Ferreira, M.R., Cleary, D.F.R., Coelho, F.J.R.C., Gomes, N.C.M., Huang, Y.M., Polónia, A.R.M., and De Voogd, N.J. (2020) Geographical location and habitat predict variation in prokaryotic community composition of *Suberites diversicolor*. *Ann Microbiol* **70**: 14.
- Schläppy, M.L., Schöttner, S.I., Lavik, G., Kuypers, M.M.M., de Beer, D., and Hoffmann, F. (2010) Evidence of nitrification and denitrification in high and low microbial abundance sponges. *Mar Biol* **157**: 593–602.
- Schmitt, S., Weisz, J.B., Lindquist, N., and Hentschel, U. (2007) Vertical transmission of a phylogenetically complex microbial consortium in the viviparous sponge *Ircinia felix*. *Appl Environ Microbiol* **73**: 2067–2078.
- Schmitt, S., Angermeier, H., Schiller, R., Lindquist, N., and Hentschel, U. (2008) Molecular microbial diversity survey of sponge reproductive stages and mechanistic insights into vertical transmission of microbial symbionts. *Appl Environ Microbiol* **74**: 7694–7708.
- Schmitt, S., Tsai, P., Bell, J., Fromont, J., Ilan, M., Lindquist, N., et al. (2012) Assessing the complex sponge microbiota: core, variable and species-specific bacterial communities in marine sponges. *ISME J* **6**: 564–576.
- Schmittmann, L., Jahn, M.T., Pita, L., and Hentschel, U. (2020) Decoding cellular dialogues between sponges, bacteria, and phages. In Bosch, T.C.G. and Hadfield, M.G. (eds) *Cellular dialogues in the holobiont*, CRC Press, Kiel, 49–63.
- Schuster, A., Strehlow, B.W., Eckford-Soper, L., McAllen, R., and Canfield, D.E. (2020) The effects of seasonal anoxia on the microbial community structure in demosponges in a marine lake (Lough Hyne, Ireland). *bioRxiv* 2020.09.09.290791.
- Shade, A. and Handelsman, J. (2012) Beyond the Venn diagram: The hunt for a core microbiome. *Environ Microbiol* **14**: 4–12.
- Siegl, A., Kamke, J., Hochmuth, T., Piel, J., Richter, M., Liang, C., et al. (2011) Single-cell genomics reveals the lifestyle of *Poribacteria*, a candidate phylum symbiotically associated with marine sponges. *ISME J* **5**: 61–70.
- Simister, R.L., Deines, P., Botté, E.S., Webster, N.S., and Taylor, M.W. (2012a) Sponge-specific clusters revisited: A comprehensive phylogeny of sponge-associated microorganisms. *Environ Microbiol* **14**: 517–524.
- Simister, R., Taylor, M.W., Tsai, P., Fan, L., Bruxner, T.J., Crowe, M.L., and Webster, N. (2012b) Thermal stress responses in the bacterial biosphere of the great barrier reef sponge, *Rhopaloeides odorabile*. *Environ Microbiol* **14**: 3232–3246.
- Simister, R., Taylor, M.W., Tsai, P., and Webster, N. (2012c) Sponge-microbe associations survive high nutrients and temperatures. *PLoS One* **7**: 21–23.
- Simister, R., Taylor, M.W., Rogers, K.M., Schupp, P.J., and Deines, P. (2013) Temporal molecular and isotopic analysis of active bacterial communities in two New Zealand sponges. *FEMS Microbiol Ecol* **85**: 195–205.
- Simonsen, A.K., Dinnage, R., Barrett, L.G., Prober, S.M., and Thrall, P.H. (2017) Symbiosis limits establishment of legumes outside their native range at a global scale. *Nat Commun* **8**: 1–9.
- Simpson, T.L. (1984) *The cell biology of sponges*, Springer, New York.
- Sipkema, D. (2004) *Cultivation of marine sponges: from sea to cell*. Dissertation, Wageningen University.
- Sipkema, D., Schippers, K., Maalcke, W.J., Yang, Y., Salim, S., and Blanch, H.W. (2011) Multiple approaches to enhance the cultivability of bacteria associated with the marine sponge *Haliclona (gellius) sp.* *Appl Environ Microbiol* **77**: 2130–2140.
- Slaby, B.M., Hackl, T., Horn, H., Bayer, K., and Hentschel, U. (2017) Metagenomic binning of a marine sponge microbiome reveals unity in defense but metabolic specialization. *ISME J* **11**: 1–14.
- Slaby, B., Franke, A., Rix, L., Pita, L., Bayer, K., Jahn, M., and Hentschel, U. (2019) Marine sponge holobionts in health and disease. In Li, Z. (ed), *Symbiotic microbiomes of coral reefs sponges and corals*. Springer, Dordrecht, Netherlands, 81–104.
- Snelgrove, P.V.R. (2016) An ocean of discovery: Biodiversity beyond the Census of Marine Life. *Planta Med* **82**: 790–799.
- Snelgrove, P.V.R. (1999) Getting to the bottom of marine biodiversity: Sedimentary habitats: Ocean bottoms are the most widespread habitat on earth and support high biodiversity and key ecosystem services. *Bioscience* **49**: 129–130.
- Van Soest, R.W.M., Cleary, D.F.R., De Kluijver, M.J., Lavaleye, M.S.S., Maier, C., and Van Duyl, F.C. (2007) Sponge diversity and community composition in Irish bathyal coral reefs. *Contrib to Zool* **76**: 121–142.
- van Soest, R.W.M., Boury-Esnault, N., Vacelet, J., Dohrmann, M., Erpenbeck, D., de Voogd, N.J., et al. (2012) Global diversity of sponges (Porifera). *PLoS One* **7**: e35105.
- Sogin, E.M., Leisch, N., and Dubilier, N. (2020) Chemosynthetic symbioses. *Curr Biol* **30**: R1137–R1142.
- Southwell, M., Weisz, J., Martens, C., and Lindquist, N. (2008) In situ fluxes of dissolved inorganic nitrogen from the sponge community on Conch Reef, Key Largo, Florida. *Limnol Oceanogr* **53**: 986–996.
- Southwell, M.W. (2007) *Sponge impacts on coral reef nitrogen cycling*. Dissertation, University of North Carolina at Chapel Hill.
- Spang, A., Saw, J.H., Jørgensen, S.L., Zaremba-Niedzwiedzka, K., Martijn, J., Lind, A.E., et al. (2015) Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* **521**: 173–179.
- Spang, A., Caceres, E.F., and Ettema, T.J.G. (2017) Genomic exploration of the diversity, ecology, and evolution of the archaeal domain of life. *Science* **357**: 563.
- Steindler, L., Huchon, D., Avni, A., and Ilan, M. (2005) 16S rRNA phylogeny of sponge-associated *Cyanobacteria*. *Appl Environ Microbiol* **71**: 4127–4131.
- Steinert, G., Whitfield, S., Taylor, M.W., Thoms, C., and Schupp, P.J. (2014) Application of diffusion growth chambers for the cultivation of marine sponge-associated bacteria. *Mar Biotechnol* **16**: 594–603.
- Steinert, G. (2015) *Microbial diversity of temperate and tropical sponges*. Dissertation, University of Oldenburg.
- Steinert, G., Busch, K., Bayer, K., Khodami, S., Arbizu, P., Kelly, M., et al. (2020) Compositional and quantitative insights into bacterial and archaeal communities of South Pacific deep-sea sponges (Demospongiae and Hexactinellida). *Front Microbiol* **11**: 1–16.
- Strand, R., Whalan, S., Webster, N.S., Kutti, T., Fang, J.K.H., Luter, H.M., and Bannister, R.J. (2017) The response of a boreal deep-sea sponge holobiont to acute thermal stress. *Sci Rep* **7**: 1660.
- Swierts, T., Cleary, D.F.R., and de Voogd, N.J. (2018) Prokaryotic communities of Indo-Pacific giant barrel sponges are more strongly influenced by geography than host phylogeny. *FEMS Microbiol Ecol* **94**: 1–12.
- Tatzel, M., Von Blanckenburg, F., Oelze, M., Bouchez, J., and Hippler, D. (2017) Late Neoproterozoic seawater oxygenation by siliceous sponges. *Nat Commun* **8**: 1–9.

- Taylor, M.W., Schupp, P.J., De Nys, R., Kjelleberg, S., and Steinberg, P.D. (2005) Biogeography of bacteria associated with the marine sponge *Cymbastela concentrica*. *Environ Microbiol* **7**: 419–433.
- Taylor, M.W., Radax, R., Steger, D., and Wagner, M. (2007) Sponge-associated microorganisms: Evolution, ecology, and biotechnological potential. *Microbiol Mol Biol Rev* **71**: 295–347.
- Taylor, M.W., Tsai, P., Simister, R.L., Deines, P., Botte, E., Ericson, G., et al. (2013) Sponge-specific bacteria are widespread (but rare) in diverse marine environments. *ISME J* **7**: 438–443.
- Thacker, R.W. and Starnes, S. (2003) Host specificity of the symbiotic cyanobacterium *Oscillatoria spongelliae* in marine sponges, *Dysidea* spp. *Mar Biol* **142**: 643–648.
- Thistle, D. (2003) The deep-sea floor: An overview. In Tyler, P.A. (ed) *Ecosystems of the deep oceans*. Elsevier Science, 5–38.
- Thomas, T., Moitinho-Silva, L., Lurgi, M., Björk, J.R., Easson, C., Astudillo-García, C., et al. (2016) Diversity, structure and convergent evolution of the global sponge microbiome. *Nat Commun* **7**: 11870.
- Thompson, L.R., Sanders, J.G., McDonald, D., Amir, A., Ladau, J., Locey, K.J., et al. (2017) A communal catalogue reveals Earth's multiscale microbial diversity. *Nature* **551**: 457–463.
- Tian, R.M., Sun, J., Cai, L., Zhang, W.P., Zhou, G.W., Qiu, J.W., and Qian, P.Y. (2016) The deep-sea glass sponge *Lophophysema eversa* harbours potential symbionts responsible for the nutrient conversions of carbon, nitrogen and sulfur. *Environ Microbiol* **18**: 2481–2494.
- Turon, X., Galera, J., and Uriz, M.J. (1997) Clearance rates and aquiferous systems in two sponges with contrasting life-history strategies. *J Exp Zool* **278**: 22–36.
- Turon, M., Cáliz, J., Garate, L., Casamayor, E.O., and Uriz, M.J. (2018) Showcasing the role of seawater in bacteria recruitment and microbiome stability in sponges. *Sci Rep* **8**: 1–10.
- Turque, A.S., Batista, D., Silveira, C.B., Cardoso, A.M., Vieira, R.P., Moraes, F.C., et al. (2010) Environmental shaping of sponge associated archaeal communities. *PLoS One* **5**.
- Unson, M.D., Holland, N.D., and Faulkner, D.J. (1994) A brominated secondary metabolite synthesized by the cyanobacterial symbiont of a marine sponge and accumulation of the crystalline metabolite in the sponge tissue. *Mar Biol* **119**: 1–11.
- Uriz, M.J., Turon, X., Becerro, M.A., and Agell, G. (2003) Siliceous spicules and skeleton frameworks in sponges: Origin, diversity, ultrastructural patterns, and biological functions. *Microsc Res Tech* **62**: 279–299.
- Vacelet, J. and Donadey, C. (1977) Electron microscope study of the association between some sponges and bacteria. *J exp mar Biol Ecol* **30**: 301–314.
- Vacelet, J. and Boury-Esnault, N. (1995) Carnivorous sponges. *Nature* **374**: 170.
- Vacelet, J., Boury-Esnault, N., Fiala-Mendoní, A., and Fisher, C.R. (1995) A methanotrophic carnivorous sponge. *Nature* **377**: 296.
- Vacelet, J., Fiala-Médioni, A., Fisher, C.R., and Boury-Esnault, N. (1996) Symbiosis between methane-oxidizing bacteria and a deep-sea carnivorous cladorhizid sponge. *Mar Ecol Prog Ser* **145**: 77–85.
- Vacelet, J. and Boury-Esnault, N. (2002) A new species of carnivorous deep-sea sponge (Demospongiae: *Cladorhizidae*) associated with methanotrophic bacteria. *Cah Biol Mar* **43**: 141–148.
- Versluis, D., McPherson, K., van Passel, M.W.J., Smidt, H., and Sipkema, D. (2017) Recovery of previously uncultured bacterial genera from three Mediterranean sponges. *Mar Biotechnol* **19**: 454–468.
- Van de Vyver, G., Vray, B., Belaouane, S., and Toussaint, D. (1990) Efficiency and selectivity of microorganism retention by *Ephydatia fluviatilis*. In Rützler, K. (ed), *New Perspectives in Sponge Biology*. Smithsonian Institution Press, Washington D.C./London, 511–515.
- Wagner, D. and Kelley, C.D. (2017) The largest sponge in the world? *Mar Biodivers* **47**: 367–368.
- Watling, L., Guinotte, J., Clark, M.R., and Smith, C.R. (2013) A proposed biogeography of the deep ocean floor. *Prog Oceanogr* **111**: 91–112.
- Webster, N.S., Webb, R.I., Ridd, M.J., Hill, R.T., and Negri, A.P. (2001) The effects of copper on the microbial community of a coral reef sponge. *Environ Microbiol* **3**: 19–31.
- Webster, N.S., Negri, A.P., Webb, R.I., and Hill, R.T. (2002) A spongin-boring α -proteobacterium is the etiological agent of disease in the Great Barrier Reef sponge *Rhopaloeides odorabile*. *Mar Ecol Prog Ser* **232**: 305–309.
- Webster, N.S. (2007) Sponge disease: A global threat? *Environ Microbiol* **9**: 1363–1375.
- Webster, N.S., Cobb, R.E., and Negri, A.P. (2008a) Temperature thresholds for bacterial symbiosis with a sponge. *ISME J* **2**: 830–842.
- Webster, N.S., Xavier, J.R., Freckleton, M., Motti, C.A., and Cobb, R. (2008b) Shifts in microbial and chemical patterns within the marine sponge *Aplysina aerophoba* during a disease outbreak. *Environ Microbiol* **10**: 3366–3376.
- Webster, N.S., Bourne, D.G., and Blackall, L.L. (2009) Impact of global climate change on marine bacterial symbioses and disease. *Microbiol Aust* **30**: 78.
- Webster, N.S., Taylor, M.W., Behnam, F., Lücker, S., Rattei, T., Whalan, S., et al. (2010) Deep sequencing reveals exceptional diversity and modes of transmission for bacterial sponge symbionts. *Environ Microbiol* **12**: 2070–2082.
- Webster, N.S., Luter, H.M., Soo, R.M., Botté, E.S., Simister, R.L., Abdo, D., and Whalan, S. (2012) Same, same but different: Symbiotic bacterial associations in GBR sponges. *Front Microbiol* **3**: 1–11.
- Webster, N., Pantile, R., Botté, E., Abdo, D., Andreakis, N., and Whalan, S. (2013) A complex life cycle in a warming planet: Gene expression in thermally stressed sponges. *Mol Ecol* **22**: 1854–1868.
- Webster, N.S. and Hill, R. (2014) Vulnerability of marine microbes on the Great Barrier Reef to climate change. In *Climate Change and the Great Barrier Reef: A Vulnerability Assessment*. The Great Barrier Reef Marine Park Authority.
- Webster, N.S. and Reusch, T.B.H. (2017) Microbial contributions to the persistence of coral reefs. *ISME J* **11**: 2167–2174.
- Webster, N.S., Wagner, M., and Negri, A.P. (2018) Microbial conservation in the Anthropocene. *Environ Microbiol* **20**: 1925–1928.
- Weigel, B.L. and Erwin, P.M. (2016) Intraspecific variation in microbial symbiont communities of the sun sponge, *Hymeniacidon heliophila*, from intertidal and subtidal habitats. *Appl Environ Microbiol* **82**: 650–658.
- Weisz, J.B., Hentschel, U., Lindquist, N., and Martens, C.S. (2007) Linking abundance and diversity of sponge-associated microbial communities to metabolic differences in host sponges. *Mar Biol* **152**: 475–483.
- Weisz, J.B., Lindquist, N., and Martens, C.S. (2008) Do associated microbial abundances impact marine demosponge pumping rates and tissue densities? *Oecologia* **155**: 367–376.
- Wennekes, P.L., Rosindell, J., and Etienne, R.S. (2012) The neutral-niche debate: A philosophical perspective. *Acta Biotheor* **60**: 257–271.
- Whittaker, R.H. (1972) Evolution and measurement of species diversity. *Taxon* **21**: 213–251.
- Wilkinson, C.R. (1978) Microbial associations in sponges I-III. *Mar Biol* **49**: 161–185.
- Wilson, M.C., Mori, T., Rückert, C., Uria, A.R., Helf, M.J., Takada, K., et al. (2014) An environmental bacterial taxon with a large and distinct metabolic repertoire. *Nature* **506**: 58–62.
- Wolfrath, B. and Barthel, D. (1989) Production of faecal pellets by the marine sponge *Halichondria panicea* Pallas (1766). *J Exp Mar Biol Ecol* **129**: 81–94.
- Wootton, J.T. and Emmerson, M. (2005) Measurement of interaction strength in nature. *Annu Rev Ecol Evol Syst* **36**: 419–444.
- Yahel, G., Sharp, J.H., Marie, D., Häse, C., and Genin, A. (2003) In situ feeding and element removal in the symbiont-bearing sponge *Theonella swinhoei*: Bulk DOC is the major source for carbon. *Limnol Oceanogr* **48**: 141–149.
- Yahel, G., Eerkes-Medrano, D.I., and Leys, S.P. (2006) Size independent selective filtration of ultraplankton by hexactinellid glass sponges. *Aquat Microb Ecol* **45**: 181–194.

Yahel, G., Whitney, F., Reiswig, H.M., Eerkes-Medrano, D.I., and Leys, S.P. (2007) In situ feeding and metabolism of glass sponges (Hexactinellida, Porifera) studied in a deep temperate fjord with a remotely operated submersible. *Limnol Oceanogr* **52**: 428–440.

Yilmaz, P., Kottmann, R., Field, D., Knight, R., Cole, J.R., Amaral-Zettler, L., et al. (2011) Minimum information about a marker gene sequence (MIMARKS) and minimum information about any (x) sequence (MIXS) specifications. *Nat Biotechnol* **29**: 415–420.

Zaremba-Niedzwiedzka, K., Caceres, E.F., Saw, J.H., Bäckström, Di., Juzokaite, L., Vancaester, E., et al. (2017) *Asgard* archaea illuminate the origin of eukaryotic cellular complexity. *Nature* **541**: 353–358.

Source Code collection : Deep-sea sponge microbiome project

Kathrin Busch

2020-09-03

<https://kathrinbusch.github.io/16S-AmpliconCorePipeline/#>

Please cite the original article of Busch et al. 2020.

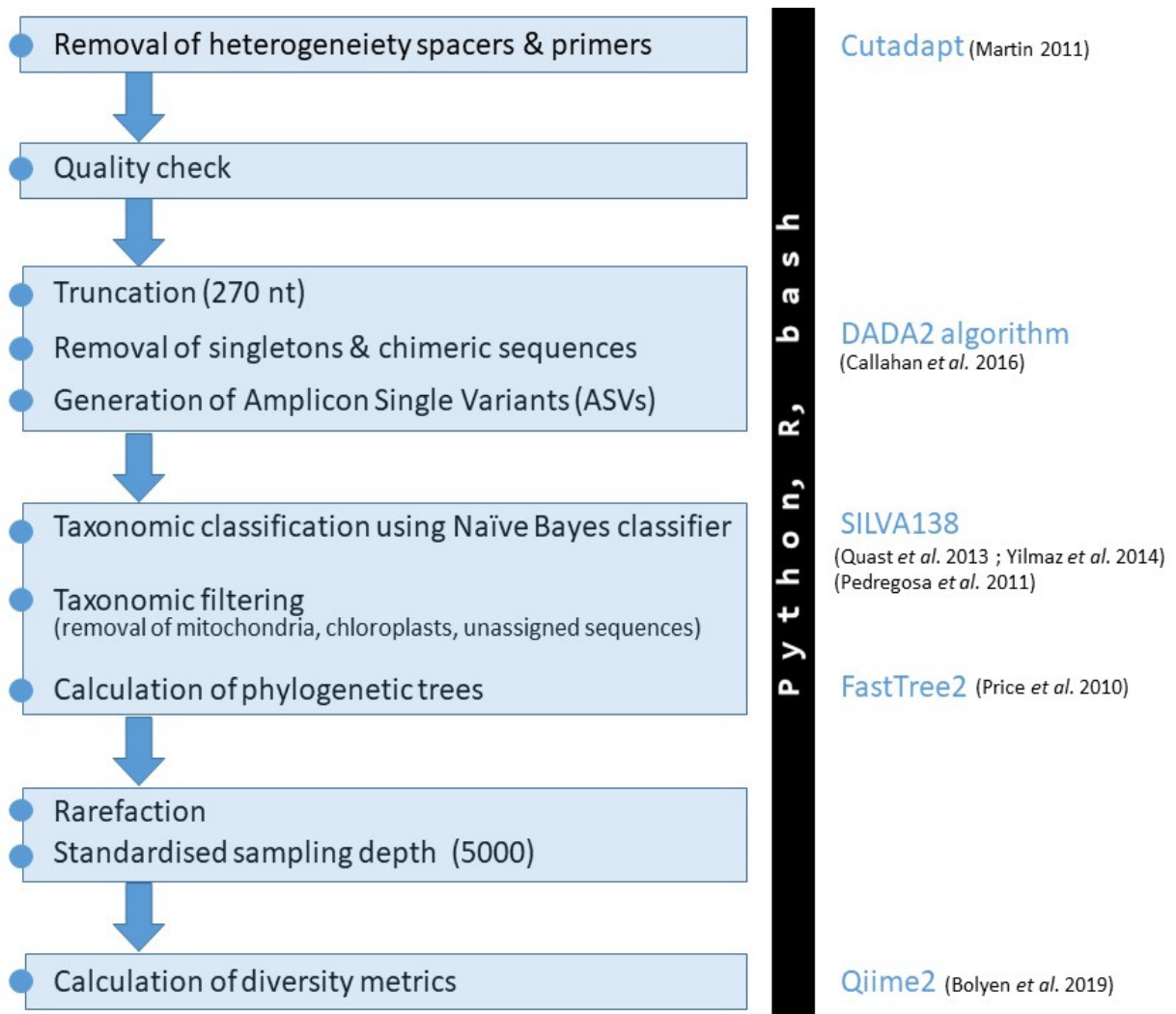


Illustrated by Kathrin Busch

16S Amplicon Core Analyses

0.1 Workflow diagram

This documentation includes the core bioinformatic SOP developed within the deep-sea sponge microbiome project. The very basic workflow looks like this:



0.2 Getting started

Operated from Ubuntu 14.04.6 LTS (GNU/Linux 4.4.0-148-generic x86_64) command line.

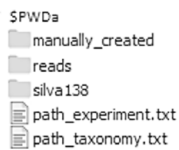
Program versions used:

- qiime2-2019.10 (via miniconda3 v4.4.10)
- R version 3.5.1 (unless other version is stated explicitly)
- SILVA_138

Input files and structure:

- Downloaded *SILVA_138_SSURef_NR99_tax_silva.fasta.gz* from https://www.arb-silva.de/no_cache/download/archive/release_138/Exports/ (accessed on the 19th of December 2019)
- Started the following script from the working directory (*path_experiment*) within a

screen session



```
$PWDa
├── manually_created
├── reads
├── silva138
└── path_experiment.txt

manually_created
├── path_experiment.txt
└── path_taxonomy.txt
```

[Download input files](#)

```
PWDa=`cat path_experiment.txt`
echo $PWDa
PWDb=`cat path_taxonomy.txt`
echo $PWDb
cd $PWDa/manually_created/
```

```
# manually filter dataset, remove outlier samples [outlier = (i) samples with unclear taxonomy (ii)
R
data<-read.table(file ="metadataV34_complete.txt", header = T, sep="\t")
dat<-read.table(file ="filter_manually.txt", header = T, sep="\t")
mydata<-merge(data, dat, by="SampleID", all=TRUE)
mydata$remove<-as.character(mydata$remove)
mydata$remove[is.na(mydata$remove)] <- "keep"
my<-mydata[!(mydata$remove=="delete"),]
my$remove<-NULL
names(my)[1] <- c('#SampleID')
write.table(my, file = "metadataV34.txt", sep="\t", col.names= T, row.names=F)
quit("no")
```

```
sed 's//g' metadataV34.txt > $PWDa/metadataV34.txt
cp $PWDa/manually_created/pre-33-manifestV34.csv $PWDa/pre-33-manifestV34.csv
rm -r -f -- - metadataV34.txt
source activate qiime2-2019.10
```

0.3 Remove heterogeneity spacers and primers

```
cd $PWDa/  
mkdir trimmed  
cd $PWDa/reads/  
  
for i in *_R1_001.fastq  
do  
cutadapt -g CCTACGGGAGGCAGCAG -o $PWDa/trimmed/$i $i  
done
```

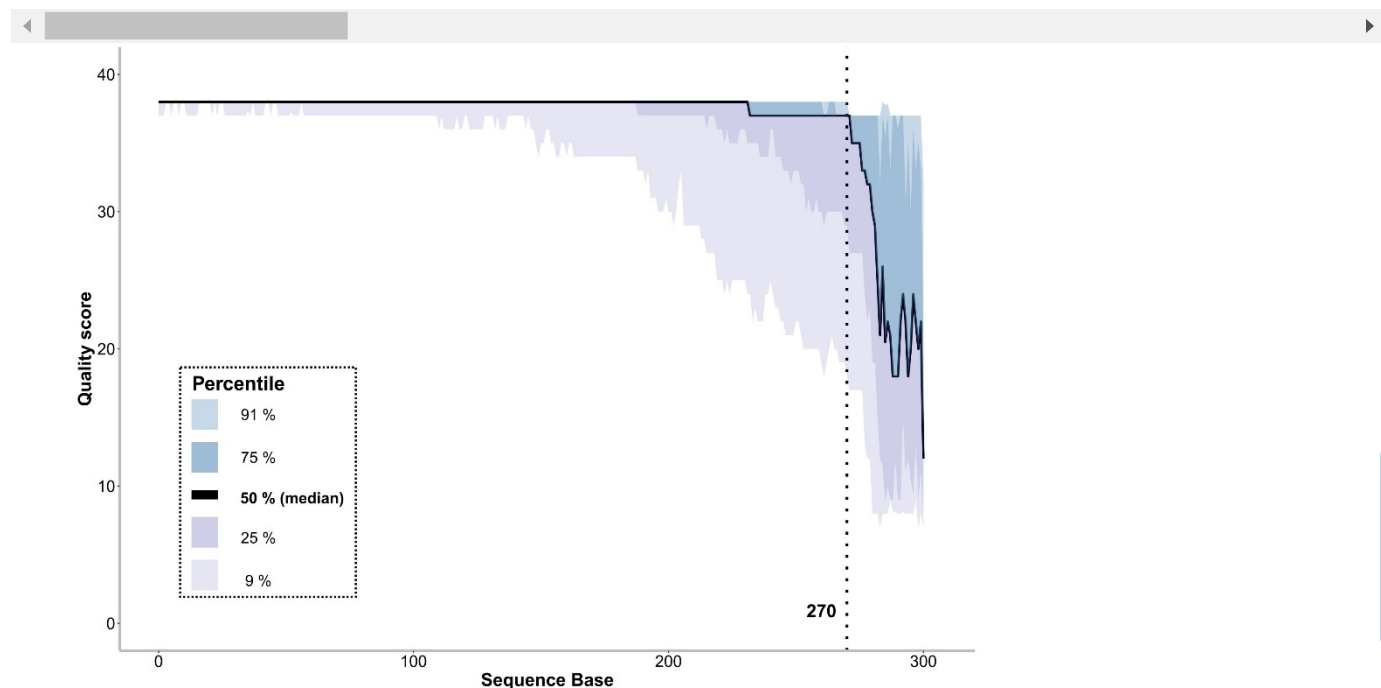
0.4 Visualise sequence quality and determine truncation length

```
cd $PWDa/  
mkdir dada2  
  
qiime tools import \  
--type 'SampleData[SequencesWithQuality]' \  
--input-path $PWDa/pre-33-manifestV34.csv \  
--output-path $PWDa/dada2/single-end-demux.qza \  
--input-format SingleEndFastqManifestPhred33  
  
qiime demux summarize \  
--i-data $PWDa/dada2/single-end-demux.qza \  
--o-visualization $PWDa/dada2/single-end-demux.qzv  
  
cd $PWDa/dada2  
unzip single-end-demux.qzv  
cd "$(ls -1dt ./ */ | head -n 1)/data"  
sed 's|,|\t|g' forward-seven-number-summaries.csv > forward-seven-number-summaries.txt  
sed -i '1s/^/sequence_base /' forward-seven-number-summaries.txt  
cp forward-seven-number-summaries.txt $PWDa/dada2/forward-seven-number-summaries.txt  
  
R  
list.of.packages <- c("tidyr", "plyr", "stringr", "ggplot2")  
new.packages <- list.of.packages[!(list.of.packages %in% installed.packages()[,"Package"])]  
if(length(new.packages)) install.packages(new.packages, repos='http://cran.us.r-project.org')  
library(tidyr)  
library(plyr)  
library(stringr)  
library(ggplot2)  
mydata = read.table(file = "forward-seven-number-summaries.txt", header = T)  
mydata=as.data.frame(mydata)  
a<-names(mydata)[1]  
b<-names(mydata)[ncol(mydata)]  
data<-gather(mydata,condition, measurement, names(mydata)[2:ncol(mydata)], factor_key=T)  
dat <- subset(data, condition != "sequence_base")  
names(dat)[names(dat) == "condition"] <- "base"  
names(dat)[names(dat) == "measurement"] <- "quality"  
dat<-rename(dat, c("sequence_base"="percentile"))  
dat<-dat[!(dat$percentile=="count"),]  
dat<-dat[!(dat$percentile=="2%"),]  
dat<-dat[!(dat$percentile=="98%"),]  
dat$percentile<-revalue(dat$percentile, c("9%"="09%"))  
data_wide <- spread(dat, percentile, quality)  
data_wide$base<-str_remove_all(data_wide$base, "[X]")  
data_wide$base<-as.numeric(data_wide$base)  
write.table(data_wide , file = "data_wide.txt", sep="\t", col.names= T, row.names=F)  
data_wide = read.table(file = "data_wide.txt", header = T)  
col <- data.frame(name = c("91%", "75%", "50% (median)", "25%", "09%"),  
start = c(11.5, 9.5, 8.25, 6,4),  
end = c(13, 11, 8.75, 7.5,5.5),
```

```

percentile = c("91%", "75%", "50% (median)", "25%", "09%"),
stringsAsFactors = FALSE)
pdf(file="sequence_quality.pdf", height= 10, width= 14)
p<-ggplot(data=data_wide, aes(x=base, y=X50.)) +
geom_line(linetype="solid", size=1.2) +
labs(x = "Sequence Base", y="Quality score") +
xlim(0, 305) +
ylim(0, 40) +
geom_vline(xintercept = 270, linetype="dotted", size=1.5)+
annotate("text", x = 260, y = 0, label = 'atop(bold("270"))', size=7, parse=TRUE) +
geom_ribbon(aes(ymin=X25., ymax=X50.), fill="darkblue",alpha=0.1) +
geom_ribbon(aes(ymin=X09., ymax=X50.), fill="darkblue",alpha=0.1) +
geom_ribbon(aes(ymin=X50., ymax=X75.), fill="#4682b4",alpha=0.3) +
geom_ribbon(aes(ymin=X50., ymax=X91.), fill="#4682b4",alpha=0.3) +
geom_rect(data=col, aes(NULL,NULL, ymin=start, ymax=end, fill=percentile), xmin=1, xmax=8,
scale_fill_manual(values=c("91%" = "#c7d9e8ff", "75%" = "#9fbdd7ff", "50% (median)" =
"#000000ff",
annotate("text", x = 13, y = 13, label = 'atop(bold("Percentile"))', size=5, parse=TRUE) +
annotate("text", x = 18, y = 12.25, label = "91 %", size=4) +
annotate("text", x = 18, y = 10.25, label = "75 %", size=4) +
annotate("text", x = 26.5, y = 7.75, label = 'atop(bold("50 % (median)"))', size=4,
parse=TRUE) +
annotate("text", x = 18, y = 6.5, label = "25 %", size=4) +
annotate("text", x = 18, y = 4.5, label = "9 %", size=4) +
theme(panel.grid.major = element_blank(),
panel.grid.minor = element_blank(),
axis.line = element_line(colour = "grey", size = 1.5, linetype = "solid"),
panel.background = element_rect(fill = "#ffffff",colour = NA),
plot.background = element_rect(fill = '#ffffff'),
legend.position = "none",
axis.title.x = element_text(color="black", size=20, face="bold"),
axis.text.x = element_text(face="plain", color="black",size=18, angle=0),
axis.title.y = element_text(color="black", size=20, face="bold"),
axis.text.y = element_text(face="plain", color="black",size=18, angle=0))
p
dev.off()
quit("no")

```



```

cp sequence_quality.pdf $PWDa/dada2/sequence_quality.pdf
cd $PWDa/dada2
rm -r -f -- - -d */
cd $PWDa/

```

0.5 Trimming, truncation, chimera removal and denoising (DADA2 algorithm)

```
qiime dada2 denoise-single \  
--i-demultiplexed-seqs $PWDa/dada2/single-end-demux.qza \  
--p-trim-left 0 \  
--p-trunc-len 270 \  
--p-chimera-method consensus \  
--p-n-threads 0 \  
--p-n-reads-learn 1000000 \  
--o-table $PWDa/dada2/dada2-table.qza \  
--o-representative-sequences $PWDa/dada2/dada2-rep-seqs.qza \  
--o-denoising-stats $PWDa/dada2/dada2-denoising-stats.qza  
  
qiime metadata tabulate \  
--m-input-file $PWDa/dada2/dada2-denoising-stats.qza \  
--o-visualization $PWDa/dada2/dada2-denoising-stats.qzv  
  
qiime feature-table summarize \  
--i-table $PWDa/dada2/dada2-table.qza \  
--o-visualization $PWDa/dada2/dada2-table.qzv \  
--m-sample-metadata-file $PWDa/metadataV34.txt
```

0.6 Taxonomic classification (using primerspecific Naïve Bayes classifier created on SILVA database) and taxonomy-based filtering

```
cd $PWDb/  
gunzip *.gz  
grep -e ">" SILVA_138_SSURef_NR99_tax_silva.fasta > taxonomy.txt  
sed -e 's/\s+/_/' taxonomy.txt > modified_taxonomy.txt
```

```
R  
list.of.packages <- c("plyr", "stringr")  
new.packages <- list.of.packages[!(list.of.packages %in% installed.packages()[,"Package"])]  
if(length(new.packages)) install.packages(new.packages, repos='http://cran.us.r-project.org')  
library(stringr)  
library(plyr)  
data<-read.table("modified_taxonomy.txt", header=FALSE)  
data[,2]<-data[,1]  
datanew<-str_split_fixed(data[,2], ";", 7)  
datanew<-as.data.frame(datanew)  
datanew[,8]<-datanew[,1]  
datanew[,1] <- lapply(datanew, gsub, pattern='>', replacement='')  
datanew[,1] <- lapply(datanew, sub, pattern='.*_', replacement='')  
datanew[,8] <-as.character(datanew[,8])  
datanew[,8] <-gsub("\\\\_.*", "", datanew[,8])  
datanew[,9]<-datanew[,1]  
datanew[,10]<-datanew[,2]  
datanew[,11]<-datanew[,3]  
datanew[,12]<-datanew[,4]  
datanew[,13]<-datanew[,5]  
datanew[,14]<-datanew[,6]  
datanew[,15]<-datanew[,7]  
datanew[,1]<-NULL  
datanew[,1]<-NULL  
datanew[,1]<-NULL  
datanew[,1]<-NULL  
datanew[,1]<-NULL  
datanew[,1]<-NULL
```

```

datanew[,1]<-NULL
datanew[,2]<- paste((rep("k_", length(datanew[,2]))),datanew[,2])
datanew[,3]<- paste((rep("p_", length(datanew[,3]))),datanew[,3])
datanew[,4]<- paste((rep("c_", length(datanew[,4]))),datanew[,4])
datanew[,5]<- paste((rep("o_", length(datanew[,5]))),datanew[,5])
datanew[,6]<- paste((rep("f_", length(datanew[,6]))),datanew[,6])
datanew[,7]<- paste((rep("g_", length(datanew[,7]))),datanew[,7])
datanew[,8]<- paste((rep("s_", length(datanew[,8]))),datanew[,8])
datanew[] <- lapply(datanew, gsub, pattern=' ', replacement='')
datanew[,9]<-
paste(datanew[,2],datanew[,3],datanew[,4],datanew[,5],datanew[,6],datanew[,7],datanew[
datanew[,2]<-NULL
datanew[,2]<-NULL
datanew[,2]<-NULL
datanew[,2]<-NULL
datanew[,2]<-NULL
datanew[,2]<-NULL
datanew[,2]<-NULL
datanew[,2]<-NULL
datanew[,3]<- paste(datanew[,1],datanew[,2], sep=" ")
datanew[,1]<-NULL
datanew[,1]<-NULL
write.table(datanew, file = "taxonomy_7levels.txt",sep="\t",col.names= F, row.names=F)
quit("no")

```

```

sed 's/"/' taxonomy_7levels.txt > taxonomy_7_levels.txt
sed 's/"$/' taxonomy_7_levels.txt > taxonomy_7_levels_fin.txt
rm -r -f -- -taxonomy_7_levels.txt
rm -r -f -- -taxonomy_7levels.txt
rm -r -f -- -modified_taxonomy.txt
rm -r -f -- -taxonomy.txt
sed 's/ /'$'\t''/' taxonomy_7_levels_fin.txt >taxonomy_7_levels.txt
rm -r -f -- -taxonomy_7_levels_fin.txt
sed '/>/s/\s.*//' ' SILVA_138_SSURef_NR99_tax_silva.fasta>silva_138_99_fin.fasta
sed '/^>/! s/[U]/\T/g' silva_138_99_fin.fasta > silva_138_99.fasta
rm -r -f -- - silva_138_99_fin.fasta

```

```

qiime tools import \
--type 'FeatureData[Sequence]' \
--input-path silva_138_99.fasta \
--output-path silva_138_99.qza

```

```

qiime tools import \
--type 'FeatureData[Taxonomy]' \
--input-format HeaderlessTSVTaxonomyFormat \
--input-path taxonomy_7_levels.txt \
--output-path silva138-ref-taxonomy.qza

```

```

qiime feature-classifier extract-reads \
--i-sequences silva_138_99.qza \
--p-f-primer CCTACGGGAGGCAGCAG \
--p-r-primer GGACTACHVGGGTWTCTAAT \
--o-reads silva138V3V4-ref-seqs.qza

```

```

qiime feature-classifier fit-classifier-naive-bayes \
--i-reference-reads silva138V3V4-ref-seqs.qza \
--i-reference-taxonomy silva138-ref-taxonomy.qza \
--o-classifier Silva138V3V4-classifier.qza

```

```

rm -r -f -- -silva138V3V4-ref-seqs.qza
rm -r -f -- -silva138-ref-taxonomy.qza
rm -r -f -- -silva_138_99.qza
rm -r -f -- -silva_138_99.fasta
rm -r -f -- -taxonomy_7_levels.txt
rm -r -f -- -SILVA_138_SSURef_NR99_tax_silva.fasta

```

```

cd $PWDa/

qiime feature-classifier classify-sklearn \
--p-n-jobs 1 \
--p-reads-per-batch 1000 \
--i-classifier $PWDb/Silva138V3V4-classifier.qza \
--i-reads $PWDa/dada2/dada2-rep-seqs.qza \
--p-confidence 0.8 \
--o-classification $PWDa/dada2/Silva138V3V4-taxonomy-dada2.qza

qiime metadata tabulate \
--m-input-file $PWDa/dada2/Silva138V3V4-taxonomy-dada2.qza \
--o-visualization $PWDa/dada2/Silva138V3V4-taxonomy-dada2.qzv

qiime feature-table filter-features \
--i-table $PWDa/dada2/dada2-table.qza \
--m-metadata-file $PWDa/dada2/Silva138V3V4-taxonomy-dada2.qza \
--p-where "Taxon NOT LIKE '%Chloroplast%'" \
--o-filtered-table $PWDa/dada2/dada2-tablenochloroplasts.qza

qiime feature-table filter-features \
--i-table $PWDa/dada2/dada2-tablenochloroplasts.qza \
--m-metadata-file $PWDa/dada2/Silva138V3V4-taxonomy-dada2.qza \
--p-where "Taxon NOT LIKE '%Mitochondria%'" \
--o-filtered-table $PWDa/dada2/dada2-tablenochloroplastsmitchondria.qza

qiime feature-table filter-features \
--i-table $PWDa/dada2/dada2-tablenochloroplastsmitchondria.qza \
--m-metadata-file $PWDa/dada2/Silva138V3V4-taxonomy-dada2.qza \
--p-where "Taxon NOT LIKE '%Unassigned%'" \
--o-filtered-table $PWDa/dada2/dada2-tablenochloroplastsmitchondriaUnassigned.qza

qiime feature-table summarize \
--i-table $PWDa/dada2/dada2-tablenochloroplastsmitchondriaUnassigned.qza \
--o-visualization $PWDa/dada2/dada2-tablenochloroplastsmitchondriaUnassigned.qzv \
--m-sample-metadata-file $PWDa/metadataV34.txt

```

0.7 Set min-frequency and subsample based on sequence frequency

```

cd $PWDa/dada2
unzip dada2-tablenochloroplastsmitchondriaUnassigned.qzv
cd "$(\ls -1dt ./*/ | head -n 1)/data"
sed 's|,|\t|g' sample-frequency-detail.csv > sample-frequency-detail.txt
cp sample-frequency-detail.txt $PWDa/dada2/sample-frequency-detail.txt
cd $PWDa/dada2
echo "5000" >> $PWDa/dada2/min-frequency.txt

R
list.of.packages <- c("tidyr", "plyr", "stringr", "ggplot2", "gridExtra")
new.packages <- list.of.packages[!(list.of.packages %in% installed.packages()[,"Package"])]
if(length(new.packages)) install.packages(new.packages, repos='http://cran.us.r-project.org')
library(tidyr)
library(plyr)
library(stringr)
library(ggplot2)
library(gridExtra)
mydata = read.table(file = "sample-frequency-detail.txt", header = F)
mydata<-rename(mydata, c("V1"="Lost_Samples", "V2"="Frequency"))
pdf(file="sample_frequencies.pdf", height= 25, width= 30)
vari = read.table(file = "min-frequency.txt", header = F)
names(vari)[1] <- "seqdepth"
p<-ggplot(data=mydata, aes(mydata$Frequency)) +

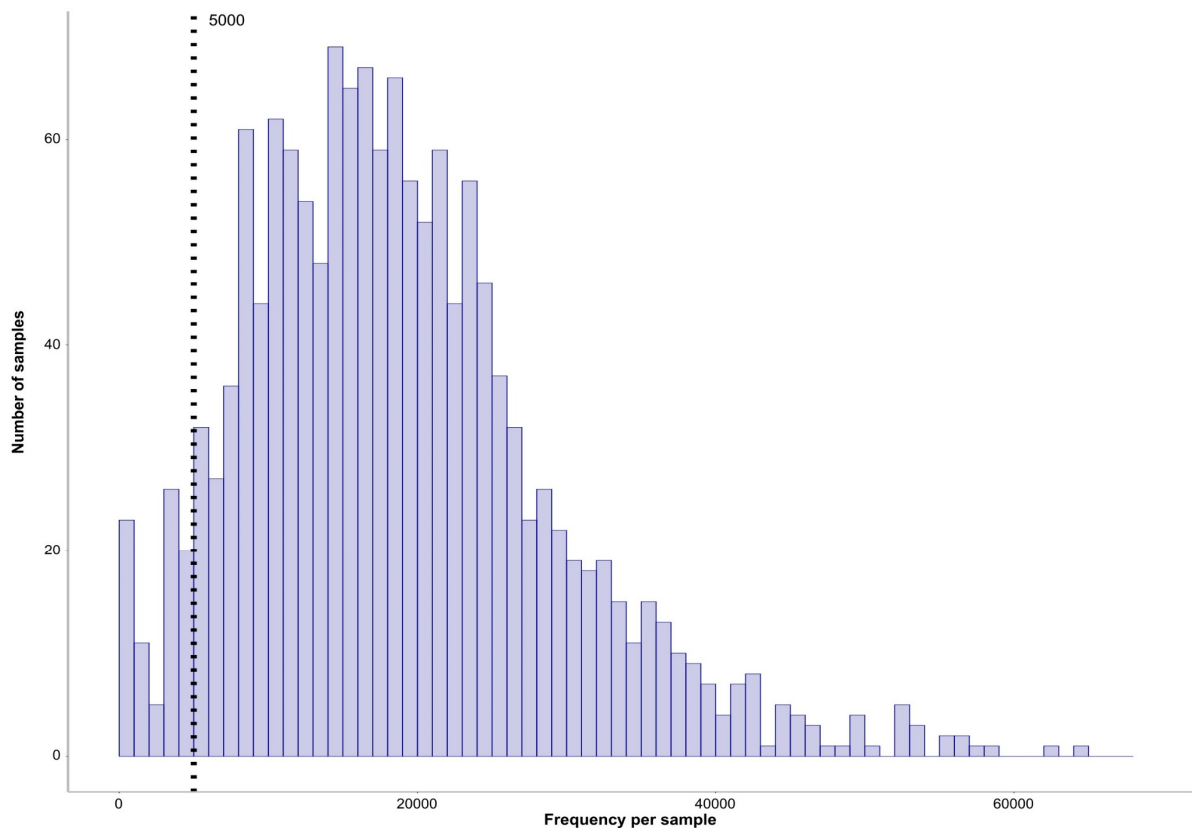
```



```

geom_histogram(breaks=seq(0, max(mydata$Frequency), by=1000), col="darkblue", fill="dark-
blue", alp
labs(x = "Frequency per sample", y="Number of samples") +
xlim(0, max(mydata$Frequency)) +
geom_vline(xintercept = vari[1,1], linetype="dotted", size=1.5)+
annotate("text", vari[1,1]+1000, Inf, label = vari[1,1], size=7, parse=TRUE, hjust = 0,
vjust = 1)
theme(panel.grid.major = element_blank(),
panel.grid.minor = element_blank(),
axis.line = element_line(colour = "grey", size = 1.5, linetype = "solid"),
panel.background = element_rect(fill = "#ffffff",colour = NA),
plot.background = element_rect(fill = '#ffffff'),
legend.position = "none",
axis.title.x = element_text(color="black", size=20, face="bold"),
axis.text.x = element_text(face="plain", color="black",size=18, angle=0),
axis.title.y = element_text(color="black", size=20, face="bold"),
axis.text.y = element_text(face="plain", color="black",size=18, angle=0))
mydata2<-mydata[mydata$Frequency < vari[1,1], ]
write.table(mydata2, file = "sample-frequency-lost_samples.txt",sep="\t",col.names= T,
row.names=F)
rownames(mydata2) <- NULL
p2 <- gridExtra::tableGrob(mydata2)
grid.arrange(p, p2 ,nrow=1, widths = 2:1)
dev.off()
quit("no")

```



```

rm -r -f -- -d */
var=`cat min-frequency.txt`
echo $var
cd $PWDa/

```

```

qiime feature-table filter-samples \
--i-table $PWDa/dada2/dada2-tablenochloroplastsmitochondriaUnassigned.qza \
--p-min-frequency $var \
--o-filtered-table $PWDa/dada2/dada2-tablenochloroplastsmitochondriaUnassigned$var.qza

```

```
qiime feature-table summarize \
--i-table $PWDa/dada2/dada2-tablenochloroplastsmitochondriaUnassigned$var.qza \
--o-visualization $PWDa/dada2/dada2-tablenochloroplastsmitochondriaUnassigned$var.qzv \
--m-sample-metadata-file $PWDa/metadataV34.txt
```

0.8 Visualise Filtering steps and Denoising stats

```
cd $PWDa/dada2/
unzip dada2-denoising-stats.qza
cd "$(\ls -ldt ./ */ | head -n 1)/data"
sed '1 s/ /_/g' stats.tsv >stats2.tsv
sed -e 's/,/\t/g' stats2.tsv > tmpfile; mv tmpfile stats.txt
sed 2d stats.txt > stats3.txt
sed '1 s/-/_/g' stats3.txt >stats4.txt

R
data<- read.table(file="stats4.txt", header =T)
data$percentage_of_input_passed_filter <-NULL
data$percentage_of_input_non_chimeric <-NULL
write.table(data, file = "0stats.txt",sep="\t",col.names= T, row.names=F)
quit("no")

sed 's/"//g' 0stats.txt > 1stats.txt
cp -r "1stats.txt" $PWDa/dada2/
cd $PWDa/dada2/

qiime feature-table summarize \
--i-table $PWDa/dada2/dada2-tablenochloroplasts.qza \
--o-visualization $PWDa/dada2/dada2-tablenochloroplasts.qzv \
--m-sample-metadata-file $PWDa/metadataV34.txt

unzip dada2-tablenochloroplasts.qzv
cd "$(\ls -ldt ./ */ | head -n 1)/data"
sed -e 's/,/\t/g' sample-frequency-detail.csv > tmpfile; mv tmpfile sample-frequency-de-
tail.txt
sed -i '1i sample-id nochloroplasts' sample-frequency-detail.txt
mv sample-frequency-detail.txt 2stats.txt
cp -r "2stats.txt" $PWDa/dada2/
cd $PWDa/dada2/

qiime feature-table summarize \
--i-table $PWDa/dada2/dada2-tablenochloroplastsmitochondria.qza \
--o-visualization $PWDa/dada2/dada2-tablenochloroplastsmitochondria.qzv \
--m-sample-metadata-file $PWDa/metadataV34.txt

unzip dada2-tablenochloroplastsmitochondria.qzv
cd "$(\ls -ldt ./ */ | head -n 1)/data"
sed -e 's/,/\t/g' sample-frequency-detail.csv > tmpfile; mv tmpfile sample-frequency-de-
tail.txt
sed -i '1i sample-id nomitochondria' sample-frequency-detail.txt
mv sample-frequency-detail.txt 3stats.txt
cp -r "3stats.txt" $PWDa/dada2/
cd $PWDa/dada2/
unzip dada2-tablenochloroplastsmitochondriaUnassigned.qzv
cd "$(\ls -ldt ./ */ | head -n 1)/data"
sed -e 's/,/\t/g' sample-frequency-detail.csv > tmpfile; mv tmpfile sample-frequency-de-
tail.txt
sed -i '1i sample-id nounassigned' sample-frequency-detail.txt
mv sample-frequency-detail.txt 4stats.txt
cp -r "4stats.txt" $PWDa/dada2/
cd $PWDa/dada2/
unzip dada2-tablenochloroplastsmitochondriaUnassigned$var.qzv
cd "$(\ls -ldt ./ */ | head -n 1)/data"
sed -e 's/,/\t/g' sample-frequency-detail.csv > tmpfile; mv tmpfile sample-frequency-de-
tail.txt
```

```
sed -i '1i sample-id samplingdepth' sample-frequency-detail.txt
```

```
R
data<- read.table(file="sample-frequency-detail.txt", header =T)
min<-min(data$samplingdepth)
write.table(min, file = "sample_read_count_min_at_seq_depth.txt", sep="\t", col.names= F,
row.names=F)
quit("no")
```

```
mv sample-frequency-detail.txt 5stats.txt
cp -r "5stats.txt" $PWDa/dada2/
cp -r "sample_read_count_min_at_seq_depth.txt" $PWDa/dada2/
cd $PWDa/dada2/
wc -l 5stats.txt > temp.txt
cut -d' ' -f1 temp.txt
vari=$(cut -d' ' -f1 temp.txt )
echo "$(($vari-1))" > sample_number_end.txt
wc -l 1stats.txt > temp2.txt
cut -d' ' -f1 temp2.txt
varis=$(cut -d' ' -f1 temp2.txt )
echo "$(($varis-1))" > sample_number_start.txt
rm -r -f -- - 5stats.txt
rm -r -f -- - temp.txt
rm -r -f -- - temp2.txt
sed '1 s/\-/_/g' 1stats.txt > 1stats2.txt
sed '1 s/\-/_/g' 2stats.txt > 2stats2.txt
sed '1 s/\-/_/g' 3stats.txt > 3stats2.txt
sed '1 s/\-/_/g' 4stats.txt > 4stats2.txt
```

```
R
list.of.packages <- c("plyr", "stringr")
new.packages <- list.of.packages[!(list.of.packages %in% installed.packages()[,"Package"])]
if(length(new.packages)) install.packages(new.packages, repos='http://cran.us.r-project.org')
library(stringr)
library(plyr)
data1<- read.table(file="1stats2.txt", header =T)
data2<- read.table(file="2stats2.txt", header =T)
data3<- read.table(file="3stats2.txt", header =T)
data4<- read.table(file="4stats2.txt", header =T)
x<-merge(data1,data2, by="sample_id")
x2<-merge(x,data3,by="sample_id")
x3<-merge(x2,data4, by="sample_id")
x4<-rename(x3, c("sample_id"="SampleID", "nochloroplasts"="non_chloroplastic", "nomitochondria"="non_m
x4$input_percent<-x4$input/x4$input*100
x4$filtered_percent<-100/x4$input*x4$filtered
x4$denoised_percent<-100/x4$input*x4$denoised
x4$non_chimeric_percent<-100/x4$input*x4$non_chimeric
x4$non_chloroplastic_percent<-100/x4$input*x4$non_chloroplastic
x4$non_mitochondric_percent<-100/x4$input*x4$non_mitochondric
x4$non_unassigned_percent<-100/x4$input*x4$non_unassigned
x4$filtered<-NULL
x4$filtered_percent<-NULL
myvars <- c("SampleID", "input", "denoised", "non_chimeric", "non_chloroplastic", "non_mito-
chondric", "n
absolute <- x4[myvars]
myvars <- c("SampleID", "input_percent", "denoised_percent", "non_chimeric_per-
cent", "non_chloroplasti
relative <- x4[myvars]
```

```
write.table(absolute, file = "denoising_absolute.txt", sep="\t", col.names= T, row.names=F)
write.table(relative, file = "denoising_relative.txt", sep="\t", col.names= T, row.names=F)
quit("no")
```

```
sed 's//g' denoising_absolute.txt > sample_denoising_absolute.txt
sed 's//g' denoising_relative.txt > sample_denoising_relative.txt
rm -r -f -- - 1stats.txt
rm -r -f -- - 1stats2.txt
rm -r -f -- - 2stats.txt
rm -r -f -- - 2stats2.txt
rm -r -f -- - 3stats.txt
rm -r -f -- - 3stats2.txt
rm -r -f -- - 4stats.txt
rm -r -f -- - 4stats2.txt
rm -r -f -- - denoising_absolute.txt
rm -r -f -- - denoising_relative.txt
rm -r -f -- - -d */
unzip dada2-tablechloroplastsmitochondriaUnassigned$var.qzv
cd "$(\ls -1dt ./ */ | head -n 1)/data"
sed -e 's/,/\t/g' feature-frequency-detail.csv > tmpfile; mv tmpfile feature-frequency-detail.txt
sed -i '1i FeatureID frequency' feature-frequency-detail.txt
cp -r "feature-frequency-detail.txt" $PWDa/dada2/
cd $PWDa/dada2/
rm -r -f -- - -d */
```

```
R
list.of.packages <- c("tidyr", "ggplot2", "grid", "gridExtra", "scales")
new.packages <- list.of.packages[!(list.of.packages %in% installed.packages()[,"Package"])]
if(length(new.packages)) install.packages(new.packages, repos='http://cran.us.r-project.org')
library(tidyr)
library(ggplot2)
library(gridExtra)
library(scales)
mydata = read.table(file="sample_denoising_absolute.txt", header =T)
mydata=as.data.frame(mydata)
a<-names(mydata)[1]
b<-names(mydata)[ncol(mydata)]
dat<-gather(mydata, condition, measurement, names(mydata)[2:ncol(mydata)], factor_key=T)
names(dat)[names(dat) == "condition"] <- "filtering_step"
names(dat)[names(dat) == "measurement"] <- "read_count"
addline_format <- function(x,...){gsub('_', '-\n',x)}
multiplot <- function(..., plotlist=NULL, file, cols=1, layout=NULL) {
library(grid)
plots <- c(list(...), plotlist)
numPlots = length(plots)
if (is.null(layout)) {
layout <- matrix(seq(1, cols * ceiling(numPlots/cols)),
ncol = cols, nrow = ceiling(numPlots/cols))
}
if (numPlots==1) {
print(plots[[1]])
} else {
grid.newpage()
pushViewport(viewport(layout = grid.layout(nrow(layout), ncol(layout))))
for (i in 1:numPlots) {
matchidx <- as.data.frame(which(layout == i, arr.ind = TRUE))
print(plots[[i]], vp = viewport(layout.pos.row = matchidx$row,
layout.pos.col = matchidx$col))
}
}
```

```

}
}
col <- data.frame(name = c("Mean"),
start = c(0.00001,0.00002),
end = c(0.00001,0.00002),
percentile = c("Mean"),
stringsAsFactors = FALSE)
pdf(file="see_denoising.pdf", height= 17, width= 25)
W <- subset(dat, filtering_step == "input")
p1<-ggplot(data=dat, aes(x=filtering_step, y=read_count)) +
stat_summary(aes(y = read_count,group=1), fun.y=sum, colour="black",
geom="line",group=1,linetype=
labs(x = "Filtering steps", y="Total read count", fill="") +
geom_rect(data=col, aes(NULL,NULL, ymin=start, ymax=end, fill=percentile), xmin=0.00001,
xmax=0.00
scale_fill_manual(values=c("Mean" = "white")) +
scale_x_discrete(breaks=unique(dat$filtering_step),
labels=addline_format(c("input", "denoised","non_chimeric","non_chloroplastic",
"non_mitochondric", "non_unassigned")))+
ylim(0, sum(W$read_count)) +
theme(panel.grid.major = element_blank(),
panel.grid.minor = element_blank(),
axis.line = element_line(colour = "grey", size = 1.5, linetype = "solid"),
panel.background = element_rect(fill = "#ffffff",colour = NA),
plot.background = element_rect(fill = '#ffffff'),
legend.text = element_text(color = "white", size = 18),
legend.position = c(0.96, 0.1),
axis.title.x = element_text(color="black", size=20, face="bold"),
axis.text.x = element_text(face="plain", color="black",size=18, angle=0),
axis.title.y = element_text(color="black", size=20, face="bold"),
axis.text.y = element_text(face="plain", color="black",size=18, angle=0))
p2<-ggplot(data=dat, aes(x=filtering_step, y=read_count)) +
geom_point(colour="black", alpha=0.2) +
stat_summary(aes(y = read_count,group=1), fun.y=mean, colour="#cecee8ff",
geom="line",group=1,line
labs(x = "Filtering steps", y="Read count per sample", fill="") +
geom_rect(data=col, aes(NULL,NULL, ymin=start, ymax=end, fill=percentile), xmin=0.00001,
xmax=0.00
scale_fill_manual(values=c("Mean" = "#cecee8ff")) +
scale_x_discrete(breaks=unique(dat$filtering_step),
labels=addline_format(c("input", "denoised","non_chimeric","non_chloroplastic",
"non_mitochondric", "non_unassigned")))+
theme(panel.grid.major = element_blank(),
panel.grid.minor = element_blank(),
axis.line = element_line(colour = "grey", size = 1.5, linetype = "solid"),
panel.background = element_rect(fill = "#ffffff",colour = NA),
plot.background = element_rect(fill = '#ffffff'),
legend.text = element_text(color = "black", size = 18),
legend.position = c(0.96, 0.1),
axis.title.x = element_text(color="black", size=20, face="bold"),
axis.text.x = element_text(face="plain", color="black",size=18, angle=0),
axis.title.y = element_text(color="black", size=20, face="bold"),
axis.text.y = element_text(face="plain", color="black",size=18, angle=0))
bar = read.table(file="sample_denoising_relative.txt", header =T)
bar=as.data.frame(bar)
a<-names(bar)[1]
b<-names(bar)[ncol(bar)]
bar2<-gather(bar,condition, measurement, names(bar)[2:ncol(bar)], factor_key=T)
names(bar2)[names(bar2) == "condition"] <- "filtering_step"
names(bar2)[names(bar2) == "measurement"] <- "read_count"
bar3<-aggregate(bar2[, 3], list(bar2$filtering_step), mean)
bar3$z<-rep(1, nrow(bar3))
#bar3$a<-(1:nrow(bar3))
names(bar3)[1] <- "filtering_step"
names(bar3)[2] <- "read_percent"
df1 <- subset(bar3,bar3$filtering_step=="non_unassigned_percent")

```

```

df1$a<-df1$read_percent-0
df2 <- subset(bar3,bar3$filtering_step=="non_mitochondric_percent")
df2$a<-df2$read_percent-df1$read_percent
df3 <- subset(bar3,bar3$filtering_step=="non_chloroplastic_percent")
df3$a<-df3$read_percent-df2$read_percent
df4 <- subset(bar3,bar3$filtering_step=="non_chimeric_percent")
df4$a<-df4$read_percent-df3$read_percent
df5 <- subset(bar3,bar3$filtering_step=="denoised_percent")
df5$a<-df5$read_percent-df4$read_percent
df6 <- subset(bar3,bar3$filtering_step=="input_percent")
df6$a<-df6$read_percent-df5$read_percent
u<- rbind(df6, df5, df4, df3, df2, df1)
p3<-ggplot(data=u, aes(x=z, y=a, fill=filtering_step)) +
geom_bar(position="stack", stat="identity",width = 0.2) +
scale_fill_manual(values=c("#cee1f0ff", "#9db8d6ff", "#667eb1ff", "#354567ff", "#657090ff",
"#e5e5f"),
name = "Filtering steps",
labels = c("Denoising", "Chimera removal", "Chloroplast removal", "Mitochondria removal",
"Unassignig"),
labs(x = "", y="Mean fraction of input read count \n [% per sample]") +
theme(panel.grid.major = element_blank(),
panel.grid.minor = element_blank(),
axis.line = element_line(colour = "grey", size = 1.5, linetype = "solid"),
panel.background = element_rect(fill = "#ffffff",colour = NA),
plot.background = element_rect(fill = '#ffffff'),
legend.title= element_text(color = "black", size = 18, face="bold"),
legend.text = element_text(color = "black", size = 18),
axis.title.x=element_blank(),
axis.text.x=element_blank(),
axis.ticks.x=element_blank(),
axis.title.y = element_text(color="black", size=20, face="bold"),
axis.text.y = element_text(face="plain", color="black",size=18, angle=0))
dat3<-aggregate(dat[, 3], list(dat$filtering_step), sum)
dat3$z<-rep(1, nrow(dat3))
names(dat3)[1] <- "filtering_step"
names(dat3)[2] <- "read_absolute"
df7 <- subset(dat3,dat3$filtering_step=="input")
df7$read_percent<-100/df7$read_absolute*df7$read_absolute
df8 <- subset(dat3,dat3$filtering_step=="non_unassigned")
df8$read_percent<-100/df7$read_absolute*df8$read_absolute
df9 <- subset(dat3,dat3$filtering_step=="non_mitochondric")
df9$read_percent<-100/df7$read_absolute*df9$read_absolute
df10 <- subset(dat3,dat3$filtering_step=="non_chloroplastic")
df10$read_percent<-100/df7$read_absolute*df10$read_absolute
df11 <- subset(dat3,dat3$filtering_step=="non_chimeric")
df11$read_percent<-100/df7$read_absolute*df11$read_absolute
df12 <- subset(dat3,dat3$filtering_step=="denoised")
df12$read_percent<-100/df7$read_absolute*df12$read_absolute
df13<-df7
df8$a<-df8$read_percent-0
df9$a<-df9$read_percent-df8$read_percent
df10$a<-df10$read_percent-df9$read_percent
df11$a<-df11$read_percent-df10$read_percent
df12$a<-df12$read_percent-df11$read_percent
df13$a<-df13$read_percent-df12$read_percent
u2<- rbind(df13, df12, df11, df10, df9, df8)
p4<-ggplot(data=u2, aes(x=z, y=a, fill=filtering_step)) +
geom_bar(position="stack", stat="identity",width = 0.2) +
scale_fill_manual(values=c("#cee1f0ff", "#9db8d6ff", "#667eb1ff", "#354567ff", "#657090ff",
"#e5e5f"),
name = "Filtering steps",
labels = c("Denoising", "Chimera removal", "Chloroplast removal", "Mitochondria
labs(x = "", y="Fraction of total input read count \n [%]") +
theme(panel.grid.major = element_blank(),
panel.grid.minor = element_blank(),
axis.line = element_line(colour = "grey", size = 1.5, linetype = "solid"),

```

```

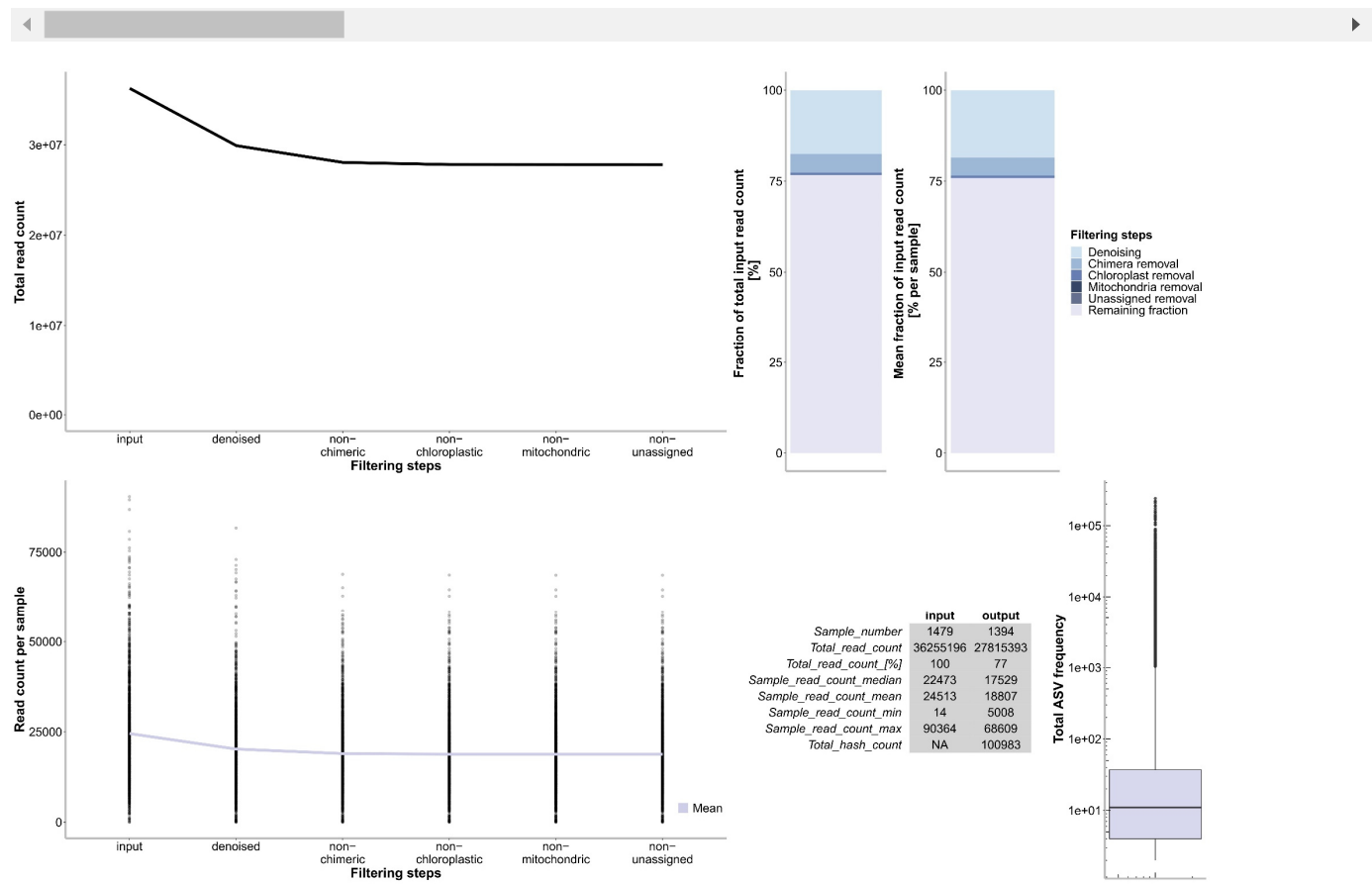
panel.background = element_rect(fill = "#ffffff", colour = NA),
plot.background = element_rect(fill = '#ffffff'),
legend.title= element_text(color = "black", size = 18, face="bold"),
legend.text = element_text(color = "black", size = 18),
legend.position = "none",
axis.title.x=element_blank(),
axis.text.x=element_blank(),
axis.ticks.x=element_blank(),
axis.title.y = element_text(color="black", size=20, face="bold"),
axis.text.y = element_text(face="plain", color="black",size=18, angle=0))
u2$z<-NULL
u2$a<-NULL
u3 <- subset(u2,filtering_step=="input"| filtering_step=="non_unassigned")
names(u3)[names(u3) == "read_absolute"] <- "Total_read_count"
names(u3)[names(u3) == "read_percent"] <- "Total_read_count_["%"
u4 <- subset(dat,filtering_step=="input")
u4$Sample_read_count_median<-median(u4$read_count)
u4$Sample_read_count_mean<-mean(u4$read_count)
u4$Sample_read_count_min<-min(u4$read_count)
u4$Sample_read_count_max<-max(u4$read_count)
u4$SampleID<-NULL
u4$read_count<-NULL
input<-u4[1,]
u5 <- subset(dat,filtering_step=="non_unassigned")
u5$Sample_read_count_median<-median(u5$read_count)
u5$Sample_read_count_mean<-mean(u5$read_count)
u5$Sample_read_count_min<-min(u5$read_count)
u5$Sample_read_count_max<-max(u5$read_count)
u5$SampleID<-NULL
u5$read_count<-NULL
output<-u5[1,]
o<-rbind(input,output)
sum<-merge(u3,o,by="filtering_step")
noin = scan("sample_number_start.txt")
nout = scan("sample_number_end.txt")
noin2 <- data.frame(matrix(unlist(noin), nrow=length(noin), byrow=T))
names(noin2)[1] <- "Sample_number"
noin2$filtering_step<-rep(c("input"), nrow(noin2))
nout2 <- data.frame(matrix(unlist(nout), nrow=length(nout), byrow=T))
names(nout2)[1] <- "Sample_number"
nout2$filtering_step<-rep(c("non_unassigned"), nrow(nout2))
fin<-rbind(noin2,nout2)
sum2<-merge(fin,sum,by="filtering_step")
sum3<-t(sum2)
sum3<-as.data.frame(sum3)
names(sum3) <- as.matrix(sum3[1, ])
sum3 <- sum3[-1, ]
names(sum3)[2] <- "output"
write.table(sum3, file = "sum3.txt", sep="\t", col.names= T, row.names=T)
sum3 = read.table(file="sum3.txt", header =T)
round_df <- function(x, digits) {
numeric_columns <- sapply(x, mode) == 'numeric'
x[numeric_columns] <- round(x[numeric_columns], digits)
x
}
sum4<-round_df(sum3, 0)
feat = read.table(file="feature-frequency-detail.txt", header =T)
feat$b<-rep(1, nrow(feat))
p6<-ggplot(feat, aes(x=b, y=frequency)) +
geom_boxplot(fill="#cecee8ff", alpha=0.8) +
scale_y_continuous(trans = 'log10') +
annotation_logticks() +
labs(x = "", y="Total ASV frequency") +
theme(panel.grid.major = element_blank(),
panel.grid.minor = element_blank(),
axis.line = element_line(colour = "grey", size = 1.5, linetype = "solid"),

```

```

panel.background = element_rect(fill = "#ffffff", colour = NA),
plot.background = element_rect(fill = '#ffffff'),
legend.title= element_text(color = "black", size = 18, face="bold"),
legend.text = element_text(color = "black", size = 18),
legend.position = "none",
axis.title.x=element_blank(),
axis.text.x=element_blank(),
axis.ticks.x=element_blank(),
axis.title.y = element_text(color="black", size=20, face="bold"),
axis.text.y = element_text(face="plain", color="black",size=18, angle=0)
l<-c("NA",length(feats$FeatureID))
l <- data.frame(matrix(unlist(l), nrow=length(l), byrow=T))
names(l)[1] <- "Total_hash_count"
l<-t(l)
l<-as.data.frame(l)
names(l)[1] <- "input"
names(l)[2] <- "output"
sum5<-rbind(sum4,l)
e = scan(file="sample_read_count_min_at_seq_depth.txt")
e2 <- data.frame(matrix(unlist(e), nrow=length(e), byrow=T))
names(e2)[1] <- "Sample_read_count_min"
sum5[6,2] <-e2[1,1]
mytheme <- gridExtra::theme_minimal(
core = list(fg_params=list(cex = 1.5),bg_params = list(fill = "lightgrey", col=NA)),
colhead = list(fg_params=list(cex = 1.5)),
rowhead = list(fg_params=list(cex = 1.5)))
p5 <- gridExtra::tableGrob(sum5, theme = mytheme)
grid.arrange(p1, arrangeGrob(p4, p3,nrow=1, widths = 1:2), p2, widths = 3:2, arrangeGrob(p5,
p6,nro
dev.off()
quit("no")

```



0.9 Create phylogenetic backbone tree and compute core-diversity metrics

```
qiime phylogeny align-to-tree-mafft-fasttree \  
--p-n-threads 0 \  
--i-sequences $PWDa/dada2/dada2-rep-seqs.qza \  
--o-alignment $PWDa/dada2/aligned-rep-seqs.qza \  
--o-masked-alignment $PWDa/dada2/masked-aligned-rep-seqs.qza \  
--o-tree $PWDa/dada2/unrooted-tree.qza \  
--o-rooted-tree $PWDa/dada2/rooted-tree.qza  
  
qiime diversity core-metrics-phylogenetic \  
--p-n-jobs 8 \  
--i-phylogeny $PWDa/dada2/rooted-tree.qza \  
--i-table $PWDa/dada2/dada2-tablenochloroplastsmitochondriaUnassigned$var.qza \  
--p-sampling-depth $var \  
--m-metadata-file $PWDa/metadataV34.txt \  
--output-dir $PWDa/dada2/core-metrics-results
```

0.10 Compute rarefaction curves

```
qiime diversity alpha-rarefaction \  
--i-table $PWDa/dada2/dada2-tablenochloroplastsmitochondriaUnassigned$var.qza \  
--i-phylogeny $PWDa/dada2/rooted-tree.qza \  
--p-max-depth $var \  
--p-steps 100 \  
--p-iterations 10 \  
--m-metadata-file $PWDa/metadataV34.txt \  
--o-visualization $PWDa/dada2/core-metrics-results/alpha-rarefaction.qzv  
  
cd $PWDa/dada2/  
awk 'BEGIN {max = 0} {if ($2>max) max=$2} END {print max}' sample-frequency-detail.txt >  
max.txt  
int=`cat max.txt`  
variable=${int%.*}  
echo $variable  
  
qiime diversity alpha-rarefaction \  
--i-table $PWDa/dada2/dada2-tablenochloroplastsmitochondriaUnassigned.qza \  
--i-phylogeny $PWDa/dada2/rooted-tree.qza \  
--p-max-depth $variable \  
--p-steps 100 \  
--p-iterations 10 \  
--m-metadata-file $PWDa/metadataV34.txt \  
--o-visualization $PWDa/dada2/core-metrics-results/alpha-rarefaction_complete.qzv  
  
cd $PWDa/dada2/core-metrics-results/  
unzip alpha-rarefaction.qzv  
cd "$(\ls -ldt ./ */ | head -n 1)/data"  
sed '1 s/-//g' observed_otus.csv > observed_otus2.csv  
sed -e 's/,/\t/g' observed_otus2.csv > tmpfile; mv tmpfile observed_otus2.txt  
cp -r "observed_otus2.txt" $PWDa/dada2/  
cd $PWDa/dada2/  
  
R  
list.of.packages <- c("ggplot2", "stringr")  
new.packages <- list.of.packages[!(list.of.packages %in% installed.packages()), "Package"]]  
if(length(new.packages)) install.packages(new.packages, repos='http://cran.us.r-project.org')  
library(stringr)  
library(ggplot2)  
data<- read.table(file="observed_otus2.txt", header =T)  
dat <- data[,-1]  
rownames(dat) <- data[,1]  
depth1<-dat[,1:10]
```

```
depth2<-dat[,11:20]
depth3<-dat[,21:30]
depth4<-dat[,31:40]
depth5<-dat[,41:50]
depth6<-dat[,51:60]
depth7<-dat[,61:70]
depth8<-dat[,71:80]
depth9<-dat[,81:90]
depth10<-dat[,91:100]
depth11<-dat[,101:110]
depth12<-dat[,111:120]
depth13<-dat[,121:130]
depth14<-dat[,131:140]
depth15<-dat[,141:150]
depth16<-dat[,151:160]
depth17<-dat[,161:170]
depth18<-dat[,171:180]
depth19<-dat[,181:190]
depth20<-dat[,191:200]
depth21<-dat[,201:210]
depth22<-dat[,211:220]
depth23<-dat[,221:230]
depth24<-dat[,231:240]
depth25<-dat[,241:250]
depth26<-dat[,251:260]
depth27<-dat[,261:270]
depth28<-dat[,271:280]
depth29<-dat[,281:290]
depth30<-dat[,291:300]
depth31<-dat[,301:310]
depth32<-dat[,311:320]
depth33<-dat[,321:330]
depth34<-dat[,331:340]
depth35<-dat[,341:350]
depth36<-dat[,351:360]
depth37<-dat[,361:370]
depth38<-dat[,371:380]
depth39<-dat[,381:390]
depth40<-dat[,391:400]
depth41<-dat[,401:410]
depth42<-dat[,411:420]
depth43<-dat[,421:430]
depth44<-dat[,431:440]
depth45<-dat[,441:450]
depth46<-dat[,451:460]
depth47<-dat[,461:470]
depth48<-dat[,471:480]
depth49<-dat[,481:490]
depth50<-dat[,491:500]
depth51<-dat[,501:510]
depth52<-dat[,511:520]
depth53<-dat[,521:530]
depth54<-dat[,531:540]
depth55<-dat[,541:550]
depth56<-dat[,551:560]
depth57<-dat[,561:570]
depth58<-dat[,571:580]
depth59<-dat[,581:590]
depth60<-dat[,591:600]
depth61<-dat[,601:610]
depth62<-dat[,611:620]
depth63<-dat[,621:630]
depth64<-dat[,631:640]
depth65<-dat[,641:650]
depth66<-dat[,651:660]
depth67<-dat[,661:670]
depth68<-dat[,671:680]
```

```

depth69<-dat[,681:690]
depth70<-dat[,691:700]
depth71<-dat[,701:710]
depth72<-dat[,711:720]
depth73<-dat[,721:730]
depth74<-dat[,731:740]
depth75<-dat[,741:750]
depth76<-dat[,751:760]
depth77<-dat[,761:770]
depth78<-dat[,771:780]
depth79<-dat[,781:790]
depth80<-dat[,791:800]
depth81<-dat[,801:810]
depth82<-dat[,811:820]
depth83<-dat[,821:830]
depth84<-dat[,831:840]
depth85<-dat[,841:850]
depth86<-dat[,851:860]
depth87<-dat[,861:870]
depth88<-dat[,871:880]
depth89<-dat[,881:890]
depth90<-dat[,891:900]
depth91<-dat[,901:910]
depth92<-dat[,911:920]
depth93<-dat[,921:930]
depth94<-dat[,931:940]
depth95<-dat[,941:950]
depth96<-dat[,951:960]
depth97<-dat[,961:970]
depth98<-dat[,971:980]
depth99<-dat[,981:990]
depth100<-dat[,991:1000]
metadata<-dat[,1001:ncol(dat)]
depth1$mean<-rowMeans(depth1)
depth2$mean<-rowMeans(depth2)
depth3$mean<-rowMeans(depth3)
depth4$mean<-rowMeans(depth4)
depth5$mean<-rowMeans(depth5)
depth6$mean<-rowMeans(depth6)
depth7$mean<-rowMeans(depth7)
depth8$mean<-rowMeans(depth8)
depth9$mean<-rowMeans(depth9)
depth10$mean<-rowMeans(depth10)
depth11$mean<-rowMeans(depth11)
depth12$mean<-rowMeans(depth12)
depth13$mean<-rowMeans(depth13)
depth14$mean<-rowMeans(depth14)
depth15$mean<-rowMeans(depth15)
depth16$mean<-rowMeans(depth16)
depth17$mean<-rowMeans(depth17)
depth18$mean<-rowMeans(depth18)
depth19$mean<-rowMeans(depth19)
depth20$mean<-rowMeans(depth20)
depth21$mean<-rowMeans(depth21)
depth22$mean<-rowMeans(depth22)
depth23$mean<-rowMeans(depth23)
depth24$mean<-rowMeans(depth24)
depth25$mean<-rowMeans(depth25)
depth26$mean<-rowMeans(depth26)
depth27$mean<-rowMeans(depth27)
depth28$mean<-rowMeans(depth28)
depth29$mean<-rowMeans(depth29)
depth30$mean<-rowMeans(depth30)
depth31$mean<-rowMeans(depth31)
depth32$mean<-rowMeans(depth32)
depth33$mean<-rowMeans(depth33)

```

```
depth34$mean<-rowMeans(depth34)
depth35$mean<-rowMeans(depth35)
depth36$mean<-rowMeans(depth36)
depth37$mean<-rowMeans(depth37)
depth38$mean<-rowMeans(depth38)
depth39$mean<-rowMeans(depth39)
depth40$mean<-rowMeans(depth40)
depth41$mean<-rowMeans(depth41)
depth42$mean<-rowMeans(depth42)
depth43$mean<-rowMeans(depth43)
depth44$mean<-rowMeans(depth44)
depth45$mean<-rowMeans(depth45)
depth46$mean<-rowMeans(depth46)
depth47$mean<-rowMeans(depth47)
depth48$mean<-rowMeans(depth48)
depth49$mean<-rowMeans(depth49)
depth50$mean<-rowMeans(depth50)
depth51$mean<-rowMeans(depth51)
depth52$mean<-rowMeans(depth52)
depth53$mean<-rowMeans(depth53)
depth54$mean<-rowMeans(depth54)
depth55$mean<-rowMeans(depth55)
depth56$mean<-rowMeans(depth56)
depth57$mean<-rowMeans(depth57)
depth58$mean<-rowMeans(depth58)
depth59$mean<-rowMeans(depth59)
depth60$mean<-rowMeans(depth60)
depth61$mean<-rowMeans(depth61)
depth62$mean<-rowMeans(depth62)
depth63$mean<-rowMeans(depth63)
depth64$mean<-rowMeans(depth64)
depth65$mean<-rowMeans(depth65)
depth66$mean<-rowMeans(depth66)
depth67$mean<-rowMeans(depth67)
depth68$mean<-rowMeans(depth68)
depth69$mean<-rowMeans(depth69)
depth70$mean<-rowMeans(depth70)
depth71$mean<-rowMeans(depth71)
depth72$mean<-rowMeans(depth72)
depth73$mean<-rowMeans(depth73)
depth74$mean<-rowMeans(depth74)
depth75$mean<-rowMeans(depth75)
depth76$mean<-rowMeans(depth76)
depth77$mean<-rowMeans(depth77)
depth78$mean<-rowMeans(depth78)
depth79$mean<-rowMeans(depth79)
depth80$mean<-rowMeans(depth80)
depth81$mean<-rowMeans(depth81)
depth82$mean<-rowMeans(depth82)
depth83$mean<-rowMeans(depth83)
depth84$mean<-rowMeans(depth84)
depth85$mean<-rowMeans(depth85)
depth86$mean<-rowMeans(depth86)
depth87$mean<-rowMeans(depth87)
depth88$mean<-rowMeans(depth88)
depth89$mean<-rowMeans(depth89)
depth90$mean<-rowMeans(depth90)
depth91$mean<-rowMeans(depth91)
depth92$mean<-rowMeans(depth92)
depth93$mean<-rowMeans(depth93)
depth94$mean<-rowMeans(depth94)
depth95$mean<-rowMeans(depth95)
depth96$mean<-rowMeans(depth96)
depth97$mean<-rowMeans(depth97)
depth98$mean<-rowMeans(depth98)
depth99$mean<-rowMeans(depth99)
```

```

depth100$mean<-rowMeans(depth100)
names(depth1)<-sapply(strsplit(names(depth1), "_iter"), `[[`, 1)
colnames(depth1)[1]
depth1$seqdepth<-rep(colnames(depth1)[1], nrow(depth1))
depth1[,1:10]<-NULL
depth1$seqdepth<-str_remove_all(depth1$seqdepth, "[depth]")
depth1<-merge(depth1, metadata, by="row.names")
names(depth2)<-sapply(strsplit(names(depth2), "_iter"), `[[`, 1)
colnames(depth2)[1]
depth2$seqdepth<-rep(colnames(depth2)[1], nrow(depth2))
depth2[,1:10]<-NULL
depth2$seqdepth<-str_remove_all(depth2$seqdepth, "[depth]")
depth2<-merge(depth2, metadata, by="row.names")
names(depth3)<-sapply(strsplit(names(depth3), "_iter"), `[[`, 1)
colnames(depth3)[1]
depth3$seqdepth<-rep(colnames(depth3)[1], nrow(depth3))
depth3[,1:10]<-NULL
depth3$seqdepth<-str_remove_all(depth3$seqdepth, "[depth]")
depth3<-merge(depth3, metadata, by="row.names")
names(depth4)<-sapply(strsplit(names(depth4), "_iter"), `[[`, 1)
colnames(depth4)[1]
depth4$seqdepth<-rep(colnames(depth4)[1], nrow(depth4))
depth4[,1:10]<-NULL
depth4$seqdepth<-str_remove_all(depth4$seqdepth, "[depth]")
depth4<-merge(depth4, metadata, by="row.names")
names(depth5)<-sapply(strsplit(names(depth5), "_iter"), `[[`, 1)
colnames(depth5)[1]
depth5$seqdepth<-rep(colnames(depth5)[1], nrow(depth5))
depth5[,1:10]<-NULL
depth5$seqdepth<-str_remove_all(depth5$seqdepth, "[depth]")
depth5<-merge(depth5, metadata, by="row.names")
names(depth6)<-sapply(strsplit(names(depth6), "_iter"), `[[`, 1)
colnames(depth6)[1]
depth6$seqdepth<-rep(colnames(depth6)[1], nrow(depth6))
depth6[,1:10]<-NULL
depth6$seqdepth<-str_remove_all(depth6$seqdepth, "[depth]")
depth6<-merge(depth6, metadata, by="row.names")
names(depth7)<-sapply(strsplit(names(depth7), "_iter"), `[[`, 1)
colnames(depth7)[1]
depth7$seqdepth<-rep(colnames(depth7)[1], nrow(depth7))
depth7[,1:10]<-NULL
depth7$seqdepth<-str_remove_all(depth7$seqdepth, "[depth]")
depth7<-merge(depth7, metadata, by="row.names")
names(depth8)<-sapply(strsplit(names(depth8), "_iter"), `[[`, 1)
colnames(depth8)[1]
depth8$seqdepth<-rep(colnames(depth8)[1], nrow(depth8))
depth8[,1:10]<-NULL
depth8$seqdepth<-str_remove_all(depth8$seqdepth, "[depth]")
depth8<-merge(depth8, metadata, by="row.names")
names(depth9)<-sapply(strsplit(names(depth9), "_iter"), `[[`, 1)
colnames(depth9)[1]
depth9$seqdepth<-rep(colnames(depth9)[1], nrow(depth9))
depth9[,1:10]<-NULL
depth9$seqdepth<-str_remove_all(depth9$seqdepth, "[depth]")
depth9<-merge(depth9, metadata, by="row.names")
names(depth10)<-sapply(strsplit(names(depth10), "_iter"), `[[`, 1)
colnames(depth10)[1]
depth10$seqdepth<-rep(colnames(depth10)[1], nrow(depth10))
depth10[,1:10]<-NULL
depth10$seqdepth<-str_remove_all(depth10$seqdepth, "[depth]")
depth10<-merge(depth10, metadata, by="row.names")
names(depth11)<-sapply(strsplit(names(depth11), "_iter"), `[[`, 1)
colnames(depth11)[1]
depth11$seqdepth<-rep(colnames(depth11)[1], nrow(depth11))
depth11[,1:10]<-NULL
depth11$seqdepth<-str_remove_all(depth11$seqdepth, "[depth]")

```

```

depth11<-merge(depth11, metadata, by="row.names")
names(depth12)<-sapply(strsplit(names(depth12), "_iter"), `[`, 1)
colnames(depth12)[1]
depth12$seqdepth<-rep(colnames(depth12)[1], nrow(depth12))
depth12[,1:10]<-NULL
depth12$seqdepth<-str_remove_all(depth12$seqdepth, "[depth]")
depth12<-merge(depth12, metadata, by="row.names")
names(depth13)<-sapply(strsplit(names(depth13), "_iter"), `[`, 1)
colnames(depth13)[1]
depth13$seqdepth<-rep(colnames(depth13)[1], nrow(depth13))
depth13[,1:10]<-NULL
depth13$seqdepth<-str_remove_all(depth13$seqdepth, "[depth]")
depth13<-merge(depth13, metadata, by="row.names")
names(depth14)<-sapply(strsplit(names(depth14), "_iter"), `[`, 1)
colnames(depth14)[1]
depth14$seqdepth<-rep(colnames(depth14)[1], nrow(depth14))
depth14[,1:10]<-NULL
depth14$seqdepth<-str_remove_all(depth14$seqdepth, "[depth]")
depth14<-merge(depth14, metadata, by="row.names")
names(depth15)<-sapply(strsplit(names(depth15), "_iter"), `[`, 1)
colnames(depth15)[1]
depth15$seqdepth<-rep(colnames(depth15)[1], nrow(depth15))
depth15[,1:10]<-NULL
depth15$seqdepth<-str_remove_all(depth15$seqdepth, "[depth]")
depth15<-merge(depth15, metadata, by="row.names")
names(depth16)<-sapply(strsplit(names(depth16), "_iter"), `[`, 1)
colnames(depth16)[1]
depth16$seqdepth<-rep(colnames(depth16)[1], nrow(depth16))
depth16[,1:10]<-NULL
depth16$seqdepth<-str_remove_all(depth16$seqdepth, "[depth]")
depth16<-merge(depth16, metadata, by="row.names")
names(depth17)<-sapply(strsplit(names(depth17), "_iter"), `[`, 1)
colnames(depth17)[1]
depth17$seqdepth<-rep(colnames(depth17)[1], nrow(depth17))
depth17[,1:10]<-NULL
depth17$seqdepth<-str_remove_all(depth17$seqdepth, "[depth]")
depth17<-merge(depth17, metadata, by="row.names")
names(depth18)<-sapply(strsplit(names(depth18), "_iter"), `[`, 1)
colnames(depth18)[1]
depth18$seqdepth<-rep(colnames(depth18)[1], nrow(depth18))
depth18[,1:10]<-NULL
depth18$seqdepth<-str_remove_all(depth18$seqdepth, "[depth]")
depth18<-merge(depth18, metadata, by="row.names")
names(depth19)<-sapply(strsplit(names(depth19), "_iter"), `[`, 1)
colnames(depth19)[1]
depth19$seqdepth<-rep(colnames(depth19)[1], nrow(depth19))
depth19[,1:10]<-NULL
depth19$seqdepth<-str_remove_all(depth19$seqdepth, "[depth]")
depth19<-merge(depth19, metadata, by="row.names")
names(depth20)<-sapply(strsplit(names(depth20), "_iter"), `[`, 1)
colnames(depth20)[1]
depth20$seqdepth<-rep(colnames(depth20)[1], nrow(depth20))
depth20[,1:10]<-NULL
depth20$seqdepth<-str_remove_all(depth20$seqdepth, "[depth]")
depth20<-merge(depth20, metadata, by="row.names")
names(depth21)<-sapply(strsplit(names(depth21), "_iter"), `[`, 1)
colnames(depth21)[1]
depth21$seqdepth<-rep(colnames(depth21)[1], nrow(depth21))
depth21[,1:10]<-NULL
depth21$seqdepth<-str_remove_all(depth21$seqdepth, "[depth]")
depth21<-merge(depth21, metadata, by="row.names")
names(depth22)<-sapply(strsplit(names(depth22), "_iter"), `[`, 1)
colnames(depth22)[1]
depth22$seqdepth<-rep(colnames(depth22)[1], nrow(depth22))
depth22[,1:10]<-NULL
depth22$seqdepth<-str_remove_all(depth22$seqdepth, "[depth]")

```

```

depth22<-merge(depth22, metadata, by="row.names")
names(depth23)<-sapply(strsplit(names(depth23), "_iter"), `[`, 1)
colnames(depth23)[1]
depth23$seqdepth<-rep(colnames(depth23)[1], nrow(depth23))
depth23[,1:10]<-NULL
depth23$seqdepth<-str_remove_all(depth23$seqdepth, "[depth]")
depth23<-merge(depth23, metadata, by="row.names")
names(depth24)<-sapply(strsplit(names(depth24), "_iter"), `[`, 1)
colnames(depth24)[1]
depth24$seqdepth<-rep(colnames(depth24)[1], nrow(depth24))
depth24[,1:10]<-NULL
depth24$seqdepth<-str_remove_all(depth24$seqdepth, "[depth]")
depth24<-merge(depth24, metadata, by="row.names")
names(depth25)<-sapply(strsplit(names(depth25), "_iter"), `[`, 1)
colnames(depth25)[1]
depth25$seqdepth<-rep(colnames(depth25)[1], nrow(depth25))
depth25[,1:10]<-NULL
depth25$seqdepth<-str_remove_all(depth25$seqdepth, "[depth]")
depth25<-merge(depth25, metadata, by="row.names")
names(depth26)<-sapply(strsplit(names(depth26), "_iter"), `[`, 1)
colnames(depth26)[1]
depth26$seqdepth<-rep(colnames(depth26)[1], nrow(depth26))
depth26[,1:10]<-NULL
depth26$seqdepth<-str_remove_all(depth26$seqdepth, "[depth]")
depth26<-merge(depth26, metadata, by="row.names")
names(depth27)<-sapply(strsplit(names(depth27), "_iter"), `[`, 1)
colnames(depth27)[1]
depth27$seqdepth<-rep(colnames(depth27)[1], nrow(depth27))
depth27[,1:10]<-NULL
depth27$seqdepth<-str_remove_all(depth27$seqdepth, "[depth]")
depth27<-merge(depth27, metadata, by="row.names")
names(depth28)<-sapply(strsplit(names(depth28), "_iter"), `[`, 1)
colnames(depth28)[1]
depth28$seqdepth<-rep(colnames(depth28)[1], nrow(depth28))
depth28[,1:10]<-NULL
depth28$seqdepth<-str_remove_all(depth28$seqdepth, "[depth]")
depth28<-merge(depth28, metadata, by="row.names")
names(depth29)<-sapply(strsplit(names(depth29), "_iter"), `[`, 1)
colnames(depth29)[1]
depth29$seqdepth<-rep(colnames(depth29)[1], nrow(depth29))
depth29[,1:10]<-NULL
depth29$seqdepth<-str_remove_all(depth29$seqdepth, "[depth]")
depth29<-merge(depth29, metadata, by="row.names")
names(depth30)<-sapply(strsplit(names(depth30), "_iter"), `[`, 1)
colnames(depth30)[1]
depth30$seqdepth<-rep(colnames(depth30)[1], nrow(depth30))
depth30[,1:10]<-NULL
depth30$seqdepth<-str_remove_all(depth30$seqdepth, "[depth]")
depth30<-merge(depth30, metadata, by="row.names")
names(depth31)<-sapply(strsplit(names(depth31), "_iter"), `[`, 1)
colnames(depth31)[1]
depth31$seqdepth<-rep(colnames(depth31)[1], nrow(depth31))
depth31[,1:10]<-NULL
depth31$seqdepth<-str_remove_all(depth31$seqdepth, "[depth]")
depth31<-merge(depth31, metadata, by="row.names")
names(depth32)<-sapply(strsplit(names(depth32), "_iter"), `[`, 1)
colnames(depth32)[1]
depth32$seqdepth<-rep(colnames(depth32)[1], nrow(depth32))
depth32[,1:10]<-NULL
depth32$seqdepth<-str_remove_all(depth32$seqdepth, "[depth]")
depth32<-merge(depth32, metadata, by="row.names")
names(depth33)<-sapply(strsplit(names(depth33), "_iter"), `[`, 1)
colnames(depth33)[1]
depth33$seqdepth<-rep(colnames(depth33)[1], nrow(depth33))
depth33[,1:10]<-NULL
depth33$seqdepth<-str_remove_all(depth33$seqdepth, "[depth]")

```

```

depth33<-merge(depth33, metadata, by="row.names")
names(depth34)<-sapply(strsplit(names(depth34), "_iter"), `[`, 1)
colnames(depth34)[1]
depth34$seqdepth<-rep(colnames(depth34)[1], nrow(depth34))
depth34[,1:10]<-NULL
depth34$seqdepth<-str_remove_all(depth34$seqdepth, "[depth]")
depth34<-merge(depth34, metadata, by="row.names")
names(depth35)<-sapply(strsplit(names(depth35), "_iter"), `[`, 1)
colnames(depth35)[1]
depth35$seqdepth<-rep(colnames(depth35)[1], nrow(depth35))
depth35[,1:10]<-NULL
depth35$seqdepth<-str_remove_all(depth35$seqdepth, "[depth]")
depth35<-merge(depth35, metadata, by="row.names")
names(depth36)<-sapply(strsplit(names(depth36), "_iter"), `[`, 1)
colnames(depth36)[1]
depth36$seqdepth<-rep(colnames(depth36)[1], nrow(depth36))
depth36[,1:10]<-NULL
depth36$seqdepth<-str_remove_all(depth36$seqdepth, "[depth]")
depth36<-merge(depth36, metadata, by="row.names")
names(depth37)<-sapply(strsplit(names(depth37), "_iter"), `[`, 1)
colnames(depth37)[1]
depth37$seqdepth<-rep(colnames(depth37)[1], nrow(depth37))
depth37[,1:10]<-NULL
depth37$seqdepth<-str_remove_all(depth37$seqdepth, "[depth]")
depth37<-merge(depth37, metadata, by="row.names")
names(depth38)<-sapply(strsplit(names(depth38), "_iter"), `[`, 1)
colnames(depth38)[1]
depth38$seqdepth<-rep(colnames(depth38)[1], nrow(depth38))
depth38[,1:10]<-NULL
depth38$seqdepth<-str_remove_all(depth38$seqdepth, "[depth]")
depth38<-merge(depth38, metadata, by="row.names")
names(depth39)<-sapply(strsplit(names(depth39), "_iter"), `[`, 1)
colnames(depth39)[1]
depth39$seqdepth<-rep(colnames(depth39)[1], nrow(depth39))
depth39[,1:10]<-NULL
depth39$seqdepth<-str_remove_all(depth39$seqdepth, "[depth]")
depth39<-merge(depth39, metadata, by="row.names")
names(depth40)<-sapply(strsplit(names(depth40), "_iter"), `[`, 1)
colnames(depth40)[1]
depth40$seqdepth<-rep(colnames(depth40)[1], nrow(depth40))
depth40[,1:10]<-NULL
depth40$seqdepth<-str_remove_all(depth40$seqdepth, "[depth]")
depth40<-merge(depth40, metadata, by="row.names")
names(depth41)<-sapply(strsplit(names(depth41), "_iter"), `[`, 1)
colnames(depth41)[1]
depth41$seqdepth<-rep(colnames(depth41)[1], nrow(depth41))
depth41[,1:10]<-NULL
depth41$seqdepth<-str_remove_all(depth41$seqdepth, "[depth]")
depth41<-merge(depth41, metadata, by="row.names")
names(depth42)<-sapply(strsplit(names(depth42), "_iter"), `[`, 1)
colnames(depth42)[1]
depth42$seqdepth<-rep(colnames(depth42)[1], nrow(depth42))
depth42[,1:10]<-NULL
depth42$seqdepth<-str_remove_all(depth42$seqdepth, "[depth]")
depth42<-merge(depth42, metadata, by="row.names")
names(depth43)<-sapply(strsplit(names(depth43), "_iter"), `[`, 1)
colnames(depth43)[1]
depth43$seqdepth<-rep(colnames(depth43)[1], nrow(depth43))
depth43[,1:10]<-NULL
depth43$seqdepth<-str_remove_all(depth43$seqdepth, "[depth]")
depth43<-merge(depth43, metadata, by="row.names")
names(depth44)<-sapply(strsplit(names(depth44), "_iter"), `[`, 1)
colnames(depth44)[1]
depth44$seqdepth<-rep(colnames(depth44)[1], nrow(depth44))
depth44[,1:10]<-NULL
depth44$seqdepth<-str_remove_all(depth44$seqdepth, "[depth]")

```



```

depth44<-merge(depth44, metadata, by="row.names")
names(depth45)<-sapply(strsplit(names(depth45), "_iter"), `[`, 1)
colnames(depth45)[1]
depth45$seqdepth<-rep(colnames(depth45)[1], nrow(depth45))
depth45[,1:10]<-NULL
depth45$seqdepth<-str_remove_all(depth45$seqdepth, "[depth]")
depth45<-merge(depth45, metadata, by="row.names")
names(depth46)<-sapply(strsplit(names(depth46), "_iter"), `[`, 1)
colnames(depth46)[1]
depth46$seqdepth<-rep(colnames(depth46)[1], nrow(depth46))
depth46[,1:10]<-NULL
depth46$seqdepth<-str_remove_all(depth46$seqdepth, "[depth]")
depth46<-merge(depth46, metadata, by="row.names")
names(depth47)<-sapply(strsplit(names(depth47), "_iter"), `[`, 1)
colnames(depth47)[1]
depth47$seqdepth<-rep(colnames(depth47)[1], nrow(depth47))
depth47[,1:10]<-NULL
depth47$seqdepth<-str_remove_all(depth47$seqdepth, "[depth]")
depth47<-merge(depth47, metadata, by="row.names")
names(depth48)<-sapply(strsplit(names(depth48), "_iter"), `[`, 1)
colnames(depth48)[1]
depth48$seqdepth<-rep(colnames(depth48)[1], nrow(depth48))
depth48[,1:10]<-NULL
depth48$seqdepth<-str_remove_all(depth48$seqdepth, "[depth]")
depth48<-merge(depth48, metadata, by="row.names")
names(depth49)<-sapply(strsplit(names(depth49), "_iter"), `[`, 1)
colnames(depth49)[1]
depth49$seqdepth<-rep(colnames(depth49)[1], nrow(depth49))
depth49[,1:10]<-NULL
depth49$seqdepth<-str_remove_all(depth49$seqdepth, "[depth]")
depth49<-merge(depth49, metadata, by="row.names")
names(depth50)<-sapply(strsplit(names(depth50), "_iter"), `[`, 1)
colnames(depth50)[1]
depth50$seqdepth<-rep(colnames(depth50)[1], nrow(depth50))
depth50[,1:10]<-NULL
depth50$seqdepth<-str_remove_all(depth50$seqdepth, "[depth]")
depth50<-merge(depth50, metadata, by="row.names")
names(depth51)<-sapply(strsplit(names(depth51), "_iter"), `[`, 1)
colnames(depth51)[1]
depth51$seqdepth<-rep(colnames(depth51)[1], nrow(depth51))
depth51[,1:10]<-NULL
depth51$seqdepth<-str_remove_all(depth51$seqdepth, "[depth]")
depth51<-merge(depth51, metadata, by="row.names")
names(depth52)<-sapply(strsplit(names(depth52), "_iter"), `[`, 1)
colnames(depth52)[1]
depth52$seqdepth<-rep(colnames(depth52)[1], nrow(depth52))
depth52[,1:10]<-NULL
depth52$seqdepth<-str_remove_all(depth52$seqdepth, "[depth]")
depth52<-merge(depth52, metadata, by="row.names")
names(depth53)<-sapply(strsplit(names(depth53), "_iter"), `[`, 1)
colnames(depth53)[1]
depth53$seqdepth<-rep(colnames(depth53)[1], nrow(depth53))
depth53[,1:10]<-NULL
depth53$seqdepth<-str_remove_all(depth53$seqdepth, "[depth]")
depth53<-merge(depth53, metadata, by="row.names")
names(depth54)<-sapply(strsplit(names(depth54), "_iter"), `[`, 1)
colnames(depth54)[1]
depth54$seqdepth<-rep(colnames(depth54)[1], nrow(depth54))
depth54[,1:10]<-NULL
depth54$seqdepth<-str_remove_all(depth54$seqdepth, "[depth]")
depth54<-merge(depth54, metadata, by="row.names")
names(depth55)<-sapply(strsplit(names(depth55), "_iter"), `[`, 1)
colnames(depth55)[1]
depth55$seqdepth<-rep(colnames(depth55)[1], nrow(depth55))
depth55[,1:10]<-NULL
depth55$seqdepth<-str_remove_all(depth55$seqdepth, "[depth]")

```

```

depth55<-merge(depth55, metadata, by="row.names")
names(depth56)<-sapply(strsplit(names(depth56), "_iter"), `[`, 1)
colnames(depth56)[1]
depth56$seqdepth<-rep(colnames(depth56)[1], nrow(depth56))
depth56[,1:10]<-NULL
depth56$seqdepth<-str_remove_all(depth56$seqdepth, "[depth]")
depth56<-merge(depth56, metadata, by="row.names")
names(depth57)<-sapply(strsplit(names(depth57), "_iter"), `[`, 1)
colnames(depth57)[1]
depth57$seqdepth<-rep(colnames(depth57)[1], nrow(depth57))
depth57[,1:10]<-NULL
depth57$seqdepth<-str_remove_all(depth57$seqdepth, "[depth]")
depth57<-merge(depth57, metadata, by="row.names")
names(depth58)<-sapply(strsplit(names(depth58), "_iter"), `[`, 1)
colnames(depth58)[1]
depth58$seqdepth<-rep(colnames(depth58)[1], nrow(depth58))
depth58[,1:10]<-NULL
depth58$seqdepth<-str_remove_all(depth58$seqdepth, "[depth]")
depth58<-merge(depth58, metadata, by="row.names")
names(depth59)<-sapply(strsplit(names(depth59), "_iter"), `[`, 1)
colnames(depth59)[1]
depth59$seqdepth<-rep(colnames(depth59)[1], nrow(depth59))
depth59[,1:10]<-NULL
depth59$seqdepth<-str_remove_all(depth59$seqdepth, "[depth]")
depth59<-merge(depth59, metadata, by="row.names")
names(depth60)<-sapply(strsplit(names(depth60), "_iter"), `[`, 1)
colnames(depth60)[1]
depth60$seqdepth<-rep(colnames(depth60)[1], nrow(depth60))
depth60[,1:10]<-NULL
depth60$seqdepth<-str_remove_all(depth60$seqdepth, "[depth]")
depth60<-merge(depth60, metadata, by="row.names")
names(depth61)<-sapply(strsplit(names(depth61), "_iter"), `[`, 1)
colnames(depth61)[1]
depth61$seqdepth<-rep(colnames(depth61)[1], nrow(depth61))
depth61[,1:10]<-NULL
depth61$seqdepth<-str_remove_all(depth61$seqdepth, "[depth]")
depth61<-merge(depth61, metadata, by="row.names")
names(depth62)<-sapply(strsplit(names(depth62), "_iter"), `[`, 1)
colnames(depth62)[1]
depth62$seqdepth<-rep(colnames(depth62)[1], nrow(depth62))
depth62[,1:10]<-NULL
depth62$seqdepth<-str_remove_all(depth62$seqdepth, "[depth]")
depth62<-merge(depth62, metadata, by="row.names")
names(depth63)<-sapply(strsplit(names(depth63), "_iter"), `[`, 1)
colnames(depth63)[1]
depth63$seqdepth<-rep(colnames(depth63)[1], nrow(depth63))
depth63[,1:10]<-NULL
depth63$seqdepth<-str_remove_all(depth63$seqdepth, "[depth]")
depth63<-merge(depth63, metadata, by="row.names")
names(depth64)<-sapply(strsplit(names(depth64), "_iter"), `[`, 1)
colnames(depth64)[1]
depth64$seqdepth<-rep(colnames(depth64)[1], nrow(depth64))
depth64[,1:10]<-NULL
depth64$seqdepth<-str_remove_all(depth64$seqdepth, "[depth]")
depth64<-merge(depth64, metadata, by="row.names")
names(depth65)<-sapply(strsplit(names(depth65), "_iter"), `[`, 1)
colnames(depth65)[1]
depth65$seqdepth<-rep(colnames(depth65)[1], nrow(depth65))
depth65[,1:10]<-NULL
depth65$seqdepth<-str_remove_all(depth65$seqdepth, "[depth]")
depth65<-merge(depth65, metadata, by="row.names")
names(depth66)<-sapply(strsplit(names(depth66), "_iter"), `[`, 1)
colnames(depth66)[1]
depth66$seqdepth<-rep(colnames(depth66)[1], nrow(depth66))
depth66[,1:10]<-NULL
depth66$seqdepth<-str_remove_all(depth66$seqdepth, "[depth]")

```

```

depth66<-merge(depth66, metadata, by="row.names")
names(depth67)<-sapply(strsplit(names(depth67), "_iter"), `[`, 1)
colnames(depth67)[1]
depth67$seqdepth<-rep(colnames(depth67)[1], nrow(depth67))
depth67[,1:10]<-NULL
depth67$seqdepth<-str_remove_all(depth67$seqdepth, "[depth]")
depth67<-merge(depth67, metadata, by="row.names")
names(depth68)<-sapply(strsplit(names(depth68), "_iter"), `[`, 1)
colnames(depth68)[1]
depth68$seqdepth<-rep(colnames(depth68)[1], nrow(depth68))
depth68[,1:10]<-NULL
depth68$seqdepth<-str_remove_all(depth68$seqdepth, "[depth]")
depth68<-merge(depth68, metadata, by="row.names")
names(depth69)<-sapply(strsplit(names(depth69), "_iter"), `[`, 1)
colnames(depth69)[1]
depth69$seqdepth<-rep(colnames(depth69)[1], nrow(depth69))
depth69[,1:10]<-NULL
depth69$seqdepth<-str_remove_all(depth69$seqdepth, "[depth]")
depth69<-merge(depth69, metadata, by="row.names")
names(depth70)<-sapply(strsplit(names(depth70), "_iter"), `[`, 1)
colnames(depth70)[1]
depth70$seqdepth<-rep(colnames(depth70)[1], nrow(depth70))
depth70[,1:10]<-NULL
depth70$seqdepth<-str_remove_all(depth70$seqdepth, "[depth]")
depth70<-merge(depth70, metadata, by="row.names")
names(depth71)<-sapply(strsplit(names(depth71), "_iter"), `[`, 1)
colnames(depth71)[1]
depth71$seqdepth<-rep(colnames(depth71)[1], nrow(depth71))
depth71[,1:10]<-NULL
depth71$seqdepth<-str_remove_all(depth71$seqdepth, "[depth]")
depth71<-merge(depth71, metadata, by="row.names")
names(depth72)<-sapply(strsplit(names(depth72), "_iter"), `[`, 1)
colnames(depth72)[1]
depth72$seqdepth<-rep(colnames(depth72)[1], nrow(depth72))
depth72[,1:10]<-NULL
depth72$seqdepth<-str_remove_all(depth72$seqdepth, "[depth]")
depth72<-merge(depth72, metadata, by="row.names")
names(depth73)<-sapply(strsplit(names(depth73), "_iter"), `[`, 1)
colnames(depth73)[1]
depth73$seqdepth<-rep(colnames(depth73)[1], nrow(depth73))
depth73[,1:10]<-NULL
depth73$seqdepth<-str_remove_all(depth73$seqdepth, "[depth]")
depth73<-merge(depth73, metadata, by="row.names")
names(depth74)<-sapply(strsplit(names(depth74), "_iter"), `[`, 1)
colnames(depth74)[1]
depth74$seqdepth<-rep(colnames(depth74)[1], nrow(depth74))
depth74[,1:10]<-NULL
depth74$seqdepth<-str_remove_all(depth74$seqdepth, "[depth]")
depth74<-merge(depth74, metadata, by="row.names")
names(depth75)<-sapply(strsplit(names(depth75), "_iter"), `[`, 1)
colnames(depth75)[1]
depth75$seqdepth<-rep(colnames(depth75)[1], nrow(depth75))
depth75[,1:10]<-NULL
depth75$seqdepth<-str_remove_all(depth75$seqdepth, "[depth]")
depth75<-merge(depth75, metadata, by="row.names")
names(depth76)<-sapply(strsplit(names(depth76), "_iter"), `[`, 1)
colnames(depth76)[1]
depth76$seqdepth<-rep(colnames(depth76)[1], nrow(depth76))
depth76[,1:10]<-NULL
depth76$seqdepth<-str_remove_all(depth76$seqdepth, "[depth]")
depth76<-merge(depth76, metadata, by="row.names")
names(depth77)<-sapply(strsplit(names(depth77), "_iter"), `[`, 1)
colnames(depth77)[1]
depth77$seqdepth<-rep(colnames(depth77)[1], nrow(depth77))
depth77[,1:10]<-NULL
depth77$seqdepth<-str_remove_all(depth77$seqdepth, "[depth]")

```

```

depth77<-merge(depth77, metadata, by="row.names")
names(depth78)<-sapply(strsplit(names(depth78), "_iter"), `[`, 1)
colnames(depth78)[1]
depth78$seqdepth<-rep(colnames(depth78)[1], nrow(depth78))
depth78[,1:10]<-NULL
depth78$seqdepth<-str_remove_all(depth78$seqdepth, "[depth]")
depth78<-merge(depth78, metadata, by="row.names")
names(depth79)<-sapply(strsplit(names(depth79), "_iter"), `[`, 1)
colnames(depth79)[1]
depth79$seqdepth<-rep(colnames(depth79)[1], nrow(depth79))
depth79[,1:10]<-NULL
depth79$seqdepth<-str_remove_all(depth79$seqdepth, "[depth]")
depth79<-merge(depth79, metadata, by="row.names")
names(depth80)<-sapply(strsplit(names(depth80), "_iter"), `[`, 1)
colnames(depth80)[1]
depth80$seqdepth<-rep(colnames(depth80)[1], nrow(depth80))
depth80[,1:10]<-NULL
depth80$seqdepth<-str_remove_all(depth80$seqdepth, "[depth]")
depth80<-merge(depth80, metadata, by="row.names")
names(depth81)<-sapply(strsplit(names(depth81), "_iter"), `[`, 1)
colnames(depth81)[1]
depth81$seqdepth<-rep(colnames(depth81)[1], nrow(depth81))
depth81[,1:10]<-NULL
depth81$seqdepth<-str_remove_all(depth81$seqdepth, "[depth]")
depth81<-merge(depth81, metadata, by="row.names")
names(depth82)<-sapply(strsplit(names(depth82), "_iter"), `[`, 1)
colnames(depth82)[1]
depth82$seqdepth<-rep(colnames(depth82)[1], nrow(depth82))
depth82[,1:10]<-NULL
depth82$seqdepth<-str_remove_all(depth82$seqdepth, "[depth]")
depth82<-merge(depth82, metadata, by="row.names")
names(depth83)<-sapply(strsplit(names(depth83), "_iter"), `[`, 1)
colnames(depth83)[1]
depth83$seqdepth<-rep(colnames(depth83)[1], nrow(depth83))
depth83[,1:10]<-NULL
depth83$seqdepth<-str_remove_all(depth83$seqdepth, "[depth]")
depth83<-merge(depth83, metadata, by="row.names")
names(depth84)<-sapply(strsplit(names(depth84), "_iter"), `[`, 1)
colnames(depth84)[1]
depth84$seqdepth<-rep(colnames(depth84)[1], nrow(depth84))
depth84[,1:10]<-NULL
depth84$seqdepth<-str_remove_all(depth84$seqdepth, "[depth]")
depth84<-merge(depth84, metadata, by="row.names")
names(depth85)<-sapply(strsplit(names(depth85), "_iter"), `[`, 1)
colnames(depth85)[1]
depth85$seqdepth<-rep(colnames(depth85)[1], nrow(depth85))
depth85[,1:10]<-NULL
depth85$seqdepth<-str_remove_all(depth85$seqdepth, "[depth]")
depth85<-merge(depth85, metadata, by="row.names")
names(depth86)<-sapply(strsplit(names(depth86), "_iter"), `[`, 1)
colnames(depth86)[1]
depth86$seqdepth<-rep(colnames(depth86)[1], nrow(depth86))
depth86[,1:10]<-NULL
depth86$seqdepth<-str_remove_all(depth86$seqdepth, "[depth]")
depth86<-merge(depth86, metadata, by="row.names")
names(depth87)<-sapply(strsplit(names(depth87), "_iter"), `[`, 1)
colnames(depth87)[1]
depth87$seqdepth<-rep(colnames(depth87)[1], nrow(depth87))
depth87[,1:10]<-NULL
depth87$seqdepth<-str_remove_all(depth87$seqdepth, "[depth]")
depth87<-merge(depth87, metadata, by="row.names")
names(depth88)<-sapply(strsplit(names(depth88), "_iter"), `[`, 1)
colnames(depth88)[1]
depth88$seqdepth<-rep(colnames(depth88)[1], nrow(depth88))
depth88[,1:10]<-NULL
depth88$seqdepth<-str_remove_all(depth88$seqdepth, "[depth]")

```

```

depth88<-merge(depth88, metadata, by="row.names")
names(depth89)<-sapply(strsplit(names(depth89), "_iter"), `[`, 1)
colnames(depth89)[1]
depth89$seqdepth<-rep(colnames(depth89)[1], nrow(depth89))
depth89[,1:10]<-NULL
depth89$seqdepth<-str_remove_all(depth89$seqdepth, "[depth]")
depth89<-merge(depth89, metadata, by="row.names")
names(depth90)<-sapply(strsplit(names(depth90), "_iter"), `[`, 1)
colnames(depth90)[1]
depth90$seqdepth<-rep(colnames(depth90)[1], nrow(depth90))
depth90[,1:10]<-NULL
depth90$seqdepth<-str_remove_all(depth90$seqdepth, "[depth]")
depth90<-merge(depth90, metadata, by="row.names")
names(depth91)<-sapply(strsplit(names(depth91), "_iter"), `[`, 1)
colnames(depth91)[1]
depth91$seqdepth<-rep(colnames(depth91)[1], nrow(depth91))
depth91[,1:10]<-NULL
depth91$seqdepth<-str_remove_all(depth91$seqdepth, "[depth]")
depth91<-merge(depth91, metadata, by="row.names")
names(depth92)<-sapply(strsplit(names(depth92), "_iter"), `[`, 1)
colnames(depth92)[1]
depth92$seqdepth<-rep(colnames(depth92)[1], nrow(depth92))
depth92[,1:10]<-NULL
depth92$seqdepth<-str_remove_all(depth92$seqdepth, "[depth]")
depth92<-merge(depth92, metadata, by="row.names")
names(depth93)<-sapply(strsplit(names(depth93), "_iter"), `[`, 1)
colnames(depth93)[1]
depth93$seqdepth<-rep(colnames(depth93)[1], nrow(depth93))
depth93[,1:10]<-NULL
depth93$seqdepth<-str_remove_all(depth93$seqdepth, "[depth]")
depth93<-merge(depth93, metadata, by="row.names")
names(depth94)<-sapply(strsplit(names(depth94), "_iter"), `[`, 1)
colnames(depth94)[1]
depth94$seqdepth<-rep(colnames(depth94)[1], nrow(depth94))
depth94[,1:10]<-NULL
depth94$seqdepth<-str_remove_all(depth94$seqdepth, "[depth]")
depth94<-merge(depth94, metadata, by="row.names")
names(depth95)<-sapply(strsplit(names(depth95), "_iter"), `[`, 1)
colnames(depth95)[1]
depth95$seqdepth<-rep(colnames(depth95)[1], nrow(depth95))
depth95[,1:10]<-NULL
depth95$seqdepth<-str_remove_all(depth95$seqdepth, "[depth]")
depth95<-merge(depth95, metadata, by="row.names")
names(depth96)<-sapply(strsplit(names(depth96), "_iter"), `[`, 1)
colnames(depth96)[1]
depth96$seqdepth<-rep(colnames(depth96)[1], nrow(depth96))
depth96[,1:10]<-NULL
depth96$seqdepth<-str_remove_all(depth96$seqdepth, "[depth]")
depth96<-merge(depth96, metadata, by="row.names")
names(depth97)<-sapply(strsplit(names(depth97), "_iter"), `[`, 1)
colnames(depth97)[1]
depth97$seqdepth<-rep(colnames(depth97)[1], nrow(depth97))
depth97[,1:10]<-NULL
depth97$seqdepth<-str_remove_all(depth97$seqdepth, "[depth]")
depth97<-merge(depth97, metadata, by="row.names")
names(depth98)<-sapply(strsplit(names(depth98), "_iter"), `[`, 1)
colnames(depth98)[1]
depth98$seqdepth<-rep(colnames(depth98)[1], nrow(depth98))
depth98[,1:10]<-NULL
depth98$seqdepth<-str_remove_all(depth98$seqdepth, "[depth]")
depth98<-merge(depth98, metadata, by="row.names")
names(depth99)<-sapply(strsplit(names(depth99), "_iter"), `[`, 1)
colnames(depth99)[1]
depth99$seqdepth<-rep(colnames(depth99)[1], nrow(depth99))
depth99[,1:10]<-NULL
depth99$seqdepth<-str_remove_all(depth99$seqdepth, "[depth]")

```

```

depth99<-merge(depth99, metadata, by="row.names")
names(depth100)<-sapply(strsplit(names(depth100), "_iter"), `[`, 1)
colnames(depth100)[1]
depth100$seqdepth<-rep(colnames(depth100)[1], nrow(depth100))
depth100[,1:10]<-NULL
depth100$seqdepth<-str_remove_all(depth100$seqdepth, "[depth]")
depth100<-merge(depth100, metadata, by="row.names")
d1<-rbind(depth1,depth2)
d2<-rbind(d1,depth3)
d3<-rbind(d2,depth4)
d4<-rbind(d3,depth5)
d5<-rbind(d4,depth6)
d6<-rbind(d5,depth7)
d7<-rbind(d6,depth8)
d8<-rbind(d7,depth9)
d9<-rbind(d8,depth10)
d10<-rbind(d9,depth11)
d11<-rbind(d10,depth12)
d12<-rbind(d11,depth13)
d13<-rbind(d12,depth14)
d14<-rbind(d13,depth15)
d15<-rbind(d14,depth16)
d16<-rbind(d15,depth17)
d17<-rbind(d16,depth18)
d18<-rbind(d17,depth19)
d19<-rbind(d18,depth20)
d20<-rbind(d19,depth21)
d21<-rbind(d20,depth22)
d22<-rbind(d21,depth23)
d23<-rbind(d22,depth24)
d24<-rbind(d23,depth25)
d25<-rbind(d24,depth26)
d26<-rbind(d25,depth27)
d27<-rbind(d26,depth28)
d28<-rbind(d27,depth29)
d29<-rbind(d28,depth30)
d30<-rbind(d29,depth31)
d31<-rbind(d30,depth32)
d32<-rbind(d31,depth33)
d33<-rbind(d32,depth34)
d34<-rbind(d33,depth35)
d35<-rbind(d34,depth36)
d36<-rbind(d35,depth37)
d37<-rbind(d36,depth38)
d38<-rbind(d37,depth39)
d39<-rbind(d38,depth40)
d40<-rbind(d39,depth41)
d41<-rbind(d40,depth42)
d42<-rbind(d41,depth43)
d43<-rbind(d42,depth44)
d44<-rbind(d43,depth45)
d45<-rbind(d44,depth46)
d46<-rbind(d45,depth47)
d47<-rbind(d46,depth48)
d48<-rbind(d47,depth49)
d49<-rbind(d48,depth50)
d50<-rbind(d49,depth51)
d51<-rbind(d50,depth52)
d52<-rbind(d51,depth53)
d53<-rbind(d52,depth54)
d54<-rbind(d53,depth55)
d55<-rbind(d54,depth56)
d56<-rbind(d55,depth57)
d57<-rbind(d56,depth58)
d58<-rbind(d57,depth59)
d59<-rbind(d58,depth60)

```

```

d60<-rbind(d59,depth61)
d61<-rbind(d60,depth62)
d62<-rbind(d61,depth63)
d63<-rbind(d62,depth64)
d64<-rbind(d63,depth65)
d65<-rbind(d64,depth66)
d66<-rbind(d65,depth67)
d67<-rbind(d66,depth68)
d68<-rbind(d67,depth69)
d69<-rbind(d68,depth70)
d70<-rbind(d69,depth71)
d71<-rbind(d70,depth72)
d72<-rbind(d71,depth73)
d73<-rbind(d72,depth74)
d74<-rbind(d73,depth75)
d75<-rbind(d74,depth76)
d76<-rbind(d75,depth77)
d77<-rbind(d76,depth78)
d78<-rbind(d77,depth79)
d79<-rbind(d78,depth80)
d80<-rbind(d79,depth81)
d81<-rbind(d80,depth82)
d82<-rbind(d81,depth83)
d83<-rbind(d82,depth84)
d84<-rbind(d83,depth85)
d85<-rbind(d84,depth86)
d86<-rbind(d85,depth87)
d87<-rbind(d86,depth88)
d88<-rbind(d87,depth89)
d89<-rbind(d88,depth90)
d90<-rbind(d89,depth91)
d91<-rbind(d90,depth92)
d92<-rbind(d91,depth93)
d93<-rbind(d92,depth94)
d94<-rbind(d93,depth95)
d95<-rbind(d94,depth96)
d96<-rbind(d95,depth97)
d97<-rbind(d96,depth98)
d98<-rbind(d97,depth99)
d99<-rbind(d98,depth100)
names(d99)[1] <- "SampleID"
names(d99)[names(d99) == "mean"] <- "observedASVs"
d99$seqdepth<-as.numeric(d99$seqdepth)
write.table(d99, file = "raref_sub.txt",sep="\t",col.names= T, row.names=F)
quit("no")

sed 's//g' raref_sub.txt > raref_sub2.txt
cd $PWDa/dada2/core-metrics-results/
unzip alpha-rarefaction_complete.qzv
cd "$(\ls -ldt ./ */ | head -n 1)/data"
sed '1 s//g' observed_otus.csv > observed_otus2.csv
sed -e 's/,/\t/g' observed_otus2.csv > tmpfile; mv tmpfile observed_otus.txt
cp -r "observed_otus.txt" $PWDa/dada2/
cd $PWDa/dada2/
awk 'BEGIN { FS = OFS = "\t" } { for(i=1; i<=NF; i++) if($i ~ /^ *$/) $i = 0 }; 1' ob-
served_otus.txt

```

```

R
list.of.packages <- c("ggplot2", "stringr")
new.packages <- list.of.packages[!(list.of.packages %in% installed.packages()),"Package"]]
if(length(new.packages)) install.packages(new.packages, repos='http://cran.us.r-project.org')
library(stringr)
library(ggplot2)
data<- read.table(file="observed_otusN.txt", header =T)

```

```

data[data == 0] <- NA
dat <- data[,-1]
rownames(dat) <- data[,1]
depth1<-dat[,1:10]
depth2<-dat[,11:20]
depth3<-dat[,21:30]
depth4<-dat[,31:40]
depth5<-dat[,41:50]
depth6<-dat[,51:60]
depth7<-dat[,61:70]
depth8<-dat[,71:80]
depth9<-dat[,81:90]
depth10<-dat[,91:100]
depth11<-dat[,101:110]
depth12<-dat[,111:120]
depth13<-dat[,121:130]
depth14<-dat[,131:140]
depth15<-dat[,141:150]
depth16<-dat[,151:160]
depth17<-dat[,161:170]
depth18<-dat[,171:180]
depth19<-dat[,181:190]
depth20<-dat[,191:200]
depth21<-dat[,201:210]
depth22<-dat[,211:220]
depth23<-dat[,221:230]
depth24<-dat[,231:240]
depth25<-dat[,241:250]
depth26<-dat[,251:260]
depth27<-dat[,261:270]
depth28<-dat[,271:280]
depth29<-dat[,281:290]
depth30<-dat[,291:300]
depth31<-dat[,301:310]
depth32<-dat[,311:320]
depth33<-dat[,321:330]
depth34<-dat[,331:340]
depth35<-dat[,341:350]
depth36<-dat[,351:360]
depth37<-dat[,361:370]
depth38<-dat[,371:380]
depth39<-dat[,381:390]
depth40<-dat[,391:400]
depth41<-dat[,401:410]
depth42<-dat[,411:420]
depth43<-dat[,421:430]
depth44<-dat[,431:440]
depth45<-dat[,441:450]
depth46<-dat[,451:460]
depth47<-dat[,461:470]
depth48<-dat[,471:480]
depth49<-dat[,481:490]
depth50<-dat[,491:500]
depth51<-dat[,501:510]
depth52<-dat[,511:520]
depth53<-dat[,521:530]
depth54<-dat[,531:540]
depth55<-dat[,541:550]
depth56<-dat[,551:560]
depth57<-dat[,561:570]
depth58<-dat[,571:580]
depth59<-dat[,581:590]
depth60<-dat[,591:600]
depth61<-dat[,601:610]
depth62<-dat[,611:620]
depth63<-dat[,621:630]
depth64<-dat[,631:640]

```



```

depth65<-dat[,641:650]
depth66<-dat[,651:660]
depth67<-dat[,661:670]
depth68<-dat[,671:680]
depth69<-dat[,681:690]
depth70<-dat[,691:700]
depth71<-dat[,701:710]
depth72<-dat[,711:720]
depth73<-dat[,721:730]
depth74<-dat[,731:740]
depth75<-dat[,741:750]
depth76<-dat[,751:760]
depth77<-dat[,761:770]
depth78<-dat[,771:780]
depth79<-dat[,781:790]
depth80<-dat[,791:800]
depth81<-dat[,801:810]
depth82<-dat[,811:820]
depth83<-dat[,821:830]
depth84<-dat[,831:840]
depth85<-dat[,841:850]
depth86<-dat[,851:860]
depth87<-dat[,861:870]
depth88<-dat[,871:880]
depth89<-dat[,881:890]
depth90<-dat[,891:900]
depth91<-dat[,901:910]
depth92<-dat[,911:920]
depth93<-dat[,921:930]
depth94<-dat[,931:940]
depth95<-dat[,941:950]
depth96<-dat[,951:960]
depth97<-dat[,961:970]
depth98<-dat[,971:980]
depth99<-dat[,981:990]
depth100<-dat[,991:1000]
metadata<-dat[,1001:ncol(dat)]
depth1$mean<-rowMeans(depth1)
depth2$mean<-rowMeans(depth2)
depth3$mean<-rowMeans(depth3)
depth4$mean<-rowMeans(depth4)
depth5$mean<-rowMeans(depth5)
depth6$mean<-rowMeans(depth6)
depth7$mean<-rowMeans(depth7)
depth8$mean<-rowMeans(depth8)
depth9$mean<-rowMeans(depth9)
depth10$mean<-rowMeans(depth10)
depth11$mean<-rowMeans(depth11)
depth12$mean<-rowMeans(depth12)
depth13$mean<-rowMeans(depth13)
depth14$mean<-rowMeans(depth14)
depth15$mean<-rowMeans(depth15)
depth16$mean<-rowMeans(depth16)
depth17$mean<-rowMeans(depth17)
depth18$mean<-rowMeans(depth18)
depth19$mean<-rowMeans(depth19)
depth20$mean<-rowMeans(depth20)
depth21$mean<-rowMeans(depth21)
depth22$mean<-rowMeans(depth22)
depth23$mean<-rowMeans(depth23)
depth24$mean<-rowMeans(depth24)
depth25$mean<-rowMeans(depth25)
depth26$mean<-rowMeans(depth26)
depth27$mean<-rowMeans(depth27)
depth28$mean<-rowMeans(depth28)
depth29$mean<-rowMeans(depth29)

```

```
depth30$mean<-rowMeans(depth30)
depth31$mean<-rowMeans(depth31)
depth32$mean<-rowMeans(depth32)
depth33$mean<-rowMeans(depth33)
depth34$mean<-rowMeans(depth34)
depth35$mean<-rowMeans(depth35)
depth36$mean<-rowMeans(depth36)
depth37$mean<-rowMeans(depth37)
depth38$mean<-rowMeans(depth38)
depth39$mean<-rowMeans(depth39)
depth40$mean<-rowMeans(depth40)
depth41$mean<-rowMeans(depth41)
depth42$mean<-rowMeans(depth42)
depth43$mean<-rowMeans(depth43)
depth44$mean<-rowMeans(depth44)
depth45$mean<-rowMeans(depth45)
depth46$mean<-rowMeans(depth46)
depth47$mean<-rowMeans(depth47)
depth48$mean<-rowMeans(depth48)
depth49$mean<-rowMeans(depth49)
depth50$mean<-rowMeans(depth50)
depth51$mean<-rowMeans(depth51)
depth52$mean<-rowMeans(depth52)
depth53$mean<-rowMeans(depth53)
depth54$mean<-rowMeans(depth54)
depth55$mean<-rowMeans(depth55)
depth56$mean<-rowMeans(depth56)
depth57$mean<-rowMeans(depth57)
depth58$mean<-rowMeans(depth58)
depth59$mean<-rowMeans(depth59)
depth60$mean<-rowMeans(depth60)
depth61$mean<-rowMeans(depth61)
depth62$mean<-rowMeans(depth62)
depth63$mean<-rowMeans(depth63)
depth64$mean<-rowMeans(depth64)
depth65$mean<-rowMeans(depth65)
depth66$mean<-rowMeans(depth66)
depth67$mean<-rowMeans(depth67)
depth68$mean<-rowMeans(depth68)
depth69$mean<-rowMeans(depth69)
depth70$mean<-rowMeans(depth70)
depth71$mean<-rowMeans(depth71)
depth72$mean<-rowMeans(depth72)
depth73$mean<-rowMeans(depth73)
depth74$mean<-rowMeans(depth74)
depth75$mean<-rowMeans(depth75)
depth76$mean<-rowMeans(depth76)
depth77$mean<-rowMeans(depth77)
depth78$mean<-rowMeans(depth78)
depth79$mean<-rowMeans(depth79)
depth80$mean<-rowMeans(depth80)
depth81$mean<-rowMeans(depth81)
depth82$mean<-rowMeans(depth82)
depth83$mean<-rowMeans(depth83)
depth84$mean<-rowMeans(depth84)
depth85$mean<-rowMeans(depth85)
depth86$mean<-rowMeans(depth86)
depth87$mean<-rowMeans(depth87)
depth88$mean<-rowMeans(depth88)
depth89$mean<-rowMeans(depth89)
depth90$mean<-rowMeans(depth90)
depth91$mean<-rowMeans(depth91)
depth92$mean<-rowMeans(depth92)
depth93$mean<-rowMeans(depth93)
depth94$mean<-rowMeans(depth94)
depth95$mean<-rowMeans(depth95)
```

```

depth96$mean<-rowMeans(depth96)
depth97$mean<-rowMeans(depth97)
depth98$mean<-rowMeans(depth98)
depth99$mean<-rowMeans(depth99)
depth100$mean<-rowMeans(depth100)
names(depth1)<-sapply(strsplit(names(depth1), "_iter"), `[[`, 1)
colnames(depth1)[1]
depth1$seqdepth<-rep(colnames(depth1)[1], nrow(depth1))
depth1[,1:10]<-NULL
depth1$seqdepth<-str_remove_all(depth1$seqdepth, "[depth]")
depth1<-merge(depth1, metadata, by="row.names")
names(depth2)<-sapply(strsplit(names(depth2), "_iter"), `[[`, 1)
colnames(depth2)[1]
depth2$seqdepth<-rep(colnames(depth2)[1], nrow(depth2))
depth2[,1:10]<-NULL
depth2$seqdepth<-str_remove_all(depth2$seqdepth, "[depth]")
depth2<-merge(depth2, metadata, by="row.names")
names(depth3)<-sapply(strsplit(names(depth3), "_iter"), `[[`, 1)
colnames(depth3)[1]
depth3$seqdepth<-rep(colnames(depth3)[1], nrow(depth3))
depth3[,1:10]<-NULL
depth3$seqdepth<-str_remove_all(depth3$seqdepth, "[depth]")
depth3<-merge(depth3, metadata, by="row.names")
names(depth4)<-sapply(strsplit(names(depth4), "_iter"), `[[`, 1)
colnames(depth4)[1]
depth4$seqdepth<-rep(colnames(depth4)[1], nrow(depth4))
depth4[,1:10]<-NULL
depth4$seqdepth<-str_remove_all(depth4$seqdepth, "[depth]")
depth4<-merge(depth4, metadata, by="row.names")
names(depth5)<-sapply(strsplit(names(depth5), "_iter"), `[[`, 1)
colnames(depth5)[1]
depth5$seqdepth<-rep(colnames(depth5)[1], nrow(depth5))
depth5[,1:10]<-NULL
depth5$seqdepth<-str_remove_all(depth5$seqdepth, "[depth]")
depth5<-merge(depth5, metadata, by="row.names")
names(depth6)<-sapply(strsplit(names(depth6), "_iter"), `[[`, 1)
colnames(depth6)[1]
depth6$seqdepth<-rep(colnames(depth6)[1], nrow(depth6))
depth6[,1:10]<-NULL
depth6$seqdepth<-str_remove_all(depth6$seqdepth, "[depth]")
depth6<-merge(depth6, metadata, by="row.names")
names(depth7)<-sapply(strsplit(names(depth7), "_iter"), `[[`, 1)
colnames(depth7)[1]
depth7$seqdepth<-rep(colnames(depth7)[1], nrow(depth7))
depth7[,1:10]<-NULL
depth7$seqdepth<-str_remove_all(depth7$seqdepth, "[depth]")
depth7<-merge(depth7, metadata, by="row.names")
names(depth8)<-sapply(strsplit(names(depth8), "_iter"), `[[`, 1)
colnames(depth8)[1]
depth8$seqdepth<-rep(colnames(depth8)[1], nrow(depth8))
depth8[,1:10]<-NULL
depth8$seqdepth<-str_remove_all(depth8$seqdepth, "[depth]")
depth8<-merge(depth8, metadata, by="row.names")
names(depth9)<-sapply(strsplit(names(depth9), "_iter"), `[[`, 1)
colnames(depth9)[1]
depth9$seqdepth<-rep(colnames(depth9)[1], nrow(depth9))
depth9[,1:10]<-NULL
depth9$seqdepth<-str_remove_all(depth9$seqdepth, "[depth]")
depth9<-merge(depth9, metadata, by="row.names")
names(depth10)<-sapply(strsplit(names(depth10), "_iter"), `[[`, 1)
colnames(depth10)[1]
depth10$seqdepth<-rep(colnames(depth10)[1], nrow(depth10))
depth10[,1:10]<-NULL
depth10$seqdepth<-str_remove_all(depth10$seqdepth, "[depth]")
depth10<-merge(depth10, metadata, by="row.names")
names(depth11)<-sapply(strsplit(names(depth11), "_iter"), `[[`, 1)

```

```

colnames(depth11)[1]
depth11$seqdepth<-rep(colnames(depth11)[1], nrow(depth11))
depth11[,1:10]<-NULL
depth11$seqdepth<-str_remove_all(depth11$seqdepth, "[depth]")
depth11<-merge(depth11, metadata, by="row.names")
names(depth12)<-sapply(strsplit(names(depth12), "_iter"), `[`, 1)
colnames(depth12)[1]
depth12$seqdepth<-rep(colnames(depth12)[1], nrow(depth12))
depth12[,1:10]<-NULL
depth12$seqdepth<-str_remove_all(depth12$seqdepth, "[depth]")
depth12<-merge(depth12, metadata, by="row.names")
names(depth13)<-sapply(strsplit(names(depth13), "_iter"), `[`, 1)
colnames(depth13)[1]
depth13$seqdepth<-rep(colnames(depth13)[1], nrow(depth13))
depth13[,1:10]<-NULL
depth13$seqdepth<-str_remove_all(depth13$seqdepth, "[depth]")
depth13<-merge(depth13, metadata, by="row.names")
names(depth14)<-sapply(strsplit(names(depth14), "_iter"), `[`, 1)
colnames(depth14)[1]
depth14$seqdepth<-rep(colnames(depth14)[1], nrow(depth14))
depth14[,1:10]<-NULL
depth14$seqdepth<-str_remove_all(depth14$seqdepth, "[depth]")
depth14<-merge(depth14, metadata, by="row.names")
names(depth15)<-sapply(strsplit(names(depth15), "_iter"), `[`, 1)
colnames(depth15)[1]
depth15$seqdepth<-rep(colnames(depth15)[1], nrow(depth15))
depth15[,1:10]<-NULL
depth15$seqdepth<-str_remove_all(depth15$seqdepth, "[depth]")
depth15<-merge(depth15, metadata, by="row.names")
names(depth16)<-sapply(strsplit(names(depth16), "_iter"), `[`, 1)
colnames(depth16)[1]
depth16$seqdepth<-rep(colnames(depth16)[1], nrow(depth16))
depth16[,1:10]<-NULL
depth16$seqdepth<-str_remove_all(depth16$seqdepth, "[depth]")
depth16<-merge(depth16, metadata, by="row.names")
names(depth17)<-sapply(strsplit(names(depth17), "_iter"), `[`, 1)
colnames(depth17)[1]
depth17$seqdepth<-rep(colnames(depth17)[1], nrow(depth17))
depth17[,1:10]<-NULL
depth17$seqdepth<-str_remove_all(depth17$seqdepth, "[depth]")
depth17<-merge(depth17, metadata, by="row.names")
names(depth18)<-sapply(strsplit(names(depth18), "_iter"), `[`, 1)
colnames(depth18)[1]
depth18$seqdepth<-rep(colnames(depth18)[1], nrow(depth18))
depth18[,1:10]<-NULL
depth18$seqdepth<-str_remove_all(depth18$seqdepth, "[depth]")
depth18<-merge(depth18, metadata, by="row.names")
names(depth19)<-sapply(strsplit(names(depth19), "_iter"), `[`, 1)
colnames(depth19)[1]
depth19$seqdepth<-rep(colnames(depth19)[1], nrow(depth19))
depth19[,1:10]<-NULL
depth19$seqdepth<-str_remove_all(depth19$seqdepth, "[depth]")
depth19<-merge(depth19, metadata, by="row.names")
names(depth20)<-sapply(strsplit(names(depth20), "_iter"), `[`, 1)
colnames(depth20)[1]
depth20$seqdepth<-rep(colnames(depth20)[1], nrow(depth20))
depth20[,1:10]<-NULL
depth20$seqdepth<-str_remove_all(depth20$seqdepth, "[depth]")
depth20<-merge(depth20, metadata, by="row.names")
names(depth21)<-sapply(strsplit(names(depth21), "_iter"), `[`, 1)
colnames(depth21)[1]
depth21$seqdepth<-rep(colnames(depth21)[1], nrow(depth21))
depth21[,1:10]<-NULL
depth21$seqdepth<-str_remove_all(depth21$seqdepth, "[depth]")
depth21<-merge(depth21, metadata, by="row.names")
names(depth22)<-sapply(strsplit(names(depth22), "_iter"), `[`, 1)

```

```

colnames(depth22)[1]
depth22$seqdepth<-rep(colnames(depth22)[1], nrow(depth22))
depth22[,1:10]<-NULL
depth22$seqdepth<-str_remove_all(depth22$seqdepth, "[depth]")
depth22<-merge(depth22, metadata, by="row.names")
names(depth23)<-sapply(strsplit(names(depth23), "_iter"), `[`, 1)
colnames(depth23)[1]
depth23$seqdepth<-rep(colnames(depth23)[1], nrow(depth23))
depth23[,1:10]<-NULL
depth23$seqdepth<-str_remove_all(depth23$seqdepth, "[depth]")
depth23<-merge(depth23, metadata, by="row.names")
names(depth24)<-sapply(strsplit(names(depth24), "_iter"), `[`, 1)
colnames(depth24)[1]
depth24$seqdepth<-rep(colnames(depth24)[1], nrow(depth24))
depth24[,1:10]<-NULL
depth24$seqdepth<-str_remove_all(depth24$seqdepth, "[depth]")
depth24<-merge(depth24, metadata, by="row.names")
names(depth25)<-sapply(strsplit(names(depth25), "_iter"), `[`, 1)
colnames(depth25)[1]
depth25$seqdepth<-rep(colnames(depth25)[1], nrow(depth25))
depth25[,1:10]<-NULL
depth25$seqdepth<-str_remove_all(depth25$seqdepth, "[depth]")
depth25<-merge(depth25, metadata, by="row.names")
names(depth26)<-sapply(strsplit(names(depth26), "_iter"), `[`, 1)
colnames(depth26)[1]
depth26$seqdepth<-rep(colnames(depth26)[1], nrow(depth26))
depth26[,1:10]<-NULL
depth26$seqdepth<-str_remove_all(depth26$seqdepth, "[depth]")
depth26<-merge(depth26, metadata, by="row.names")
names(depth27)<-sapply(strsplit(names(depth27), "_iter"), `[`, 1)
colnames(depth27)[1]
depth27$seqdepth<-rep(colnames(depth27)[1], nrow(depth27))
depth27[,1:10]<-NULL
depth27$seqdepth<-str_remove_all(depth27$seqdepth, "[depth]")
depth27<-merge(depth27, metadata, by="row.names")
names(depth28)<-sapply(strsplit(names(depth28), "_iter"), `[`, 1)
colnames(depth28)[1]
depth28$seqdepth<-rep(colnames(depth28)[1], nrow(depth28))
depth28[,1:10]<-NULL
depth28$seqdepth<-str_remove_all(depth28$seqdepth, "[depth]")
depth28<-merge(depth28, metadata, by="row.names")
names(depth29)<-sapply(strsplit(names(depth29), "_iter"), `[`, 1)
colnames(depth29)[1]
depth29$seqdepth<-rep(colnames(depth29)[1], nrow(depth29))
depth29[,1:10]<-NULL
depth29$seqdepth<-str_remove_all(depth29$seqdepth, "[depth]")
depth29<-merge(depth29, metadata, by="row.names")
names(depth30)<-sapply(strsplit(names(depth30), "_iter"), `[`, 1)
colnames(depth30)[1]
depth30$seqdepth<-rep(colnames(depth30)[1], nrow(depth30))
depth30[,1:10]<-NULL
depth30$seqdepth<-str_remove_all(depth30$seqdepth, "[depth]")
depth30<-merge(depth30, metadata, by="row.names")
names(depth31)<-sapply(strsplit(names(depth31), "_iter"), `[`, 1)
colnames(depth31)[1]
depth31$seqdepth<-rep(colnames(depth31)[1], nrow(depth31))
depth31[,1:10]<-NULL
depth31$seqdepth<-str_remove_all(depth31$seqdepth, "[depth]")
depth31<-merge(depth31, metadata, by="row.names")
names(depth32)<-sapply(strsplit(names(depth32), "_iter"), `[`, 1)
colnames(depth32)[1]
depth32$seqdepth<-rep(colnames(depth32)[1], nrow(depth32))
depth32[,1:10]<-NULL
depth32$seqdepth<-str_remove_all(depth32$seqdepth, "[depth]")
depth32<-merge(depth32, metadata, by="row.names")
names(depth33)<-sapply(strsplit(names(depth33), "_iter"), `[`, 1)

```

```

colnames(depth33)[1]
depth33$seqdepth<-rep(colnames(depth33)[1], nrow(depth33))
depth33[,1:10]<-NULL
depth33$seqdepth<-str_remove_all(depth33$seqdepth, "[depth]")
depth33<-merge(depth33, metadata, by="row.names")
names(depth34)<-sapply(strsplit(names(depth34), "_iter"), `[`, 1)
colnames(depth34)[1]
depth34$seqdepth<-rep(colnames(depth34)[1], nrow(depth34))
depth34[,1:10]<-NULL
depth34$seqdepth<-str_remove_all(depth34$seqdepth, "[depth]")
depth34<-merge(depth34, metadata, by="row.names")
names(depth35)<-sapply(strsplit(names(depth35), "_iter"), `[`, 1)
colnames(depth35)[1]
depth35$seqdepth<-rep(colnames(depth35)[1], nrow(depth35))
depth35[,1:10]<-NULL
depth35$seqdepth<-str_remove_all(depth35$seqdepth, "[depth]")
depth35<-merge(depth35, metadata, by="row.names")
names(depth36)<-sapply(strsplit(names(depth36), "_iter"), `[`, 1)
colnames(depth36)[1]
depth36$seqdepth<-rep(colnames(depth36)[1], nrow(depth36))
depth36[,1:10]<-NULL
depth36$seqdepth<-str_remove_all(depth36$seqdepth, "[depth]")
depth36<-merge(depth36, metadata, by="row.names")
names(depth37)<-sapply(strsplit(names(depth37), "_iter"), `[`, 1)
colnames(depth37)[1]
depth37$seqdepth<-rep(colnames(depth37)[1], nrow(depth37))
depth37[,1:10]<-NULL
depth37$seqdepth<-str_remove_all(depth37$seqdepth, "[depth]")
depth37<-merge(depth37, metadata, by="row.names")
names(depth38)<-sapply(strsplit(names(depth38), "_iter"), `[`, 1)
colnames(depth38)[1]
depth38$seqdepth<-rep(colnames(depth38)[1], nrow(depth38))
depth38[,1:10]<-NULL
depth38$seqdepth<-str_remove_all(depth38$seqdepth, "[depth]")
depth38<-merge(depth38, metadata, by="row.names")
names(depth39)<-sapply(strsplit(names(depth39), "_iter"), `[`, 1)
colnames(depth39)[1]
depth39$seqdepth<-rep(colnames(depth39)[1], nrow(depth39))
depth39[,1:10]<-NULL
depth39$seqdepth<-str_remove_all(depth39$seqdepth, "[depth]")
depth39<-merge(depth39, metadata, by="row.names")
names(depth40)<-sapply(strsplit(names(depth40), "_iter"), `[`, 1)
colnames(depth40)[1]
depth40$seqdepth<-rep(colnames(depth40)[1], nrow(depth40))
depth40[,1:10]<-NULL
depth40$seqdepth<-str_remove_all(depth40$seqdepth, "[depth]")
depth40<-merge(depth40, metadata, by="row.names")
names(depth41)<-sapply(strsplit(names(depth41), "_iter"), `[`, 1)
colnames(depth41)[1]
depth41$seqdepth<-rep(colnames(depth41)[1], nrow(depth41))
depth41[,1:10]<-NULL
depth41$seqdepth<-str_remove_all(depth41$seqdepth, "[depth]")
depth41<-merge(depth41, metadata, by="row.names")
names(depth42)<-sapply(strsplit(names(depth42), "_iter"), `[`, 1)
colnames(depth42)[1]
depth42$seqdepth<-rep(colnames(depth42)[1], nrow(depth42))
depth42[,1:10]<-NULL
depth42$seqdepth<-str_remove_all(depth42$seqdepth, "[depth]")
depth42<-merge(depth42, metadata, by="row.names")
names(depth43)<-sapply(strsplit(names(depth43), "_iter"), `[`, 1)
colnames(depth43)[1]
depth43$seqdepth<-rep(colnames(depth43)[1], nrow(depth43))
depth43[,1:10]<-NULL
depth43$seqdepth<-str_remove_all(depth43$seqdepth, "[depth]")
depth43<-merge(depth43, metadata, by="row.names")
names(depth44)<-sapply(strsplit(names(depth44), "_iter"), `[`, 1)

```

```

colnames(depth44)[1]
depth44$seqdepth<-rep(colnames(depth44)[1], nrow(depth44))
depth44[,1:10]<-NULL
depth44$seqdepth<-str_remove_all(depth44$seqdepth, "[depth]")
depth44<-merge(depth44, metadata, by="row.names")
names(depth45)<-sapply(strsplit(names(depth45), "_iter"), `[`, 1)
colnames(depth45)[1]
depth45$seqdepth<-rep(colnames(depth45)[1], nrow(depth45))
depth45[,1:10]<-NULL
depth45$seqdepth<-str_remove_all(depth45$seqdepth, "[depth]")
depth45<-merge(depth45, metadata, by="row.names")
names(depth46)<-sapply(strsplit(names(depth46), "_iter"), `[`, 1)
colnames(depth46)[1]
depth46$seqdepth<-rep(colnames(depth46)[1], nrow(depth46))
depth46[,1:10]<-NULL
depth46$seqdepth<-str_remove_all(depth46$seqdepth, "[depth]")
depth46<-merge(depth46, metadata, by="row.names")
names(depth47)<-sapply(strsplit(names(depth47), "_iter"), `[`, 1)
colnames(depth47)[1]
depth47$seqdepth<-rep(colnames(depth47)[1], nrow(depth47))
depth47[,1:10]<-NULL
depth47$seqdepth<-str_remove_all(depth47$seqdepth, "[depth]")
depth47<-merge(depth47, metadata, by="row.names")
names(depth48)<-sapply(strsplit(names(depth48), "_iter"), `[`, 1)
colnames(depth48)[1]
depth48$seqdepth<-rep(colnames(depth48)[1], nrow(depth48))
depth48[,1:10]<-NULL
depth48$seqdepth<-str_remove_all(depth48$seqdepth, "[depth]")
depth48<-merge(depth48, metadata, by="row.names")
names(depth49)<-sapply(strsplit(names(depth49), "_iter"), `[`, 1)
colnames(depth49)[1]
depth49$seqdepth<-rep(colnames(depth49)[1], nrow(depth49))
depth49[,1:10]<-NULL
depth49$seqdepth<-str_remove_all(depth49$seqdepth, "[depth]")
depth49<-merge(depth49, metadata, by="row.names")
names(depth50)<-sapply(strsplit(names(depth50), "_iter"), `[`, 1)
colnames(depth50)[1]
depth50$seqdepth<-rep(colnames(depth50)[1], nrow(depth50))
depth50[,1:10]<-NULL
depth50$seqdepth<-str_remove_all(depth50$seqdepth, "[depth]")
depth50<-merge(depth50, metadata, by="row.names")
names(depth51)<-sapply(strsplit(names(depth51), "_iter"), `[`, 1)
colnames(depth51)[1]
depth51$seqdepth<-rep(colnames(depth51)[1], nrow(depth51))
depth51[,1:10]<-NULL
depth51$seqdepth<-str_remove_all(depth51$seqdepth, "[depth]")
depth51<-merge(depth51, metadata, by="row.names")
names(depth52)<-sapply(strsplit(names(depth52), "_iter"), `[`, 1)
colnames(depth52)[1]
depth52$seqdepth<-rep(colnames(depth52)[1], nrow(depth52))
depth52[,1:10]<-NULL
depth52$seqdepth<-str_remove_all(depth52$seqdepth, "[depth]")
depth52<-merge(depth52, metadata, by="row.names")
names(depth53)<-sapply(strsplit(names(depth53), "_iter"), `[`, 1)
colnames(depth53)[1]
depth53$seqdepth<-rep(colnames(depth53)[1], nrow(depth53))
depth53[,1:10]<-NULL
depth53$seqdepth<-str_remove_all(depth53$seqdepth, "[depth]")
depth53<-merge(depth53, metadata, by="row.names")
names(depth54)<-sapply(strsplit(names(depth54), "_iter"), `[`, 1)
colnames(depth54)[1]
depth54$seqdepth<-rep(colnames(depth54)[1], nrow(depth54))
depth54[,1:10]<-NULL
depth54$seqdepth<-str_remove_all(depth54$seqdepth, "[depth]")
depth54<-merge(depth54, metadata, by="row.names")
names(depth55)<-sapply(strsplit(names(depth55), "_iter"), `[`, 1)

```

```

colnames(depth55)[1]
depth55$seqdepth<-rep(colnames(depth55)[1], nrow(depth55))
depth55[,1:10]<-NULL
depth55$seqdepth<-str_remove_all(depth55$seqdepth, "[depth]")
depth55<-merge(depth55, metadata, by="row.names")
names(depth56)<-sapply(strsplit(names(depth56), "_iter"), `[`, 1)
colnames(depth56)[1]
depth56$seqdepth<-rep(colnames(depth56)[1], nrow(depth56))
depth56[,1:10]<-NULL
depth56$seqdepth<-str_remove_all(depth56$seqdepth, "[depth]")
depth56<-merge(depth56, metadata, by="row.names")
names(depth57)<-sapply(strsplit(names(depth57), "_iter"), `[`, 1)
colnames(depth57)[1]
depth57$seqdepth<-rep(colnames(depth57)[1], nrow(depth57))
depth57[,1:10]<-NULL
depth57$seqdepth<-str_remove_all(depth57$seqdepth, "[depth]")
depth57<-merge(depth57, metadata, by="row.names")
names(depth58)<-sapply(strsplit(names(depth58), "_iter"), `[`, 1)
colnames(depth58)[1]
depth58$seqdepth<-rep(colnames(depth58)[1], nrow(depth58))
depth58[,1:10]<-NULL
depth58$seqdepth<-str_remove_all(depth58$seqdepth, "[depth]")
depth58<-merge(depth58, metadata, by="row.names")
names(depth59)<-sapply(strsplit(names(depth59), "_iter"), `[`, 1)
colnames(depth59)[1]
depth59$seqdepth<-rep(colnames(depth59)[1], nrow(depth59))
depth59[,1:10]<-NULL
depth59$seqdepth<-str_remove_all(depth59$seqdepth, "[depth]")
depth59<-merge(depth59, metadata, by="row.names")
names(depth60)<-sapply(strsplit(names(depth60), "_iter"), `[`, 1)
colnames(depth60)[1]
depth60$seqdepth<-rep(colnames(depth60)[1], nrow(depth60))
depth60[,1:10]<-NULL
depth60$seqdepth<-str_remove_all(depth60$seqdepth, "[depth]")
depth60<-merge(depth60, metadata, by="row.names")
names(depth61)<-sapply(strsplit(names(depth61), "_iter"), `[`, 1)
colnames(depth61)[1]
depth61$seqdepth<-rep(colnames(depth61)[1], nrow(depth61))
depth61[,1:10]<-NULL
depth61$seqdepth<-str_remove_all(depth61$seqdepth, "[depth]")
depth61<-merge(depth61, metadata, by="row.names")
names(depth62)<-sapply(strsplit(names(depth62), "_iter"), `[`, 1)
colnames(depth62)[1]
depth62$seqdepth<-rep(colnames(depth62)[1], nrow(depth62))
depth62[,1:10]<-NULL
depth62$seqdepth<-str_remove_all(depth62$seqdepth, "[depth]")
depth62<-merge(depth62, metadata, by="row.names")
names(depth63)<-sapply(strsplit(names(depth63), "_iter"), `[`, 1)
colnames(depth63)[1]
depth63$seqdepth<-rep(colnames(depth63)[1], nrow(depth63))
depth63[,1:10]<-NULL
depth63$seqdepth<-str_remove_all(depth63$seqdepth, "[depth]")
depth63<-merge(depth63, metadata, by="row.names")
names(depth64)<-sapply(strsplit(names(depth64), "_iter"), `[`, 1)
colnames(depth64)[1]
depth64$seqdepth<-rep(colnames(depth64)[1], nrow(depth64))
depth64[,1:10]<-NULL
depth64$seqdepth<-str_remove_all(depth64$seqdepth, "[depth]")
depth64<-merge(depth64, metadata, by="row.names")
names(depth65)<-sapply(strsplit(names(depth65), "_iter"), `[`, 1)
colnames(depth65)[1]
depth65$seqdepth<-rep(colnames(depth65)[1], nrow(depth65))
depth65[,1:10]<-NULL
depth65$seqdepth<-str_remove_all(depth65$seqdepth, "[depth]")
depth65<-merge(depth65, metadata, by="row.names")
names(depth66)<-sapply(strsplit(names(depth66), "_iter"), `[`, 1)

```



```

colnames(depth66)[1]
depth66$seqdepth<-rep(colnames(depth66)[1], nrow(depth66))
depth66[,1:10]<-NULL
depth66$seqdepth<-str_remove_all(depth66$seqdepth, "[depth]")
depth66<-merge(depth66, metadata, by="row.names")
names(depth67)<-sapply(strsplit(names(depth67), "_iter"), `[`, 1)
colnames(depth67)[1]
depth67$seqdepth<-rep(colnames(depth67)[1], nrow(depth67))
depth67[,1:10]<-NULL
depth67$seqdepth<-str_remove_all(depth67$seqdepth, "[depth]")
depth67<-merge(depth67, metadata, by="row.names")
names(depth68)<-sapply(strsplit(names(depth68), "_iter"), `[`, 1)
colnames(depth68)[1]
depth68$seqdepth<-rep(colnames(depth68)[1], nrow(depth68))
depth68[,1:10]<-NULL
depth68$seqdepth<-str_remove_all(depth68$seqdepth, "[depth]")
depth68<-merge(depth68, metadata, by="row.names")
names(depth69)<-sapply(strsplit(names(depth69), "_iter"), `[`, 1)
colnames(depth69)[1]
depth69$seqdepth<-rep(colnames(depth69)[1], nrow(depth69))
depth69[,1:10]<-NULL
depth69$seqdepth<-str_remove_all(depth69$seqdepth, "[depth]")
depth69<-merge(depth69, metadata, by="row.names")
names(depth70)<-sapply(strsplit(names(depth70), "_iter"), `[`, 1)
colnames(depth70)[1]
depth70$seqdepth<-rep(colnames(depth70)[1], nrow(depth70))
depth70[,1:10]<-NULL
depth70$seqdepth<-str_remove_all(depth70$seqdepth, "[depth]")
depth70<-merge(depth70, metadata, by="row.names")
names(depth71)<-sapply(strsplit(names(depth71), "_iter"), `[`, 1)
colnames(depth71)[1]
depth71$seqdepth<-rep(colnames(depth71)[1], nrow(depth71))
depth71[,1:10]<-NULL
depth71$seqdepth<-str_remove_all(depth71$seqdepth, "[depth]")
depth71<-merge(depth71, metadata, by="row.names")
names(depth72)<-sapply(strsplit(names(depth72), "_iter"), `[`, 1)
colnames(depth72)[1]
depth72$seqdepth<-rep(colnames(depth72)[1], nrow(depth72))
depth72[,1:10]<-NULL
depth72$seqdepth<-str_remove_all(depth72$seqdepth, "[depth]")
depth72<-merge(depth72, metadata, by="row.names")
names(depth73)<-sapply(strsplit(names(depth73), "_iter"), `[`, 1)
colnames(depth73)[1]
depth73$seqdepth<-rep(colnames(depth73)[1], nrow(depth73))
depth73[,1:10]<-NULL
depth73$seqdepth<-str_remove_all(depth73$seqdepth, "[depth]")
depth73<-merge(depth73, metadata, by="row.names")
names(depth74)<-sapply(strsplit(names(depth74), "_iter"), `[`, 1)
colnames(depth74)[1]
depth74$seqdepth<-rep(colnames(depth74)[1], nrow(depth74))
depth74[,1:10]<-NULL
depth74$seqdepth<-str_remove_all(depth74$seqdepth, "[depth]")
depth74<-merge(depth74, metadata, by="row.names")
names(depth75)<-sapply(strsplit(names(depth75), "_iter"), `[`, 1)
colnames(depth75)[1]
depth75$seqdepth<-rep(colnames(depth75)[1], nrow(depth75))
depth75[,1:10]<-NULL
depth75$seqdepth<-str_remove_all(depth75$seqdepth, "[depth]")
depth75<-merge(depth75, metadata, by="row.names")
names(depth76)<-sapply(strsplit(names(depth76), "_iter"), `[`, 1)
colnames(depth76)[1]
depth76$seqdepth<-rep(colnames(depth76)[1], nrow(depth76))
depth76[,1:10]<-NULL
depth76$seqdepth<-str_remove_all(depth76$seqdepth, "[depth]")
depth76<-merge(depth76, metadata, by="row.names")
names(depth77)<-sapply(strsplit(names(depth77), "_iter"), `[`, 1)

```

```

colnames(depth77)[1]
depth77$seqdepth<-rep(colnames(depth77)[1], nrow(depth77))
depth77[,1:10]<-NULL
depth77$seqdepth<-str_remove_all(depth77$seqdepth, "[depth]")
depth77<-merge(depth77, metadata, by="row.names")
names(depth78)<-sapply(strsplit(names(depth78), "_iter"), `[`, 1)
colnames(depth78)[1]
depth78$seqdepth<-rep(colnames(depth78)[1], nrow(depth78))
depth78[,1:10]<-NULL
depth78$seqdepth<-str_remove_all(depth78$seqdepth, "[depth]")
depth78<-merge(depth78, metadata, by="row.names")
names(depth79)<-sapply(strsplit(names(depth79), "_iter"), `[`, 1)
colnames(depth79)[1]
depth79$seqdepth<-rep(colnames(depth79)[1], nrow(depth79))
depth79[,1:10]<-NULL
depth79$seqdepth<-str_remove_all(depth79$seqdepth, "[depth]")
depth79<-merge(depth79, metadata, by="row.names")
names(depth80)<-sapply(strsplit(names(depth80), "_iter"), `[`, 1)
colnames(depth80)[1]
depth80$seqdepth<-rep(colnames(depth80)[1], nrow(depth80))
depth80[,1:10]<-NULL
depth80$seqdepth<-str_remove_all(depth80$seqdepth, "[depth]")
depth80<-merge(depth80, metadata, by="row.names")
names(depth81)<-sapply(strsplit(names(depth81), "_iter"), `[`, 1)
colnames(depth81)[1]
depth81$seqdepth<-rep(colnames(depth81)[1], nrow(depth81))
depth81[,1:10]<-NULL
depth81$seqdepth<-str_remove_all(depth81$seqdepth, "[depth]")
depth81<-merge(depth81, metadata, by="row.names")
names(depth82)<-sapply(strsplit(names(depth82), "_iter"), `[`, 1)
colnames(depth82)[1]
depth82$seqdepth<-rep(colnames(depth82)[1], nrow(depth82))
depth82[,1:10]<-NULL
depth82$seqdepth<-str_remove_all(depth82$seqdepth, "[depth]")
depth82<-merge(depth82, metadata, by="row.names")
names(depth83)<-sapply(strsplit(names(depth83), "_iter"), `[`, 1)
colnames(depth83)[1]
depth83$seqdepth<-rep(colnames(depth83)[1], nrow(depth83))
depth83[,1:10]<-NULL
depth83$seqdepth<-str_remove_all(depth83$seqdepth, "[depth]")
depth83<-merge(depth83, metadata, by="row.names")
names(depth84)<-sapply(strsplit(names(depth84), "_iter"), `[`, 1)
colnames(depth84)[1]
depth84$seqdepth<-rep(colnames(depth84)[1], nrow(depth84))
depth84[,1:10]<-NULL
depth84$seqdepth<-str_remove_all(depth84$seqdepth, "[depth]")
depth84<-merge(depth84, metadata, by="row.names")
names(depth85)<-sapply(strsplit(names(depth85), "_iter"), `[`, 1)
colnames(depth85)[1]
depth85$seqdepth<-rep(colnames(depth85)[1], nrow(depth85))
depth85[,1:10]<-NULL
depth85$seqdepth<-str_remove_all(depth85$seqdepth, "[depth]")
depth85<-merge(depth85, metadata, by="row.names")
names(depth86)<-sapply(strsplit(names(depth86), "_iter"), `[`, 1)
colnames(depth86)[1]
depth86$seqdepth<-rep(colnames(depth86)[1], nrow(depth86))
depth86[,1:10]<-NULL
depth86$seqdepth<-str_remove_all(depth86$seqdepth, "[depth]")
depth86<-merge(depth86, metadata, by="row.names")
names(depth87)<-sapply(strsplit(names(depth87), "_iter"), `[`, 1)
colnames(depth87)[1]
depth87$seqdepth<-rep(colnames(depth87)[1], nrow(depth87))
depth87[,1:10]<-NULL
depth87$seqdepth<-str_remove_all(depth87$seqdepth, "[depth]")
depth87<-merge(depth87, metadata, by="row.names")
names(depth88)<-sapply(strsplit(names(depth88), "_iter"), `[`, 1)

```

```

colnames(depth88)[1]
depth88$seqdepth<-rep(colnames(depth88)[1], nrow(depth88))
depth88[,1:10]<-NULL
depth88$seqdepth<-str_remove_all(depth88$seqdepth, "[depth]")
depth88<-merge(depth88, metadata, by="row.names")
names(depth89)<-sapply(strsplit(names(depth89), "_iter"), `[`, 1)
colnames(depth89)[1]
depth89$seqdepth<-rep(colnames(depth89)[1], nrow(depth89))
depth89[,1:10]<-NULL
depth89$seqdepth<-str_remove_all(depth89$seqdepth, "[depth]")
depth89<-merge(depth89, metadata, by="row.names")
names(depth90)<-sapply(strsplit(names(depth90), "_iter"), `[`, 1)
colnames(depth90)[1]
depth90$seqdepth<-rep(colnames(depth90)[1], nrow(depth90))
depth90[,1:10]<-NULL
depth90$seqdepth<-str_remove_all(depth90$seqdepth, "[depth]")
depth90<-merge(depth90, metadata, by="row.names")
names(depth91)<-sapply(strsplit(names(depth91), "_iter"), `[`, 1)
colnames(depth91)[1]
depth91$seqdepth<-rep(colnames(depth91)[1], nrow(depth91))
depth91[,1:10]<-NULL
depth91$seqdepth<-str_remove_all(depth91$seqdepth, "[depth]")
depth91<-merge(depth91, metadata, by="row.names")
names(depth92)<-sapply(strsplit(names(depth92), "_iter"), `[`, 1)
colnames(depth92)[1]
depth92$seqdepth<-rep(colnames(depth92)[1], nrow(depth92))
depth92[,1:10]<-NULL
depth92$seqdepth<-str_remove_all(depth92$seqdepth, "[depth]")
depth92<-merge(depth92, metadata, by="row.names")
names(depth93)<-sapply(strsplit(names(depth93), "_iter"), `[`, 1)
colnames(depth93)[1]
depth93$seqdepth<-rep(colnames(depth93)[1], nrow(depth93))
depth93[,1:10]<-NULL
depth93$seqdepth<-str_remove_all(depth93$seqdepth, "[depth]")
depth93<-merge(depth93, metadata, by="row.names")
names(depth94)<-sapply(strsplit(names(depth94), "_iter"), `[`, 1)
colnames(depth94)[1]
depth94$seqdepth<-rep(colnames(depth94)[1], nrow(depth94))
depth94[,1:10]<-NULL
depth94$seqdepth<-str_remove_all(depth94$seqdepth, "[depth]")
depth94<-merge(depth94, metadata, by="row.names")
names(depth95)<-sapply(strsplit(names(depth95), "_iter"), `[`, 1)
colnames(depth95)[1]
depth95$seqdepth<-rep(colnames(depth95)[1], nrow(depth95))
depth95[,1:10]<-NULL
depth95$seqdepth<-str_remove_all(depth95$seqdepth, "[depth]")
depth95<-merge(depth95, metadata, by="row.names")
names(depth96)<-sapply(strsplit(names(depth96), "_iter"), `[`, 1)
colnames(depth96)[1]
depth96$seqdepth<-rep(colnames(depth96)[1], nrow(depth96))
depth96[,1:10]<-NULL
depth96$seqdepth<-str_remove_all(depth96$seqdepth, "[depth]")
depth96<-merge(depth96, metadata, by="row.names")
names(depth97)<-sapply(strsplit(names(depth97), "_iter"), `[`, 1)
colnames(depth97)[1]
depth97$seqdepth<-rep(colnames(depth97)[1], nrow(depth97))
depth97[,1:10]<-NULL
depth97$seqdepth<-str_remove_all(depth97$seqdepth, "[depth]")
depth97<-merge(depth97, metadata, by="row.names")
names(depth98)<-sapply(strsplit(names(depth98), "_iter"), `[`, 1)
colnames(depth98)[1]
depth98$seqdepth<-rep(colnames(depth98)[1], nrow(depth98))
depth98[,1:10]<-NULL
depth98$seqdepth<-str_remove_all(depth98$seqdepth, "[depth]")
depth98<-merge(depth98, metadata, by="row.names")
names(depth99)<-sapply(strsplit(names(depth99), "_iter"), `[`, 1)

```

```

colnames(depth99)[1]
depth99$seqdepth<-rep(colnames(depth99)[1], nrow(depth99))
depth99[,1:10]<-NULL
depth99$seqdepth<-str_remove_all(depth99$seqdepth, "[depth]")
depth99<-merge(depth99, metadata, by="row.names")
names(depth100)<-sapply(strsplit(names(depth100), "_iter"), `[[`, 1)
colnames(depth100)[1]
depth100$seqdepth<-rep(colnames(depth100)[1], nrow(depth100))
depth100[,1:10]<-NULL
depth100$seqdepth<-str_remove_all(depth100$seqdepth, "[depth]")
depth100<-merge(depth100, metadata, by="row.names")
d1<-rbind(depth1,depth2)
d2<-rbind(d1,depth3)
d3<-rbind(d2, depth4)
d4<-rbind(d3,depth5)
d5<-rbind(d4,depth6)
d6<-rbind(d5,depth7)
d7<-rbind(d6,depth8)
d8<-rbind(d7,depth9)
d9<-rbind(d8,depth10)
d10<-rbind(d9,depth11)
d11<-rbind(d10,depth12)
d12<-rbind(d11,depth13)
d13<-rbind(d12,depth14)
d14<-rbind(d13,depth15)
d15<-rbind(d14,depth16)
d16<-rbind(d15,depth17)
d17<-rbind(d16,depth18)
d18<-rbind(d17,depth19)
d19<-rbind(d18,depth20)
d20<-rbind(d19,depth21)
d21<-rbind(d20,depth22)
d22<-rbind(d21,depth23)
d23<-rbind(d22,depth24)
d24<-rbind(d23,depth25)
d25<-rbind(d24,depth26)
d26<-rbind(d25,depth27)
d27<-rbind(d26,depth28)
d28<-rbind(d27,depth29)
d29<-rbind(d28,depth30)
d30<-rbind(d29,depth31)
d31<-rbind(d30,depth32)
d32<-rbind(d31,depth33)
d33<-rbind(d32,depth34)
d34<-rbind(d33,depth35)
d35<-rbind(d34,depth36)
d36<-rbind(d35,depth37)
d37<-rbind(d36,depth38)
d38<-rbind(d37,depth39)
d39<-rbind(d38,depth40)
d40<-rbind(d39,depth41)
d41<-rbind(d40,depth42)
d42<-rbind(d41,depth43)
d43<-rbind(d42,depth44)
d44<-rbind(d43,depth45)
d45<-rbind(d44,depth46)
d46<-rbind(d45,depth47)
d47<-rbind(d46,depth48)
d48<-rbind(d47,depth49)
d49<-rbind(d48,depth50)
d50<-rbind(d49,depth51)
d51<-rbind(d50,depth52)
d52<-rbind(d51,depth53)
d53<-rbind(d52,depth54)
d54<-rbind(d53,depth55)
d55<-rbind(d54,depth56)

```

```

d56<-rbind(d55,depth57)
d57<-rbind(d56,depth58)
d58<-rbind(d57,depth59)
d59<-rbind(d58,depth60)
d60<-rbind(d59,depth61)
d61<-rbind(d60,depth62)
d62<-rbind(d61,depth63)
d63<-rbind(d62,depth64)
d64<-rbind(d63,depth65)
d65<-rbind(d64,depth66)
d66<-rbind(d65,depth67)
d67<-rbind(d66,depth68)
d68<-rbind(d67,depth69)
d69<-rbind(d68,depth70)
d70<-rbind(d69,depth71)
d71<-rbind(d70,depth72)
d72<-rbind(d71,depth73)
d73<-rbind(d72,depth74)
d74<-rbind(d73,depth75)
d75<-rbind(d74,depth76)
d76<-rbind(d75,depth77)
d77<-rbind(d76,depth78)
d78<-rbind(d77,depth79)
d79<-rbind(d78,depth80)
d80<-rbind(d79,depth81)
d81<-rbind(d80,depth82)
d82<-rbind(d81,depth83)
d83<-rbind(d82,depth84)
d84<-rbind(d83,depth85)
d85<-rbind(d84,depth86)
d86<-rbind(d85,depth87)
d87<-rbind(d86,depth88)
d88<-rbind(d87,depth89)
d89<-rbind(d88,depth90)
d90<-rbind(d89,depth91)
d91<-rbind(d90,depth92)
d92<-rbind(d91,depth93)
d93<-rbind(d92,depth94)
d94<-rbind(d93,depth95)
d95<-rbind(d94,depth96)
d96<-rbind(d95,depth97)
d97<-rbind(d96,depth98)
d98<-rbind(d97,depth99)
d99<-rbind(d98,depth100)
names(d99)[1] <- "SampleID"
names(d99)[names(d99) == "mean"] <- "observedASVs"
d99$seqdepth<-as.numeric(d99$seqdepth)
write.table(d99, file = "raref.txt",sep="\t",col.names= T, row.names=F)
quit("no")

```

```
sed 's/"//g' raref.txt > raref2.txt
```

```

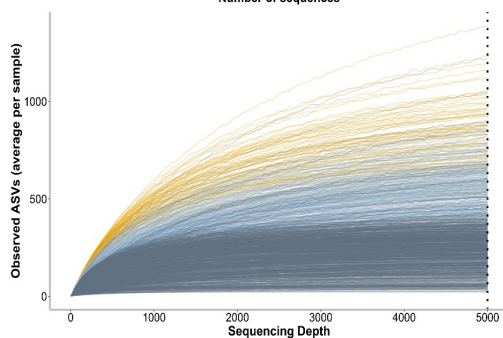
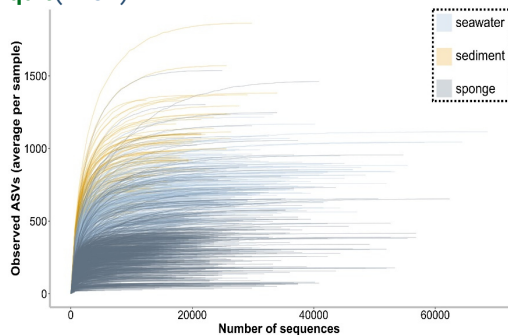
R
list.of.packages <- c("ggplot2","gridExtra")
new.packages <- list.of.packages[!(list.of.packages %in% installed.packages()[,"Package"])]
if(length(new.packages)) install.packages(new.packages, repos='http://cran.us.r-project.org')
library(ggplot2)
library(gridExtra)
sub<- read.table(file="raref2.txt", header =T)
comp<- read.table(file="raref_sub2.txt", header =T)
pdf(file="rarefaction.pdf", height= 14, width= 14)
p1<-ggplot(data=comp, aes(x=seqdepth, y=observedASVs, group=SampleID, color=samplotype)) +
geom_line(alpha=0.3) +
labs(x = "Sequencing Depth", y="Observed ASVs (average per sample)") +
scale_color_manual(values=c('#a0bed8ff', '#E69F00', '#607281'))+
geom_vline(xintercept = max(comp$seqdepth), linetype="dotted", size=1.5)+

```

```

theme(panel.grid.major = element_blank(),
panel.grid.minor = element_blank(),
axis.line = element_line(colour = "grey", size = 1.5, linetype = "solid"),
panel.background = element_rect(fill = "#ffffff",colour = NA),
plot.background = element_rect(fill = '#ffffff'),
legend.title= element_blank(),
legend.text = element_text(color = "black", size = 18),
axis.title.x = element_text(color="black", size=20, face="bold"),
axis.text.x = element_text(face="plain", color="black",size=18, angle=0),
axis.title.y = element_text(color="black", size=20, face="bold"),
axis.text.y = element_text(face="plain", color="black",size=18, angle=0),
legend.key=element_rect(fill='white'),
legend.key.size = unit(3,"line"))
p2<-ggplot(data=sub, aes(x=seqdepth, y=observedASVs, group=SampleID, color=samplotype)) +
geom_line(alpha=0.3) +
labs(x = "Number of sequences", y="Observed ASVs (average per sample)") +
scale_color_manual(values=c('#a0bed8ff', '#E69F00', '#607281'))+
theme(panel.grid.major = element_blank(),
panel.grid.minor = element_blank(),
axis.line = element_line(colour = "grey", size = 1.5, linetype = "solid"),
panel.background = element_rect(fill = "#ffffff",colour = NA),
plot.background = element_rect(fill = '#ffffff'),
legend.title= element_blank(),
legend.text = element_text(color = "black", size = 18),
axis.title.x = element_text(color="black", size=20, face="bold"),
axis.text.x = element_text(face="plain", color="black",size=18, angle=0),
axis.title.y = element_text(color="black", size=20, face="bold"),
axis.text.y = element_text(face="plain", color="black",size=18, angle=0),
legend.key=element_rect(fill='white'),
legend.key.size = unit(3,"line"))
grid.arrange(p2, p1, nrow=2)
dev.off()
quit("no")

```



```

rm -r -f -- - max.txt
rm -r -f -- - observed_otus.txt
rm -r -f -- - observed_otus2.txt
rm -r -f -- - raref2.txt
rm -r -f -- - raref.txt
rm -r -f -- - observed_otusN.txt
rm -r -f -- - raref_sub2.txt
rm -r -f -- - raref_sub.txt
cd $PWDa/dada2/core-metrics-results/
rm -r -f -- - -d */

```

0.11 Export clustering dendrogram

```
qiime diversity beta-rarefaction \  
--i-table $PWDa/dada2/dada2-tablenochloroplastsmitochondriaUnassigned$var.qza \  
--i-phylogeny $PWDa/dada2/rooted-tree.qza \  
--p-sampling-depth $var \  
--p-metric weighted_unifrac \  
--p-clustering-method upgma \  
--m-metadata-file $PWDa/metadataV34.txt \  
--o-visualization $PWDa/dada2/core-metrics-results/beta-rarefaction-upgma.qzv  
  
cd $PWDa/  
mkdir R  
cd $PWDa/dada2/core-metrics-results  
unzip beta-rarefaction-upgma.qzv  
cd "$(\ls -1dt ./*/ | head -n 1)/data"  
cp sample-clustering-upgma.tre $PWDa/R/sample-clustering-upgma.tre  
cd $PWDa/dada2/core-metrics-results  
rm -r -f -- - -d */
```

0.12 Data crunching for format compatibility

0.12.1 Export ASV Table

```
cd $PWDa/dada2/  
cp -r "dada2-tablenochloroplastsmitochondriaUnassigned$var.qza" $PWDa/R/  
cd $PWDa/R/  
unzip dada2-tablenochloroplastsmitochondriaUnassigned$var.qza  
cd "$(\ls -1dt ./*/ | head -n 1)/data"  
biom convert -i feature-table.biom -o table.tsv --to-tsv  
sed '1d' table.tsv > tmpfile; mv tmpfile table.txt  
sed -e '1s/#OTU ID/Feature ID/' table.txt > tmpfile; mv tmpfile tableforR.txt  
cp -r "tableforR.txt" $PWDa/R/  
cd $PWDa/R/  
echo "Feature ID" | awk '{ gsub (" ", "", $0); print}' tableforR.txt > tableforR2.txt  
rm -r -f -- - tableforR.txt  
mv tableforR2.txt tableforR.txt
```

0.12.2 Export Taxonomy Table

```
qiime tools export \  
--input-path $PWDa/dada2/Silva138V3V4-taxonomy-dada2.qza \  
--output-path $PWDa/R/taxonomy  
  
cd $PWDa/R/taxonomy/  
awk -F, -v OFS="," '{gsub(/ /, "", $Taxon)}1' taxonomy.tsv > tmpfile; mv tmpfile taxonomy.txt  
  
R  
list.of.packages <- c("stringr", "plyr")  
new.packages <- list.of.packages[!(list.of.packages %in% installed.packages()[,"Package"])]  
if(length(new.packages)) install.packages(new.packages, repos='http://cran.us.r-project.org')  
library(stringr)  
library(plyr)  
taxonomy<-read.table(file = "taxonomy.txt", header = T, sep="\t")  
datnew<-str_split_fixed(taxonomy$Taxon, ";", 7)  
d=as.data.frame(datnew)  
d$V1 <- sub("^$", "___", d$V1)  
d$V2 <- sub("^$", "___", d$V2)  
d$V3 <- sub("^$", "___", d$V3)  
d$V4 <- sub("^$", "___", d$V4)  
d$V5 <- sub("^$", "___", d$V5)  
d$V6 <- sub("^$", "___", d$V6)  
d$V7 <- sub("^$", "___", d$V7)
```

```

a<- paste(d$V1,d$V2,d$V3,d$V4,d$V5,d$V6,d$V7,sep=";")
b<- cbind(taxonomy,a)
b<- cbind(b,d$V1,d$V2,d$V3,d$V4,d$V5,d$V6,d$V7)
b$Taxon<-NULL
b<-rename(b, c("a"="Taxon"))
b<-b[!grepl("Chloroplast", b$Taxon),]
b<-b[!grepl("Mitochondria", b$Taxon),]
b<-b[!grepl("Unassigned", b$Taxon),]
b<-rename(b, c("d$V1"="kingdom","d$V2"="phylum","d$V3"="class","d$V4"="order","d$V5"="family","d$V6"
write.table(b, file = "taxonomyforR.txt",sep="\t",col.names= T, row.names=F)
quit("no")

```

```

cp -r "taxonomyforR.txt" $PWDa/R/
cd $PWDa/R/

```

0.12.3 Export Metadata file

```

qiime taxa collapse \
--i-table $PWDa/dada2/dada2-tablechloroplastsmitchondriaUnassigned$var.qza \
--i-taxonomy $PWDa/dada2/Silva138V3V4-taxonomy-dada2.qza \
--p-level 7 \
--output-dir taxtable

```

```

qiime tools export \
--input-path $PWDa/R/taxtable/collapsed_table.qza \
--output-path $PWDa/R/taxtable/collapsed_table

```

```

cd $PWDa/R/taxtable/collapsed_table/
biom convert -i feature-table.biom -o table.tsv --to-tsv
sed '1d' table.tsv > tmpfile; mv tmpfile table.txt
cut -d'\t' -f2- table.txt > table2.txt

```

```

R
data<-read.table(file ="table2.txt", header = F)
select<-data[1,]
turn<-t(select)
write.table(turn, file = "list_of_remaining_samples.txt",sep="\t",col.names= T, row.names=F)
quit("no")

```

```

cp -r "list_of_remaining_samples.txt" $PWDa/R/
cd $PWDa/
cp -r "metadataV34.txt" $PWDa/R/
cd $PWDa/R/
awk -F, -v OFS="," '{gsub(/ /,"")}1' metadataV34.txt > tmpfile; mv tmpfile metadataV34.txt
awk -F, -v OFS="," '{gsub(/#/,")}1' metadataV34.txt > tmpfile; mv tmpfile metadataV34.txt

```

```

R
list.of.packages <- c("stringr", "plyr")
new.packages <- list.of.packages[!(list.of.packages %in% installed.packages()[,"Package"])]
if(length(new.packages)) install.packages(new.packages, repos='http://cran.us.r-project.org')
library(stringr)
library(plyr)
metadata<-read.table(file ="list_of_remaining_samples.txt", header = T, sep="\t")
metadata2<-read.table(file ="metadataV34.txt", header = T, sep="\t")
names(metadata)[1] <- "SampleID"
data<-merge(metadata, metadata2, by="SampleID")
write.table(data, file = "metadataforR.txt",sep="\t",col.names= T, row.names=F)
quit("no")

```


0.12.4 Export UniFrac Matrix

```
cd $PWDa/dada2/core-metrics-results/
cp -r "weighted_unifrac_distance_matrix.qza" $PWDa/R/
cd $PWDa/R/
unzip weighted_unifrac_distance_matrix.qza
cd "$(\ls -1dt ./*/ | head -n 1)/data"
sed -e 's/,/\t/g' distance-matrix.tsv > tmpfile; mv tmpfile weighted_unifrac_distance_ma-
trix.txt
cp -r "weighted_unifrac_distance_matrix.txt" $PWDa/R/
```

0.12.5 Create local blastable database of representative sequences

```
cd $PWDa/dada2/
cp -r "dada2-rep-seqs.qza" $PWDa/R/
cd $PWDa/R/
unzip dada2-rep-seqs.qza
cd "$(\ls -1dt ./*/ | head -n 1)/data"
cp -r "dna-sequences.fasta" $PWDa/R/
cd $PWDa/R/
makeblastdb -in dna-sequences.fasta -dbtype nucl -out RepresentativeSequencesDatabase
#blastn -db $PWDa/R/RepresentativeSequencesDatabase -query $PWDa/R/QUERYNAME.fasta -outfmt 6 -num
```

0.12.6 Export table with relative abundances on phylum levels

```
cd $PWDa/R/

R
mydata =read.table(file = "tableforR.txt", header = T)
my = read.table(file = "taxonomyforR.txt", header = T)
x<-merge(my, mydata, by="FeatureID")
x$FeatureID<-NULL
x$Confidence<-NULL
x$Taxon<-NULL
x$kingdom<-NULL
x$class<-NULL
x$order<-NULL
x$family<-NULL
x$genus<-NULL
x$species<-NULL
fin<-aggregate(. ~ phylum, data = x, sum)
rownames(fin) <- fin[,1]
fin[,1] <- NULL
fin <- droplevels(fin[rowSums(fin) != 0,])
fin <- droplevels(fin[,colSums(fin) != 0])
write.table(fin, file = "fin.txt", sep="\t", col.names= T, row.names=T)
quit("no")

sed -i -e 's/p__/g' fin.txt
echo "__" | awk '{ gsub ("__", "Unclassified", $0); print}' fin.txt > fin2.txt
sed "s/([^\]]*)/()/g" fin2.txt > fin3.txt
cat fin3.txt | tr -d '(' > fin4.txt

R
fin = read.table(file = "fin4.txt", header = T, row.names=1)
phylum <- rownames(fin)
rela<-sapply(fin, function(x) x/sum(x), USE.NAMES = TRUE)
rela2<-cbind(phylum, rela)
write.table(rela2, file = "phylum_relabund.txt", sep="\t", col.names= T, row.names=F)
quit("no")
```

```

R
list.of.packages <- c("tidyr")
new.packages <- list.of.packages[!(list.of.packages %in% installed.packages()[,"Package"])]
if(length(new.packages)) install.packages(new.packages, repos='http://cran.us.r-project.org')
library(tidyr)
mydata = read.table(file ="phylum_relabund.txt", header = T)
mydata=as.data.frame(mydata)
a<-names(mydata)[1]
b<-names(mydata)[ncol(mydata)]
data<-gather(mydata,condition, measurement, names(mydata)[2:ncol(mydata)], factor_key=T)
dat <- subset(data, condition != "phylum")
names(dat)[names(dat) == "condition"] <- "SampleID"
names(dat)[names(dat) == "measurement"] <- "relabund"
write.table(dat, file = "phylum_tidlyformat.txt",sep="\t",col.names= T, row.names=F)
quit("no")

```

0.12.7 Export data for SVampEx & Clean-up

```

cd $PWDa/R/
rm -- - dada2-tablechloroplastsmitochondriaUnassigned$var.qza
rm -- - weighted_unifrac_distance_matrix.qza
rm -r -f -- - taxtable
rm -r -f -- - taxonomy
rm -- - metadataV34.txt
rm -- - list_of_remaining_samples.txt
rm -- - dada2-rep-seqs.qza
rm -r -f -- - fin.txt
rm -r -f -- - fin2.txt
rm -r -f -- - fin3.txt
rm -r -f -- - fin4.txt
rm -r -f -- - -d */
cd $PWDa/
mkdir svampex
cd $PWDa/svampex
mkdir taxonomy
cd $PWDa/R
cp -r "phylum_relabund.txt" $PWDa/svampex/taxonomy
cd $PWDa/svampex/taxonomy

```

```

R
phy = read.table(file ="phylum_relabund.txt", header = F, row.names=1)
names(phy) = phy[1,] [phy[1,] !="phylum"]
for(i in 1:ncol(phy)){
write.table(phy[,i],row.names = row.names(phy), col.names
=F,file=paste0(names(phy)[i],".txt"))
}
quit("no")

```

```

rm -r -f -- - phylum_relabund.txt
cd $PWDa/svampex
mkdir amplitudes
cd $PWDa/R
cp -r "phylum_tidlyformat.txt" $PWDa/svampex/amplitudes
cp -r "metadataforR.txt" $PWDa/svampex/amplitudes
cd $PWDa/svampex/amplitudes

```

```

R
dat=read.table(file ="phylum_tidlyformat.txt", header = T)
meta = read.table(file ="metadataforR.txt", header = T)
fin<-merge(dat, meta, by="SampleID")
write.table(fin, file = "amplitudes.txt",sep="\t",col.names= T, row.names=F)
quit("no")

```

```

rm -r -f -- - phylum_tidlyformat.txt
rm -r -f -- - metadataforR.txt
cd $PWDa/

```

```
mv $PWDa/trimmed $PWDa/dada2/trimmed
```

0.13 Diversity indices

```
cd $PWDa/dada2/core-metrics-results/  
unzip evenness_vector.qza  
cd "$(\ls -1dt ./*/ | head -n 1)/data"  
sed 's|,|\t|g' alpha-diversity.tsv > evenness.txt  
sed -i '1s/^/SampleID /' evenness.txt  
cp evenness.txt $PWDa/dada2/evenness.txt  
cd $PWDa/dada2/core-metrics-results/  
rm -r -f -- - -d */  
unzip faith_pd_vector.qza  
cd "$(\ls -1dt ./*/ | head -n 1)/data"  
sed 's|,|\t|g' alpha-diversity.tsv > faith_pd_vector.txt  
sed -i '1s/^/SampleID /' faith_pd_vector.txt  
cp faith_pd_vector.txt $PWDa/dada2/faith_pd_vector.txt  
cd $PWDa/dada2/core-metrics-results/  
rm -r -f -- - -d */  
unzip observed_otus_vector.qza  
cd "$(\ls -1dt ./*/ | head -n 1)/data"  
sed 's|,|\t|g' alpha-diversity.tsv > observed_otus_vector.txt  
sed -i '1s/^/SampleID /' observed_otus_vector.txt  
cp observed_otus_vector.txt $PWDa/dada2/observed_otus_vector.txt  
cd $PWDa/dada2/core-metrics-results/  
rm -r -f -- - -d */  
unzip shannon_vector.qza  
cd "$(\ls -1dt ./*/ | head -n 1)/data"  
sed 's|,|\t|g' alpha-diversity.tsv > shannon_vector.txt  
sed -i '1s/^/SampleID /' shannon_vector.txt  
cp shannon_vector.txt $PWDa/dada2/shannon_vector.txt  
cd $PWDa/dada2/core-metrics-results/  
rm -r -f -- - -d */  
cd $PWDa/dada2/
```

R

```
data1<- read.table(file="evenness.txt", header =T)  
data2<- read.table(file="faith_pd_vector.txt", header =T)  
data3<- read.table(file="observed_otus_vector.txt", header =T)  
data4<- read.table(file="shannon_vector.txt", header =T)  
d<-merge(data1,data2, by="SampleID")  
d1<-merge(d,data3, by="SampleID")  
d2<-merge(d1,data4, by="SampleID")  
names(d2)[names(d2) == "observed_otus"] <- "observed_asvs"  
write.table(d2, file = "alpha_diversity.txt",sep="\t",col.names= T, row.names=F)  
quit("no")
```

```
sed 's|s|//g' alpha_diversity.txt > alpha_diversity_indices.txt  
cp alpha_diversity_indices.txt $PWDa/R/alpha_diversity_indices.txt  
rm -r -f -- - evenness.txt  
rm -r -f -- - faith_pd_vector.txt  
rm -r -f -- - observed_otus_vector.txt  
rm -r -f -- - shannon_vector.txt  
rm -r -f -- - alpha_diversity.txt  
rm -r -f -- - alpha_diversity_indices.txt  
cd $PWDa/R/
```

R

```
list.of.packages <- c("gridExtra", "ggplot2")  
new.packages <- list.of.packages[!(list.of.packages %in% installed.packages()[,"Package"])]  
if(length(new.packages)) install.packages(new.packages, repos='http://cran.us.r-project.org')  
library(ggplot2)  
library(gridExtra)  
dat<-read.table("alpha_diversity_indices.txt", header=TRUE)  
metadata<-read.table("metadataforR.txt", header=TRUE)  
data<-merge(dat, metadata, by="SampleID")
```

```

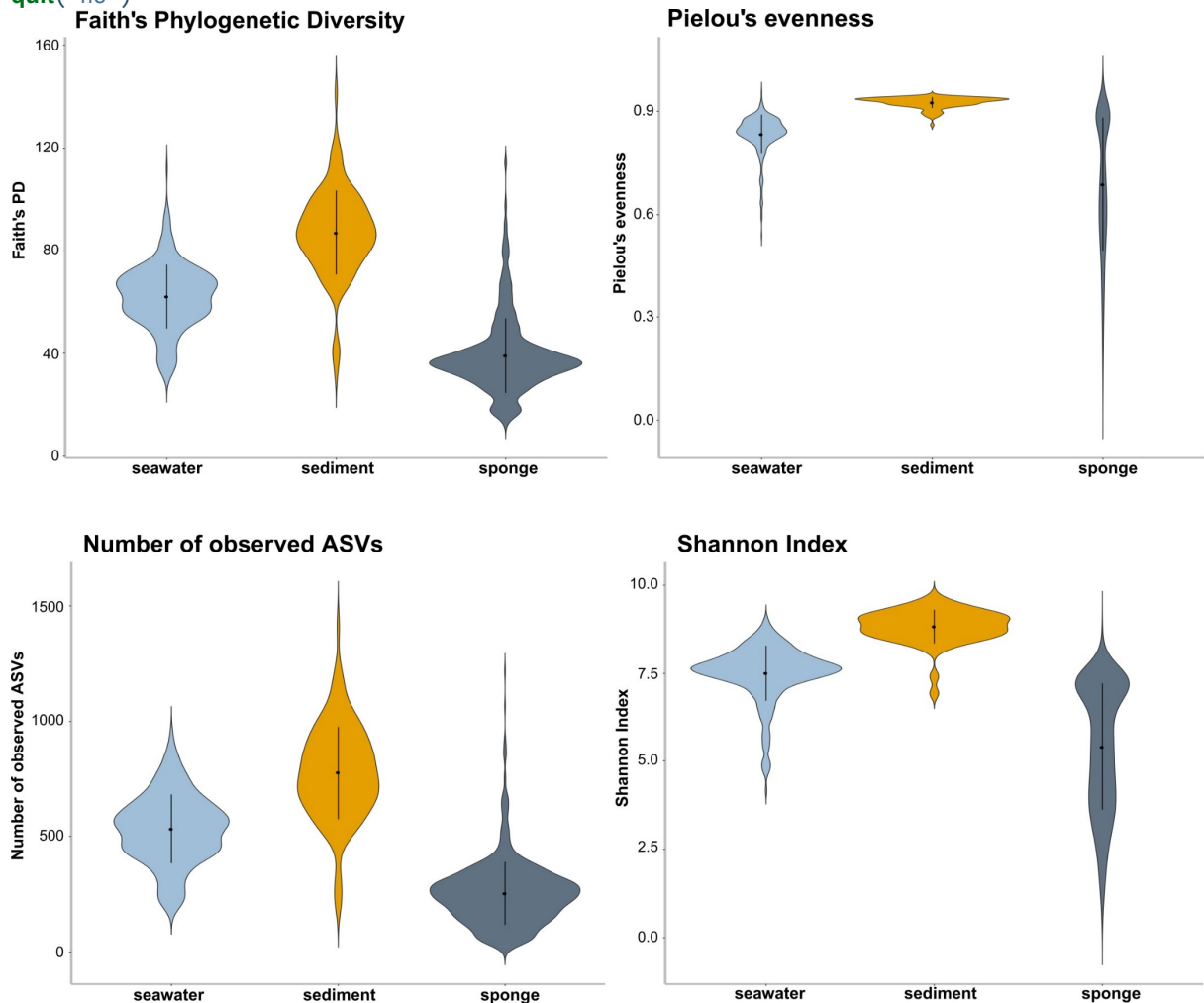
data_summary <- function(x) {
  m <- mean(x)
  ymin <- m-sd(x)
  ymax <- m+sd(x)
  return(c(y=m,ymin=ymin,ymax=ymax))
}
p1 <- ggplot(data, aes(x=samplotype, y=faith_pd, fill=samplotype)) +
  geom_violin(trim=FALSE) +
  stat_summary(fun.data=data_summary) +
  labs(x = "nMDS1", y="Faith's PD", title="Faith's Phylogenetic Diversity") +
  scale_fill_manual(values=c('#a0bed8ff', '#E69F00', '#607281'))+
  theme(panel.grid.major = element_blank(),
  panel.grid.minor = element_blank(),
  axis.line = element_line(colour = "grey", size = 1.5, linetype = "solid"),
  panel.background = element_rect(fill = "#ffffff",colour = NA),
  plot.background = element_rect(fill = '#ffffff'),
  plot.title = element_text(color="black", size=20, face="bold"),
  legend.position="none",
  legend.title= element_blank(),
  legend.text=element_text(size=20),
  legend.background = element_rect(linetype = 'dotted', size = 1, colour = 1),
  axis.title.x = element_blank(),
  axis.text.x = element_text(color="black", size=20, face="bold"),
  axis.title.y = element_text(color="black", size=20, face="bold"),
  axis.text.y = element_text(face="plain", color="black",size=18, angle=0))
p2 <- ggplot(data, aes(x=samplotype, y=pielou_e, fill=samplotype)) +
  geom_violin(trim=FALSE) +
  stat_summary(fun.data=data_summary) +
  labs(x = "nMDS1", y="Pielou's evenness", title="Pielou's evenness") +
  scale_fill_manual(values=c('#a0bed8ff', '#E69F00', '#607281'))+
  theme(panel.grid.major = element_blank(),
  panel.grid.minor = element_blank(),
  axis.line = element_line(colour = "grey", size = 1.5, linetype = "solid"),
  panel.background = element_rect(fill = "#ffffff",colour = NA),
  plot.background = element_rect(fill = '#ffffff'),
  plot.title = element_text(color="black", size=20, face="bold"),
  legend.position="none",
  legend.title= element_blank(),
  legend.text=element_text(size=20),
  legend.background = element_rect(linetype = 'dotted', size = 1, colour = 1),
  axis.title.x = element_blank(),
  axis.text.x = element_text(color="black", size=20, face="bold"),
  axis.title.y = element_text(color="black", size=20, face="bold"),
  axis.text.y = element_text(face="plain", color="black",size=18, angle=0))
p3 <- ggplot(data, aes(x=samplotype, y=observed_asvs, fill=samplotype)) +
  geom_violin(trim=FALSE) +
  stat_summary(fun.data=data_summary) +
  labs(x = "nMDS1", y="Number of observed ASVs", title="Number of observed ASVs") +
  scale_fill_manual(values=c('#a0bed8ff', '#E69F00', '#607281'))+
  theme(panel.grid.major = element_blank(),
  panel.grid.minor = element_blank(),
  axis.line = element_line(colour = "grey", size = 1.5, linetype = "solid"),
  panel.background = element_rect(fill = "#ffffff",colour = NA),
  plot.background = element_rect(fill = '#ffffff'),
  plot.title = element_text(color="black", size=20, face="bold"),
  legend.position="none",
  legend.title= element_blank(),
  legend.text=element_text(size=20),
  legend.background = element_rect(linetype = 'dotted', size = 1, colour = 1),
  axis.title.x = element_blank(),
  axis.text.x = element_text(color="black", size=20, face="bold"),
  axis.title.y = element_text(color="black", size=20, face="bold"),
  axis.text.y = element_text(face="plain", color="black",size=18, angle=0))
p4 <- ggplot(data, aes(x=samplotype, y=shannon, fill=samplotype)) +
  geom_violin(trim=FALSE) +
  stat_summary(fun.data=data_summary) +

```

```

labs(x = "nMDS1", y="Shannon Index", title="Shannon Index") +
scale_fill_manual(values=c('#a0bed8ff', '#E69F00','#607281'))+
theme(panel.grid.major = element_blank(),
panel.grid.minor = element_blank(),
axis.line = element_line(colour = "grey", size = 1.5, linetype = "solid"),
panel.background = element_rect(fill = "#ffffff",colour = NA),
plot.background = element_rect(fill = '#ffffff'),
plot.title = element_text(color="black", size=20, face="bold"),
legend.position="none",
legend.title= element_blank(),
legend.text=element_text(size=20),
legend.background = element_rect(linetype = 'dotted', size = 1, colour = 1),
axis.title.x = element_blank(),
axis.text.x = element_text(color="black", size=20, face="bold"),
axis.title.y = element_text(color="black", size=20, face="bold"),
axis.text.y = element_text(face="plain", color="black",size=18, angle=0))
pdf(file="alpha_diversity.pdf", height= 20, width= 20)
grid.arrange(p1, p2 , p3, p4, nrow=2)
dev.off()
quit("no")

```



```

cd $PWDa/dada2/core-metrics-results/
unzip bray_curtis_distance_matrix.qza
cd "$(\ls -ldt ./ */ | head -n 1)/data"
sed -e 's/,/\t/g' distance-matrix.tsv > tmpfile; mv tmpfile bray_curtis_distance_matrix.txt
cp -r "bray_curtis_distance_matrix.txt" $PWDa/R/
cd $PWDa/dada2/core-metrics-results/
rm -r -f -- -d */
unzip jaccard_distance_matrix.qza
cd "$(\ls -ldt ./ */ | head -n 1)/data"
sed -e 's/,/\t/g' distance-matrix.tsv > tmpfile; mv tmpfile jaccard_distance_matrix.txt

```

```

cp -r "jaccard_distance_matrix.txt" $PWDa/R/
cd $PWDa/dada2/core-metrics-results/
rm -r -f -- - -d */
unzip unweighted_unifrac_distance_matrix.qza
cd "$(\ls -ldt ./ */ | head -n 1)/data"
sed -e 's/,/\t/g' distance-matrix.tsv > tmpfile; mv tmpfile unweighted_unifrac_distance_ma-
trix.txt
cp -r "unweighted_unifrac_distance_matrix.txt" $PWDa/R/
cd $PWDa/dada2/core-metrics-results/
rm -r -f -- - -d */
cd $PWDa/R/

```

```

R
list.of.packages <- c("gridExtra", "ggplot2", "MASS", "vegan")
new.packages <- list.of.packages[!(list.of.packages %in% installed.packages()), "Package"]]
if(length(new.packages)) install.packages(new.packages, repos='http://cran.us.r-project.org')
library(vegan)
library(MASS)
library(ggplot2)
library(gridExtra)
set.seed(201)
community_matrix = read.table(file = "weighted_unifrac_distance_matrix.txt", header = T)
mydata = read.table(file = "metadataforR.txt", header = T)
weighted_uni_NMDS <- metaMDS(community_matrix, k=50, maxit=100, trace=TRUE)
stress_NMDS_weighted_uni <- weighted_uni_NMDS$stress
stress_NMDS_weighted_uni <- round(stress_NMDS_weighted_uni, digits = 3)
NMDS_weighted_uni = data.frame(MDS1 = weighted_uni_NMDS$points[,1], MDS2 =
weighted_uni_NMDS$points[
p1 <- ggplot(data = NMDS_weighted_uni, aes(MDS1, MDS2, col=mydata$samplotype, shape=my-
data$blank)) +
geom_point(size=6, alpha=0.5) +
labs(x = "nMDS1", y="nMDS2", title="Weighted UniFrac distances") +
annotate("text", max(NMDS_weighted_uni$MDS1)-0.05, min(NMDS_weighted_uni$MDS2), label =
paste("Str
scale_color_manual(values=c('#a0bed8ff', '#E69F00', '#607281'))+
theme(panel.grid.major = element_blank(),
panel.grid.minor = element_blank(),
axis.line = element_line(colour = "grey", size = 1.5, linetype = "solid"),
panel.background = element_rect(fill = "#ffffff", colour = NA),
plot.background = element_rect(fill = '#ffffff'),
legend.position="top",
legend.title= element_blank(),
legend.text=element_text(size=20),
legend.background = element_rect(linetype = 'dotted', size = 1, colour = 1),
plot.title = element_text(color="black", size=20, face="bold"),
axis.title.x = element_text(color="black", size=20, face="bold"),
axis.text.x = element_text(face="plain", color="black", size=18, angle=0),
axis.title.y = element_text(color="black", size=20, face="bold"),
axis.text.y = element_text(face="plain", color="black", size=18, angle=0))
community_matrix = read.table(file = "unweighted_unifrac_distance_matrix.txt", header = T)
mydata = read.table(file = "metadataforR.txt", header = T)
unweighted_uni_NMDS <- metaMDS(community_matrix, k=50, maxit=100, trace=TRUE)
stress_NMDS_unweighted_uni <- unweighted_uni_NMDS$stress
stress_NMDS_unweighted_uni <- round(stress_NMDS_unweighted_uni, digits = 3)
NMDS_unweighted_uni = data.frame(MDS1 = unweighted_uni_NMDS$points[,1], MDS2 = un-
weighted_uni_NMDS$p
p2 <- ggplot(data = NMDS_unweighted_uni, aes(MDS1, MDS2, col=mydata$samplotype, shape=my-
data$blank))
geom_point(size=6, alpha=0.5) +
labs(x = "nMDS1", y="nMDS2", title="Unweighted UniFrac distances") +
annotate("text", max(NMDS_unweighted_uni$MDS1)-0.05, min(NMDS_unweighted_uni$MDS2), label =
scale_color_manual(values=c('#a0bed8ff', '#E69F00', '#607281'))+
theme(panel.grid.major = element_blank(),
panel.grid.minor = element_blank(),

```

```

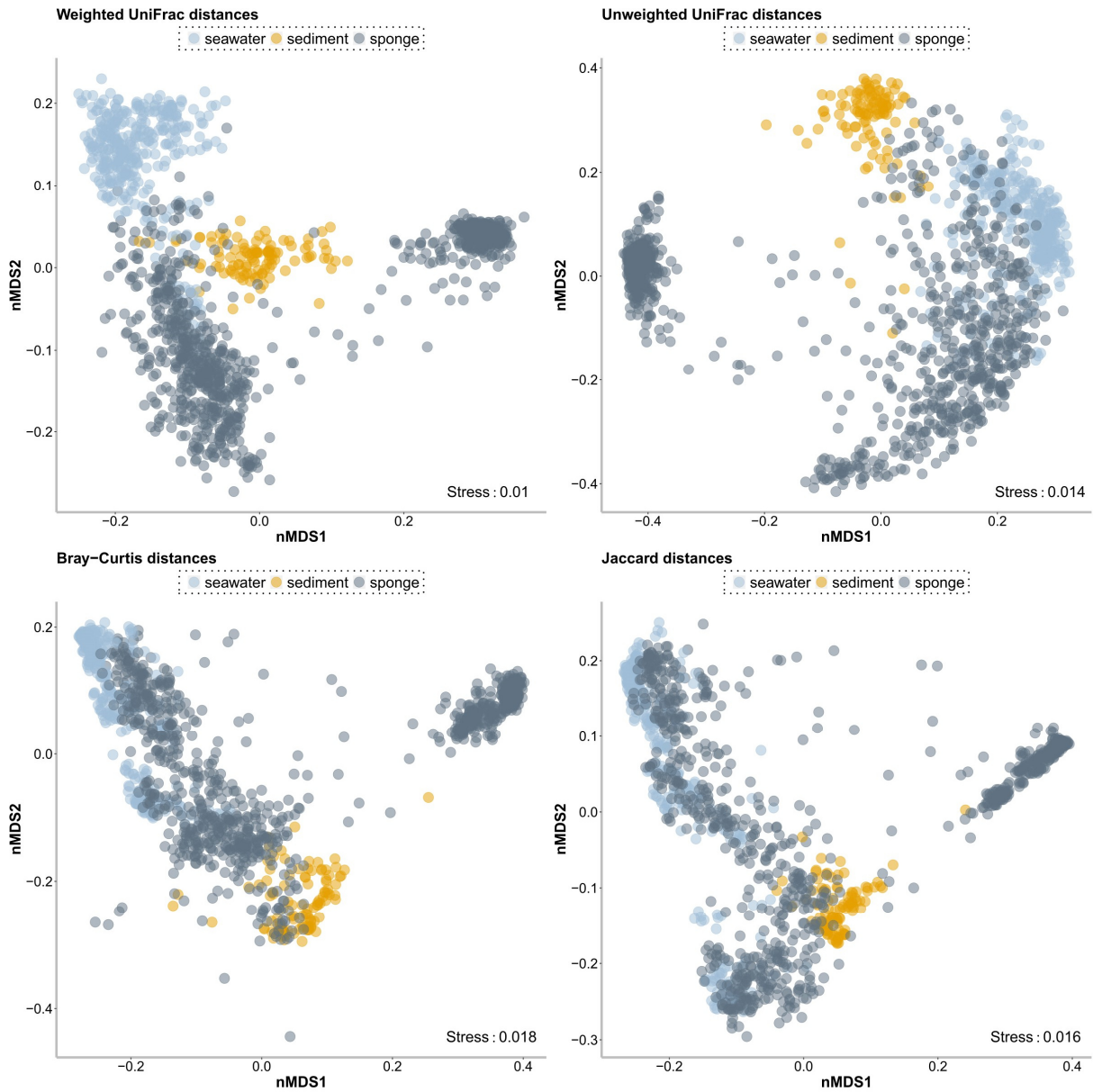
axis.line = element_line(colour = "grey", size = 1.5, linetype = "solid"),
panel.background = element_rect(fill = "#ffffff", colour = NA),
plot.background = element_rect(fill = '#ffffff'),
legend.position="top",
legend.title= element_blank(),
legend.text=element_text(size=20),
legend.background = element_rect(linetype = 'dotted', size = 1, colour = 1),
plot.title = element_text(color="black", size=20, face="bold"),
axis.title.x = element_text(color="black", size=20, face="bold"),
axis.text.x = element_text(face="plain", color="black",size=18, angle=0),
axis.title.y = element_text(color="black", size=20, face="bold"),
axis.text.y = element_text(face="plain", color="black",size=18, angle=0))
community_matrix = read.table(file = "bray_curtis_distance_matrix.txt", header = T)
mydata = read.table(file = "metadataforR.txt", header = T)
bray_curtis_NMDS <- metaMDS(community_matrix,k=50, maxit=100,trace=TRUE)
stress_NMDS_bray_curtis <-bray_curtis_NMDS$stress
stress_NMDS_bray_curtis <-round(stress_NMDS_bray_curtis, digits = 3)
NMDS_bray_curtis = data.frame(MDS1 = bray_curtis_NMDS$points[,1], MDS2 = bray_curtis_NMDS$points[,2])
p3 <-ggplot(data = NMDS_bray_curtis, aes(MDS1, MDS2, col=mydata$samplotype, shape=mydata$blank)) +
geom_point(size=6, alpha=0.5) +
labs(x = "nMDS1", y="nMDS2", title="Bray-Curtis distances") +
annotate("text", max(NMDS_bray_curtis$MDS1)-0.05, min(NMDS_bray_curtis$MDS2), label =
paste("Stres
scale_color_manual(values=c('#a0bed8ff', '#E69F00', '#607281')))+
theme(panel.grid.major = element_blank(),
panel.grid.minor = element_blank(),
axis.line = element_line(colour = "grey", size = 1.5, linetype = "solid"),
panel.background = element_rect(fill = "#ffffff", colour = NA),
plot.background = element_rect(fill = '#ffffff'),
legend.position="top",
legend.title= element_blank(),
legend.text=element_text(size=20),
legend.background = element_rect(linetype = 'dotted', size = 1, colour = 1),
plot.title = element_text(color="black", size=20, face="bold"),
axis.title.x = element_text(color="black", size=20, face="bold"),
axis.text.x = element_text(face="plain", color="black",size=18, angle=0),
axis.title.y = element_text(color="black", size=20, face="bold"),
axis.text.y = element_text(face="plain", color="black",size=18, angle=0))
community_matrix = read.table(file = "jaccard_distance_matrix.txt", header = T)
mydata = read.table(file = "metadataforR.txt", header = T)
jaccard_NMDS <- metaMDS(community_matrix,k=50, maxit=100,trace=TRUE)
stress_NMDS_jaccard <-jaccard_NMDS$stress
stress_NMDS_jaccard <-round(stress_NMDS_jaccard, digits = 3)
NMDS_jaccard = data.frame(MDS1 = jaccard_NMDS$points[,1], MDS2 = jaccard_NMDS$points[,2])
p4 <-ggplot(data = NMDS_jaccard, aes(MDS1, MDS2, col=mydata$samplotype, shape=mydata$blank)) +
geom_point(size=6, alpha=0.5) +
labs(x = "nMDS1", y="nMDS2", title="Jaccard distances") +
annotate("text", max(NMDS_jaccard$MDS1)-0.05, min(NMDS_jaccard$MDS2), label =
paste("Stress: ", st
scale_color_manual(values=c('#a0bed8ff', '#E69F00', '#607281')))+
theme(panel.grid.major = element_blank(),
panel.grid.minor = element_blank(),
axis.line = element_line(colour = "grey", size = 1.5, linetype = "solid"),
panel.background = element_rect(fill = "#ffffff", colour = NA),
plot.background = element_rect(fill = '#ffffff'),
plot.title = element_text(color="black", size=20, face="bold"),
legend.position="top",
legend.title= element_blank(),
legend.text=element_text(size=20),
legend.background = element_rect(linetype = 'dotted', size = 1, colour = 1),
axis.title.x = element_text(color="black", size=20, face="bold"),
axis.text.x = element_text(face="plain", color="black",size=18, angle=0),
axis.title.y = element_text(color="black", size=20, face="bold"),

```

```

axis.text.y = element_text(face="plain", color="black",size=18, angle=0))
pdf(file="beta_diversity.pdf", height= 20, width= 20)
grid.arrange(p1, p2 , p3, p4, nrow=2)
dev.off()
quit("no")

```

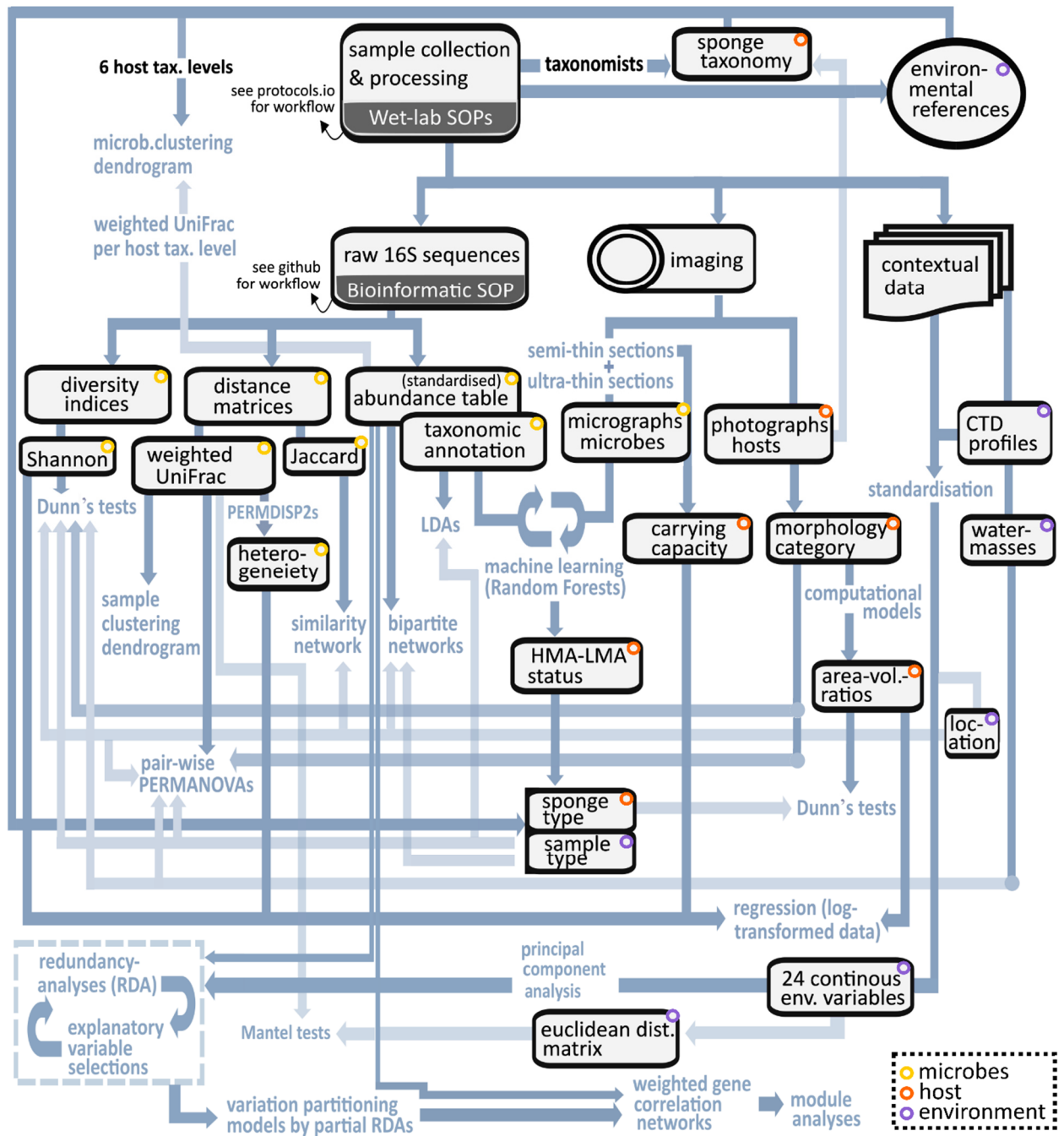


```

cd $PWDa/
mv $PWDa/silva138 $PWDa/dada2/silva138

```


Supplements Chapter 7



Method S1 Designed workflow for statistical analyses conducted in this study. More details about developed standard operation procedures (SOPs) have been deposited on protocols.io (wet-lab SOP) and github (bioinformatic processing of raw 16S sequences).

Table S1 Key data for the 21 conducted ship expeditions. Cruise leg is the official campaign name used in public repositories (PANGAEA).

	Cruise leg	Cruise platform	Departure date	Departure location	Return date	Return location
1	Hudson2016-019	CCGS Hudson	2016-07-19	Dartmouth (Canada)	2016-08-16	Dartmouth (Canada)
2	PS96 (ANT XXXI/2)	RV Polarstern	2015-12-06	Cape Town (South Africa)	2016-02-14	Punta Arenas (Chile)
3	SO254	RV Sonne	2017-01-27	Auckland (New Zealand)	2017-02-26	Auckland (New Zealand)
4	PS80 (ARK-XXVII/2)	RV Polarstern	2012-07-15	Longyearbyen (Spitsbergen)	2012-07-30	Tromsø (Norway)
5	PS107 (ARK-XXXI/2)	RV Polarstern	2017-07-23	Tromsø (Norway)	2017-08-19	Tromsø (Norway)
6	KB2017610	RV Kristine Bonnevie	2017-04-26	Bergen (Norway)	2017-05-02	Bergen (Norway)
7	MLB2017001	CCGS Martha L. Black	2017-08-31	Sydney (Canada)	2017-09-07	Dartmouth (Canada)
8	GS2016109A	RV G.O.Sars	2016-06-18	Tromsø (Norway)	2016-06-27	Tromsø (Norway)
9	GS2017110	RV G.O.Sars	2017-07-20	Bergen (Norway)	2017-08-06	Tromsø (Norway)
10	SPONGES 0617	B/O Ángeles Alvariño	2017-06-11	Gijón (Spain)	2017-06-25	Gijón (Spain)
11	HB2016952	RV Hans Brattstrøm	2016-09-08	Marineholmen (Norway)	2016-09-09	Marineholmen (Norway)
13	0915S	FRV Scotia	2015-07-16	Aberdeen (UK)	2015-07-27	Aberdeen (UK)
14	HB27102017	RV Hans Brattstrøm	2017-10-27	Bergen (Norway)	2017-10-27	Bergen (Norway)
15	JR17003a	RRS James Clark Ross	2018-02-18	Stanley (Falkland Islands)	2018-03-12	Stanley (Falkland Islands)
16	PAA2014007	GINR Paamiut	2014-09-22	Nuuk (Greenland)	2014-10-19	Nuuk (Greenland)
17	GS2018108	RV G.O.Sars	2018-07-28	Tromsø (Norway)	2018-08-14	Tromsø (Norway)
18	Azores2018	RV Ada Rebikoff	2018-07-04	Pico (Azores)	2018-07-11	Pico (Azores)
19	PS101 (ARK-XXX/3)	RV Polarstern	2016-09-09	Tromsø (Norway)	2016-10-23	Bremerhaven (Germany)
20	MSM86	RV Maria S. Merian	2019-08-14	Longyearbyen (Spitsbergen)	2019-09-17	Emden (Germany)
21	H045	fishing vessel	2018-05	NA	2018-05	NA

Table S2 Compilation of sampling instruments used in this study. Numbers of samples derived from each instrument, as well as frequencies of gear deployments are shown.

Gear	Number of samples	Number of gear deployments
Bottle, Niskin (ROV)	86	30
Bottle, Niskin (SHIP)	2	2
Bottle, Niskin (SUB)	6	3
Corer, Box (BC)	74	21
Corer, Giant Box (GBC)	18	9
Corer, Multi (MUC)	1	1
Corer, Push (PUC_ROV)	18	6
Corer, TV Multi (TVMUC)	2	2
CTD/Rosette (CTD-RO)	252	69
Dredge, anchor (DRG_A)	2	1
Dredge, Chain Bag (DCB)	6	1
Dredge, rock (DRG_R)	147	11
Dredge, triangular (DRG_T)	29	8
Grab (G)	3	1
Mooring (long time), acoustic (MOORY)	8	1
Remotely operated vehicle (ROV)	535	64
Sampling by diver (DIVER)	3	1
Sediment net (ROV)	12	4
Submersible, LULA1000 (SUB)	8	3
Television-Grab (TVG)	93	29
Trawl , Bottom (BT)	18	1
Trawl, Agassiz (AGT)	138	14
Trawl, Alfredo (AT)	18	15
Trawl, Beam (BEAM)	66	10
Trawl, Otter (OT)	1	1
	1546 = total Σ deep-sea samples	308 = total Σ deep-sea operations

Table S3 Taxonomic affiliations of the 169 deep-sea sponge species covered in this study. AphiaID is the official species identifier used in the World Register of Marine Species.

AphiaID	Host class	Host order	Host family	Host genus	Host species
132064	Demospongiae	Suberitida	Suberitidae	Aptos	<i>Aptos</i> sp
184559	Demospongiae	Poecilosclerida	Cladorhizidae	Abyssocladia	<i>Abyssocladia</i> sp1
184559	Demospongiae	Poecilosclerida	Cladorhizidae	Abyssocladia	<i>Abyssocladia</i> sp2
171966	Hexactinellida	Lyssacosida	Rossellidae	Acanthascus	<i>Acanthascus</i> sp
736912	Demospongiae	Axinellida	Raspailiidae	Acanthurypion	<i>Acanthurypion spinipinosum</i>
171850	Hexactinellida	Lyssacosida	Euplectellidae	Amphidiscella	<i>Amphidiscella caledonica</i>
1287788	Hexactinellida	Lyssacosida	Euplectellidae	Amphidiscella	<i>Amphidiscella sonnae</i>
171849	Hexactinellida	Lyssacosida	Euplectellidae	Amphidiscella	<i>Amphidiscella</i> sp
172012	Hexactinellida	Lyssacosida	Rossellidae	Anoxycalyx	<i>Anoxycalyx joubini</i>
134380	Hexactinellida	Sceptrulophora	Aphrocallistidae	Aphrocallistes	<i>Aphrocallistes beatrix</i>
131686	Hexactinellida	Sceptrulophora	Aphrocallistidae	Aphrocallistidae indet	<i>Aphrocallistidae</i> indet
236120	Demospongiae	Dendroceratida	Darwinellidae	Aplysilla	<i>Aplysilla sulfurea</i>
133172	Demospongiae	Poecilosclerida	Cladorhizidae	Asbestopluma	<i>Asbestopluma furcata</i>
235782	Demospongiae	Poecilosclerida	Cladorhizidae	Asbestopluma	<i>Asbestopluma</i> sp1
235782	Demospongiae	Poecilosclerida	Cladorhizidae	Asbestopluma	<i>Asbestopluma</i> sp2
235782	Demospongiae	Poecilosclerida	Cladorhizidae	Asbestopluma	<i>Asbestopluma</i> sp3
172017	Hexactinellida	Lyssacosida	Rossellidae	Asconema	<i>Asconema foliatum</i>
132122	Hexactinellida	Lyssacosida	Rossellidae	Asconema	<i>Asconema</i> indet
255149	Hexactinellida	Lyssacosida	Rossellidae	Asconema	<i>Asconema megaatrialia</i>
132122	Hexactinellida	Lyssacosida	Rossellidae	Asconema	<i>Asconema</i> sp
847773	Demospongiae	Astrophorida	Astrophorida indet	Astrophorida indet	<i>Astrophorida</i> indet
171610	Hexactinellida	Lyssacosida	Aulocalycidae	Aulocalyx	<i>Aulocalyx serialis</i>
171608	Hexactinellida	Lyssacosida	Aulocalycidae	Aulocalyx	<i>Aulocalyx</i> sp
132480	Demospongiae	Axinellida	Axinellidae	Axinella	<i>Axinella infundibuliformis</i>
131629	Demospongiae	Axinellida	Axinellidae	Axinellidae indet1	<i>Axinellidae</i> indet1
131629	Demospongiae	Axinellida	Axinellidae	Axinellidae indet2	<i>Axinellidae</i> indet2
131629	Demospongiae	Axinellida	Axinellidae	Axinellidae indet3	<i>Axinellidae</i> indet3
131629	Demospongiae	Axinellida	Axinellidae	Axinellidae indet4	<i>Axinellidae</i> indet4
131629	Demospongiae	Axinellida	Axinellidae	Axinellidae indet5	<i>Axinellidae</i> indet5
131629	Demospongiae	Axinellida	Axinellidae	Axinellidae indet6	<i>Axinellidae</i> indet6
131629	Demospongiae	Axinellida	Axinellidae	Axinellidae indet7	<i>Axinellidae</i> indet7
131629	Demospongiae	Axinellida	Axinellidae	Axinellidae indet8	<i>Axinellidae</i> indet8
132123	Hexactinellida	Lyssacosida	Rossellidae	Bathydorus	<i>Bathydorus</i> sp
131897	Demospongiae	Biemnida	Biemnidae	Biemna	<i>Biemna</i> sp
389041	Hexactinellida	Lyssacosida	Euplectellidae	Bolosoma	<i>Bolosoma cyanae</i>
171837	Hexactinellida	Lyssacosida	Euplectellidae	Bolosoma	<i>Bolosoma</i> indet
171837	Hexactinellida	Lyssacosida	Euplectellidae	Bolosoma	<i>Bolosoma</i> indet1
559	Calcarea	Calcarea indet1	Calcarea indet1	Calcarea indet1	<i>Calcarea</i> indet1
555794	Demospongiae	Tetractinellida	Calthropellidae	Calthropella	<i>Calthropella geodioides</i>
134398	Hexactinellida	Lyssacosida	Rossellidae	Caulophacus	<i>Caulophacus arcticus</i>
172053	Hexactinellida	Lyssacosida	Rossellidae	Caulophacus	<i>Caulophacus discohexactinus</i>
190321	Hexactinellida	Lyssacosida	Rossellidae	Caulophacus	<i>Caulophacus lotifolium</i>
132112	Hexactinellida	Lyssacosida	Rossellidae	Caulophacus	<i>Caulophacus</i> sp
170240	Demospongiae	Tetractinellida	Pachastrellidae	Characella	<i>Characella pachastrelloides</i>
171935	Hexactinellida	Lyssacosida	Leucopsacidae	Chaunoplectella	<i>Chaunoplectella</i> sp
817978	Demospongiae	Poecilosclerida	Cladorhizidae	Chondrocladia	<i>Chondrocladia robertballardi</i>
131894	Demospongiae	Poecilosclerida	Cladorhizidae	Chondrocladia	<i>Chondrocladia</i> sp
131832	Demospongiae	Haplosclerida	Chalinidae	Cladocroce	<i>Cladocroce</i> sp1
131832	Demospongiae	Haplosclerida	Chalinidae	Cladocroce	<i>Cladocroce</i> sp2
131895	Demospongiae	Poecilosclerida	Cladorhizidae	Cladorhiza	<i>Cladorhiza</i> sp
131868	Demospongiae	Poecilosclerida	Microcionidae	Clathria	<i>Clathria</i> sp
1051759	Calcarea	Clathrinida	Clathrinidae	Clathrina	<i>Clathrina pellucida</i>
168820	Demospongiae	Poecilosclerida	Coelosphaeridae	Coelosphaera	<i>Coelosphaera bullata</i>
1287715	Hexactinellida	Lyssacosida	Euplectellidae	Corbitella	<i>Corbitella plagiariorum</i>
171330	Demospongiae	Tetractinellida	Tetillidae	Craniella	<i>Craniella infrequens</i>

132093	Demospongiae	Tetractinellida	Tetillidae	Craniella	<i>Craniella</i> sp
171358	Demospongiae	Tetractinellida	Tetillidae	Craniella	<i>Craniella zetlandica</i>
410699	Demospongiae	Axinellida	Axinellidae	Cymbastela	<i>Cymbastela tricalyciformis</i>
131898	Demospongiae	Desmacellida	Desmacellidae	Desmacella	<i>Desmacella</i> sp
171858	Hexactinellida	Lyssacosida	Euplectellidae	Dictyaulus	<i>Dictyaulus</i> sp
132104	Hexactinellida	Sceptrulophora	Euretidae	Eurete	<i>Eurete</i> sp
577457	Hexactinellida	Sceptrulophora	Farreidae	Farrea	<i>Farrea ananchorata</i>
171750	Hexactinellida	Sceptrulophora	Farreidae	Farrea	<i>Farrea mexicana</i>
134391	Hexactinellida	Sceptrulophora	Farreidae	Farrea	<i>Farrea occa</i>
577453	Hexactinellida	Sceptrulophora	Farreidae	Farrea	<i>Farrea similaris</i>
132107	Hexactinellida	Sceptrulophora	Farreidae	Farrea	<i>Farrea</i> sp
134022	Demospongiae	Tetractinellida	Geodiidae	Geodia	<i>Geodia atlantica</i>
134023	Demospongiae	Tetractinellida	Geodiidae	Geodia	<i>Geodia barretti</i>
449128	Demospongiae	Tetractinellida	Geodiidae	Geodia	<i>Geodia hentscheli</i>
134033	Demospongiae	Tetractinellida	Geodiidae	Geodia	<i>Geodia macandrewii</i>
597246	Demospongiae	Tetractinellida	Geodiidae	Geodia	<i>Geodia pachydermata</i>
191325	Demospongiae	Tetractinellida	Geodiidae	Geodia	<i>Geodia parva</i>
254643	Demospongiae	Tetractinellida	Geodiidae	Geodia	<i>Geodia phlegraei</i>
132005	Demospongiae	Tetractinellida	Geodiidae	Geodia	<i>Geodia</i> sp
170192	Demospongiae	Tetractinellida	Geodiidae	Geodia	<i>Geodia vaubani</i>
170193	Demospongiae	Tetractinellida	Geodiidae	Geodia	<i>Geodia vestigifera</i>
131807	Demospongiae	Haplosclerida	Halichondriidae	Halichondria	<i>Halichondria</i> sp
842854	Demospongiae	Haplosclerida	Chalinidae	Haliclona	<i>Haliclona magna</i>
131834	Demospongiae	Haplosclerida	Chalinidae	Haliclona	<i>Haliclona</i> sp
131902	Demospongiae	Merliida	Hamacanthidae	Hamacantha	<i>Hamacantha</i> sp
131843	Demospongiae	Haplosclerida	Niphataidae	Hemigellius	<i>Hemigellius</i> sp
171080	Demospongiae	Tetractinellida	Corallistidae	Herengeria	<i>Herengeria</i> sp
577473	Hexactinellida	Sceptrulophora	Tretodictyidae	Hexactinella	<i>Hexactinella simplex</i>
22612	Hexactinellida	Hexactinellida sp nov	Hexactinellida sp nov	Hexactinellida sp nov	<i>Hexactinellida</i> sp nov
169682	Demospongiae	Verongiida	Ianthellidae	Hexadella	<i>Hexadella dedritifera</i>
169685	Demospongiae	Verongiida	Ianthellidae	Hexadella	<i>Hexadella pruvoti</i>
131777	Demospongiae	Suberitida	Suberitidae	Homaxinella	<i>Homaxinella</i> sp
171452	Hexactinellida	Amphidiscosida	Hyalonematidae	Hyalonema	<i>Hyalonema apertum</i>
132096	Hexactinellida	Amphidiscosida	Hyalonematidae	Hyalonema	<i>Hyalonema</i> sp
133621	Demospongiae	Poecilosclerida	Hymedesmiidae	Hymedesmia	<i>Hymedesmia paupertas</i>
131863	Demospongiae	Poecilosclerida	Acanthidae	Iophon	<i>Iophon</i> sp
131905	Demospongiae	Poecilosclerida	Isodictyidae	Isodictya	<i>Isodictya</i> sp
171949	Hexactinellida	Lyssacosida	Rosellidae	Lanuginellinae indet	<i>Lanuginellinae indet</i>
132038	Demospongiae	Poecilosclerida	Latrunculiidae	Latrunculia	<i>Latrunculia morrisoni</i>
389055	Hexactinellida	Lyssacosida	Leucopsacidae	Leucopsacus	<i>Leucopsacus distantus</i>
168914	Demospongiae	Poecilosclerida	Coelosphaeridae	Lissodendoryx	<i>Lissodendoryx bifacialis</i>
168950	Demospongiae	Poecilosclerida	Coelosphaeridae	Lissodendoryx	<i>Lissodendoryx complicata</i>
131930	Demospongiae	Poecilosclerida	Coelosphaeridae	Lissodendoryx	<i>Lissodendoryx</i> sp
131930	Demospongiae	Poecilosclerida	Coelosphaeridae	Lissodendoryx	<i>Lissodendoryx</i> sp2
171955	Hexactinellida	Lyssacosida	Rosellidae	Lophocalyx	<i>Lophocalyx</i> sp
168640	Demospongiae	Poecilosclerida	Mycalidae	Mycale	<i>Mycale lingua</i>
131658	Demospongiae	Poecilosclerida	Myxillidae	Myxillidae indet1	<i>Myxillidae</i> indet1
131658	Demospongiae	Poecilosclerida	Myxillidae	Myxillidae indet2	<i>Myxillidae</i> indet2
131658	Demospongiae	Poecilosclerida	Myxillidae	Myxillidae indet3	<i>Myxillidae</i> indet3
131658	Demospongiae	Poecilosclerida	Myxillidae	Myxillidae indet4	<i>Myxillidae</i> indet4
166798	Demospongiae	Haplosclerida	Petrosiidae	Neopetrosia	<i>Neopetrosia</i> sp
558215	Demospongiae	Tetractinellida	Corallistidae	Neoschrammeniella	<i>Neoschrammeniella bowerbankii</i>
596978	Demospongiae	Tetractinellida	Pachastrellidae	Pachastrella	<i>Pachastrella nodulosa</i>
134079	Demospongiae	Tetractinellida	Pachastrellidae	Pachastrella	<i>Pachastrella ovisternata</i>
131665	Demospongiae	Tetractinellida	Pachastrellidae	Pachastrellidae indet	<i>Pachastrellidae</i> indet
132036	Demospongiae	Axinellida	Stelligeridae	Paratimea	<i>Paratimea</i> sp
1363689	Demospongiae	Tetractinellida	Geodiidae	Penares	<i>Penares turmericolor</i>
166828	Demospongiae	Haplosclerida	Petrosiidae	Petrosia	<i>Petrosia crassa</i>
132506	Demospongiae	Axinellida	Axinellidae	Phakellia	<i>Phakellia bowerbanki</i>
234410	Demospongiae	Axinellida	Axinellidae	Phakellia	<i>Phakellia hironellei</i>

132509	Demospongiae	Axinellida	Axinellidae	Phakellia	<i>Phakellia robusta</i>
132510	Demospongiae	Axinellida	Axinellidae	Phakellia	<i>Phakellia rugosa</i>
132511	Demospongiae	Axinellida	Axinellidae	Phakellia	<i>Phakellia ventilabrum</i>
182738	Demospongiae	Axinellida	Axinellidae	Phakellia	<i>Phakellia vermiculata</i>
134378	Hexactinellida	Amphidiscosida	Phoronematidae	Phoronema	<i>Phoronema carpenteri</i>
171560	Hexactinellida	Amphidiscosida	Phoronematidae	Phoronema	<i>Phoronema conicum</i>
131953	Demospongiae	Poecilosclerida	Hymedesmiidae	Phorbos	<i>Phorbos</i> sp
133938	Homoscleromorpha	Homosclerophorida	Plakinidae	Plakortis	<i>Plakortis simplex</i>
171169	Demospongiae	Tetractinellida	Pleromidae	Pleroma	<i>Pleroma menoui</i>
171167	Demospongiae	Tetractinellida	Pleromidae	Pleroma	<i>Pleroma</i> sp
171171	Demospongiae	Tetractinellida	Pleromidae	Pleroma	<i>Pleroma turbinatum</i>
134083	Demospongiae	Tetractinellida	Vulcanellidae	Poecillastra	<i>Poecillastra compressa</i>
170263	Demospongiae	Tetractinellida	Vulcanellidae	Poecillastra	<i>Poecillastra laminaris</i>
132016	Demospongiae	Tetractinellida	Vulcanellidae	Poecillastra	<i>Poecillastra</i> sp
171579	Hexactinellida	Amphidiscosida	Phoronematidae	Poliopogon	<i>Poliopogon</i> sp
170639	Demospongiae	Polymastiida	Polymastiidae	Polymastia	<i>Polymastia invaginata</i>
132046	Demospongiae	Polymastiida	Polymastiidae	Polymastia	<i>Polymastia</i> sp
169618	Demospongiae	Verongiida	Aplysineidae	Porphyria	<i>Porphyria</i> sp
171875	Hexactinellida	Lyssacosida	Euplectellidae	Regadrella	<i>Regadrella okinoseana</i>
1287775	Hexactinellida	Lyssacosida	Euplectellidae	Regadrella	<i>Regadrella pedunculata</i>
132101	Hexactinellida	Lyssacosida	Aulocalycidae	Rhabdodictyum	<i>Rhabdodictyum</i> sp
132071	Demospongiae	Poecilosclerida	Suberitidae	Rhizaxinella	<i>Rhizaxinella</i> sp
172090	Hexactinellida	Lyssacosida	Rossellidae	Rossella	<i>Rossella antarctica</i>
172093	Hexactinellida	Lyssacosida	Rossellidae	Rossella	<i>Rossella fibulata</i>
172109	Hexactinellida	Lyssacosida	Rossellidae	Rossella	<i>Rossella vanhoeffeni</i>
171965	Hexactinellida	Lyssacosida	Rossellidae	Rossellinae indet	<i>Rossellinae</i> indet
171844	Hexactinellida	Lyssacosida	Euplectellidae	Saccocalyx	<i>Saccocalyx</i> sp
1287781	Hexactinellida	Lyssacosida	Euplectellidae	Saccocalyx	<i>Saccocalyx tetractinus</i>
134418	Hexactinellida	Lyssacosida	Rossellidae	Schaudinna	<i>Schaudinna rosea</i>
387351	Hexactinellida	Lyssacosida	Rossellidae	Scyphidium	<i>Scyphidium australiense</i>
134419	Hexactinellida	Lyssacosida	Rossellidae	Scyphidium	<i>Scyphidium septentrionale</i>
132051	Demospongiae	Polymastiida	Polymastiidae	Sphaerotylus	<i>Sphaerotylus</i> sp
131818	Demospongiae	Suberitida	Halichondriidae	Spongosorites	<i>Spongosorites</i> sp
133980	Demospongiae	Tetractinellida	Ancorinidae	Stelletta	<i>Stelletta normani</i>
133983	Demospongiae	Tetractinellida	Ancorinidae	Stelletta	<i>Stelletta raphidiophora</i>
133988	Demospongiae	Tetractinellida	Ancorinidae	Stryphnus	<i>Stryphnus fortis</i>
134240	Demospongiae	Suberitida	Stylocordylidae	Stylocordyla	<i>Stylocordyla borealis</i>
132113	Hexactinellida	Lyssacosida	Rossellidae	Sympagella	<i>Sympagella</i> sp
NA	Demospongiae	TAP	TAP	TAP	TAP
131974	Demospongiae	Poecilosclerida	Tedaniidae	Tedania	<i>Tedania</i> sp
131974	Demospongiae	Poecilosclerida	Tedaniidae	Tedania	<i>Tedania</i> sp2
170689	Demospongiae	Polymastiida	Polymastiidae	Tentorium	<i>Tentorium papillatum</i>
132053	Demospongiae	Polymastiida	Polymastiidae	Tentorium	<i>Tentorium</i> sp
132053	Demospongiae	Polymastiida	Polymastiidae	Tentorium	<i>Tentorium</i> sp2
134107	Demospongiae	Tetractinellida	Theneidae	Thenea	<i>Thenea schmidti</i>
132019	Demospongiae	Tetractinellida	Theneidae	Thenea	<i>Thenea</i> sp
134108	Demospongiae	Tetractinellida	Theneidae	Thenea	<i>Thenea valdiviae</i>
134113	Demospongiae	Chondrillida	Chondrillidae	Thymosia	<i>Thymosia guernei</i>
134420	Hexactinellida	Lyssacosida	Rossellidae	Trichasterina	<i>Trichasterina borealis</i>
172121	Hexactinellida	Lyssacosida	Rossellidae	Vazella	<i>Vazella pourtalesii</i>
597279	Demospongiae	Tetractinellida	Vulcanellidae	Vulcanella	<i>Vulcanella gracilis</i>
170325	Demospongiae	Tetractinellida	Vulcanellidae	Vulcanella	<i>Vulcanella</i> sp
171885	Hexactinellida	Lyssacosida	Euplectellidae	Walteria	<i>Walteria leuckarti</i>
134232	Demospongiae	Polymastiida	Polymastiidae	Weberella	<i>Weberella bursa</i>

Table S4 Results of statistical testing (Dunn's tests) conducted to assess variations in the area:volume ratio between different sponge types.

Group1	Group2	Sample number	Dunns Z	p-value
HMA	LMA_demo	670	-16.43	<0.001
HMA	LMA_glass	602	-17.70	<0.001
LMA_demo	LMA_glass	554	7.52	<0.001

Table S5 Results of statistical testing (PERMANOVAs based on weighted UniFrac distances) conducted to assess differences in the microbial community composition between sampletypes.

Group1	Group2	Sample number	Permuta- tions	pseudo F	p-value
seawater	sponge	1286	999	257.6	0.001
sediment	sponge	1039	999	62.1	0.001
seawater	sediment	463	999	205.5	0.001
HMA	LMA_demo	670	999	668.5	0.001
HMA	LMA_glass	602	999	515.9	0.001
LMA_demo	LMA_glass	554	999	65.8	0.001

Table S6 Results of statistical testing (Dunn's tests) conducted to assess variations in microbial alpha diversity (Shannon index) between sampletypes.

Group1	Group2	Sample number	Dunns Z	p-value
seawater	sponge	1286	21.49	<0.001
sediment	sponge	1039	16.75	<0.001
seawater	sediment	463	-14.19	<0.001
HMA	LMA_demo	670	21.82	<0.001
HMA	LMA_glass	602	18.42	<0.001
LMA_demo	LMA_glass	554	-1.98	0.048

Table S7 Fractions of amplicon single nucleotide variants (ASVs) shared between different sampletypes. Values are percentages [%] calculated based on the total number of ASVs present in the two compared groups.

HMA	LMA_demo	LMA_glass	seawater	sediment	
-	4.3	2.8	1.8	1.4	HMA
	-	12.4	14.6	5.4	LMA_demo
		-	16.2	12.1	LMA_glass
			-	7.2	seawater
				-	sediment

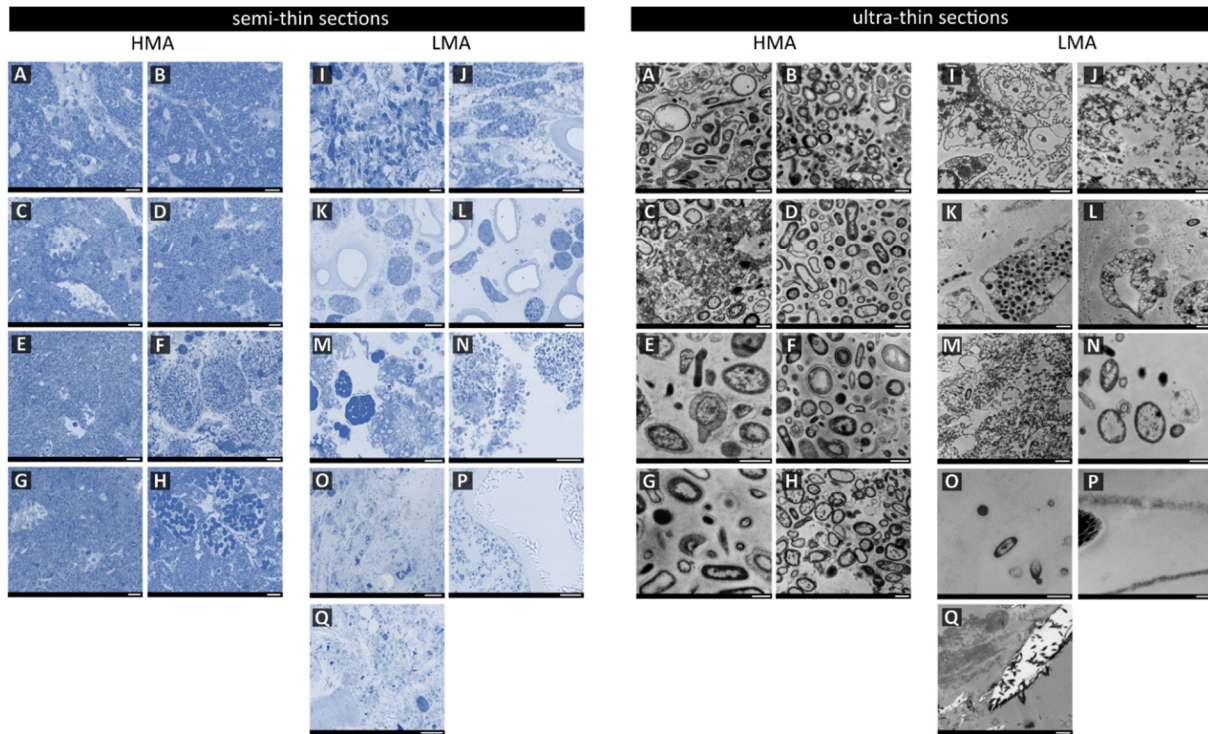


Figure S1 Micrographs showing microbial cells inside the tissue of 17 sponge species. Left panel shows semi-thin sections (blue; all scale bars= 10 μm), while right panel shows ultra-thin sections visualised by transmission electron microscopy (grey; all scale bars= 1 μm). Within each panel High Microbial Abundance (HMA) sponges are shown left (A-H) and Low Microbial Abundance (LMA) sponges are shown right (I-Q). **A)** *Geodia hentscheli*. **B)** *Geodia parva*. **C)** *Plakortis simplex*. **D)** *Geodia phlegraei*. **E)** *Craniella infrequens*. **F)** *Petrosia crassa*. **G)** *Stelletta raphidiophora*. **H)** *Stryphnus fortis*. **I)** *Acanthephyron spinispinosum*. **J)** *Axinella infundibuliformis*. **K)** *Phakellia hirondellei*. **L)** *Phakellia ventilabrum*. **M)** *Schaudinnia rosea*. **N)** *Vazella pourtalesii*. **O)** *Thenea schmidtii*. **P)** *Aphrocallistes beatrix*. **Q)** *Pheronema carpenteri*.

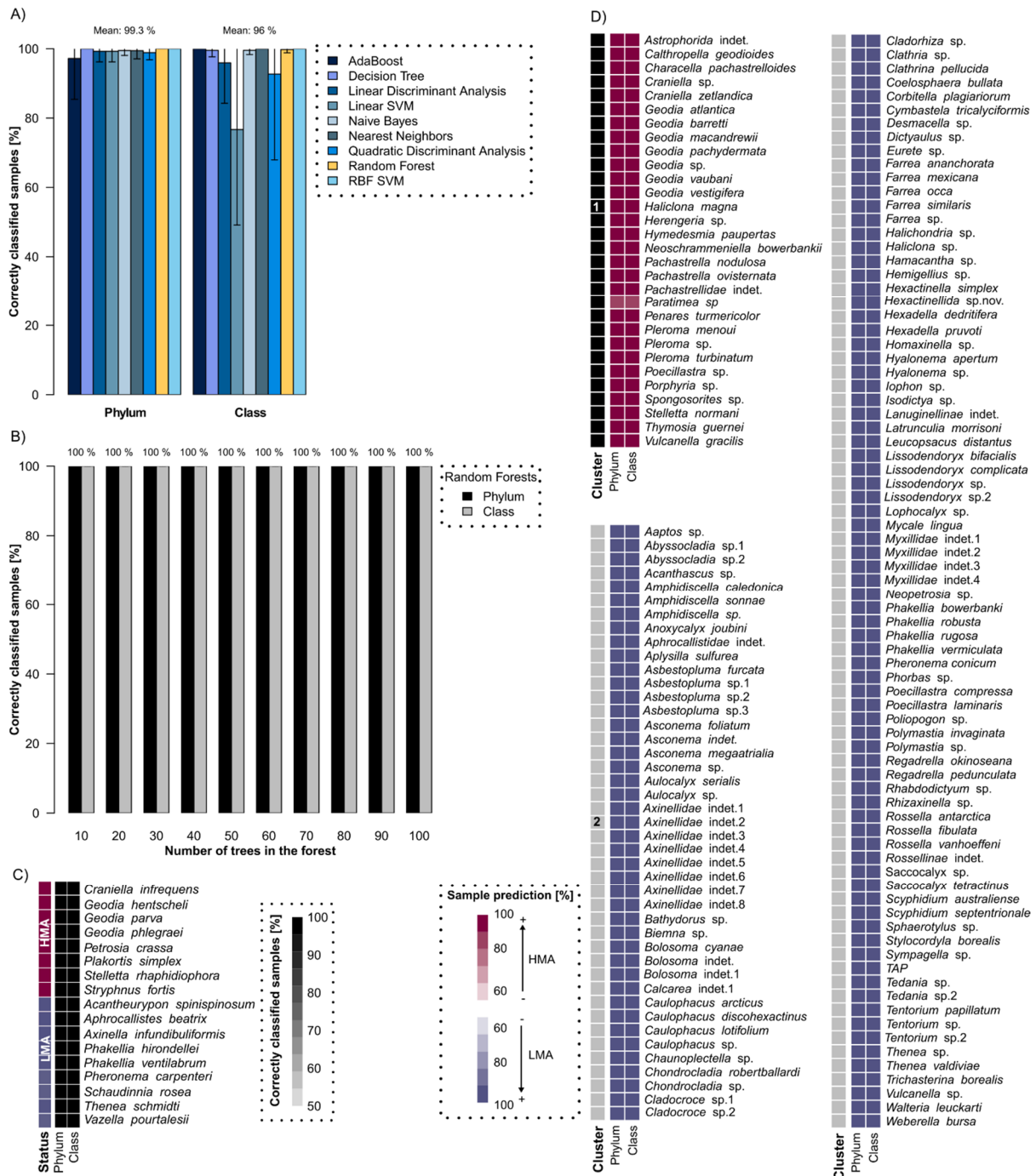


Figure S2 Visual outputs of machine learning approach to predict the HMA-LMA status of sponge species following Moitinho-Silva *et al.* (2017). A-C) Selection and standadisation of classifiers: **A)** comparing the performance of different algorithms and taxonomic levels on the training dataset (i.e. sponge species with HMA-LMA status determined by microscopy). Error bars stand for weighted standard deviations. **B)** Performance of Random Forests in relation to number of trees in the forest. Percentages above the bars represent means of weighted averages. **C)** Performance of Random Forests algorithm in predicting the HMA-LMA status of sponge species with known status. **D)** Predictions of the HMA-LMA status of previously uncharacterised sponge species by the Random Forests algorithm. The applied number of trees in the forest was 50 in C) and D).

Table S8 List of computational models for calculation of volumes for each sponge morphological category. Some equations require usage of constants (check **Table S9**): c_{tc} =species-specific constant of cone-thickness, c_{te} =species-specific constant of encrustation-thickness, c_{tf} =species-specific constant of fan-thickness, c_{tr} =species-specific constant of rectangle-thickness, c_{tt} =species-specific constant of tube-thickness, c_{nt} = species-specific constant for number of tubes per individual, c_{nr} = species-specific constant for number of rectangles per individual.


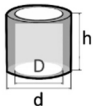
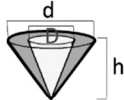

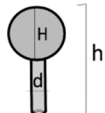

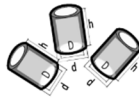
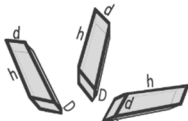
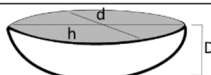
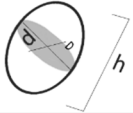
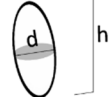

category name	geometry	pictogram	computational model /equation
A ball	sphere		$V = \frac{\pi}{6} h^3$
B vase	hollow cylinder		$V = \frac{\pi h}{4} (d^2 - D^2)$ with $D = d - c_{tt}$ where $c_{tt} > 0.01$
C flabellate: cone	hollow cone		$V = \frac{\pi h}{3} \left(\left(\frac{d}{2}\right)^2 - \left(\frac{D}{2}\right)^2 \right)$ with $D = d - c_{tc}$ where $0.01 < c_{tc} < 0.5$
D flabellate: fan	rectangular pyramid		$V = \frac{1}{3} d D h$ with $D = c_{tr}$ where $0.01 < c_{tr} < 0.5$
E stalked	sphere + cylinder		$V = \left(\frac{\pi}{6} H\right) + \left(\frac{\pi}{4} d^2 (h - H)\right)$
F stick	cylinder		$V = \frac{\pi}{4} d^2 h$
G bushy (tubes)	multiple hollow cylinders		$V = c_{nt} \left(\frac{\pi h}{4} (d^2 - D^2)\right)$ with $c_{nt} > 0$ and $D = d - c_{tt}$ where $0.01 < c_{tt} < 1$
H branching (no tubes)	multiple rectangles		$V = c_{nr} (hdD)$ with $c_{nr} > 0$ and $D = c_{tr}$ where $0.01 < c_{tr} < 1$
I bowl	half ellipsoid		$V = \frac{\pi}{12} d D h$
J encrusting	ellipsoid		$V = \frac{\pi}{6} d D h$ with $D = c_{te}$ where $0.01 < c_{te} < 0.5$
K other: compact	prolate spheroid		$V = \frac{\pi}{6} d^2 h$
L other: fluffy	prolate spheroid		$V = \frac{\pi}{6} d^2 h$

Table S9 Overview of species-specific constants required in computational models for calculation of sponge volumes. ctc=species-specific constant of cone-thickness, cte=species-specific constant of encrustation-thickness, ctf=species-specific constant of fan-thickness, ctr=species-specific constant of rectangle-thickness, ctt=species-specific constant of tube-thickness, cnt = species-specific constant for number of tubes per individual, cnr = species-specific constant for number of rectangles per individual.

Host species	ctc [cm]	cte [cm]	ctf [cm]	ctr [cm]	ctt [cm]	cnr	cnt
<i>Acanthascus</i> sp.	-	-	-	-	0.5	-	-
<i>Acantheurypon spinispinosum</i>	-	0.5	-	-	-	-	-
<i>Anoxycalyx joubini</i>	-	-	-	-	1.5	-	-
<i>Aphrocallistes beatrix</i>	0.5	-	-	-	-	-	-
<i>Aplysilla sulfurea</i>	-	0.5	-	-	-	-	-
<i>Asconema foliatum</i>	-	-	-	-	0.5	-	15
<i>Asconema megaatrialia</i>	-	-	-	-	1	-	-
<i>Aulocalyx</i> sp.	-	-	-	-	0.5	-	21
<i>Axinella infundibuliformis</i>	-	-	0.5	-	-	-	-
<i>Bathydorus</i> sp.	-	-	-	-	1	-	-
<i>Cladocroce</i> sp.1	0.5	-	-	-	-	-	-
<i>Cladorhiza</i> sp.	-	-	-	0.5	-	49	-
<i>Cymbastela tricalyciformis</i>	0.5	-	-	-	-	-	-
<i>Desmacella</i> sp.	-	-	0.5	-	-	-	-
<i>Dictyaulus</i> sp.	-	-	-	-	1	-	-
<i>Eurete</i> sp.	-	-	-	-	0.5	-	33
<i>Farrea</i> sp.	-	-	-	-	0.5	-	22
<i>Hamacantha</i> sp.	-	-	-	1	-	12	-
<i>Hexadella dedritifera</i>	-	0.5	-	-	-	-	-
<i>Hexadella pruvoti</i>	-	0.5	-	-	-	-	-
<i>Homaxinella</i> sp.	-	-	0.5	-	-	-	-
<i>Lissodendoryx complicata</i>	-	-	-	0.5	-	23	-
<i>Lophocalyx</i> sp.	-	-	-	-	1	-	-
<i>Neopetrosia</i> sp.	-	-	-	0.5	-	28	-
<i>Phakellia hironellei</i>	-	-	0.5	-	-	-	-
<i>Phakellia robusta</i>	-	-	0.5	-	-	-	-
<i>Phakellia ventilabrum</i>	-	-	0.5	-	-	-	-
<i>Phakellia vermiculata</i>	-	-	0.5	-	-	-	-
<i>Pheronema carpenteri</i>	-	-	-	-	1.5	-	-
<i>Pheronema conicum</i>	-	-	-	-	1.5	-	-
<i>Poecillastra laminaris</i>	0.5	-	-	-	-	-	-
<i>Poliopogon</i> sp.	-	-	-	-	1	-	-
<i>Regadrella okinoseana</i>	-	-	-	-	1	-	-
<i>Rossella antarctica</i>	-	-	-	-	1.5	-	-
<i>Rossella vanhoeffeni</i>	-	-	-	-	1.5	-	-
<i>Rossellinae</i> indet.	-	-	-	-	1	-	-
<i>Schaudinnia rosea</i>	-	-	-	-	1	-	-
<i>Scyphidium australiense</i>	-	-	-	-	1	-	-
TAP	-	-	0.5	-	-	-	-
<i>Trichasterina borealis</i>	-	-	-	-	1	-	4
<i>Vazella pourtalesii</i>	-	-	-	-	1	-	-

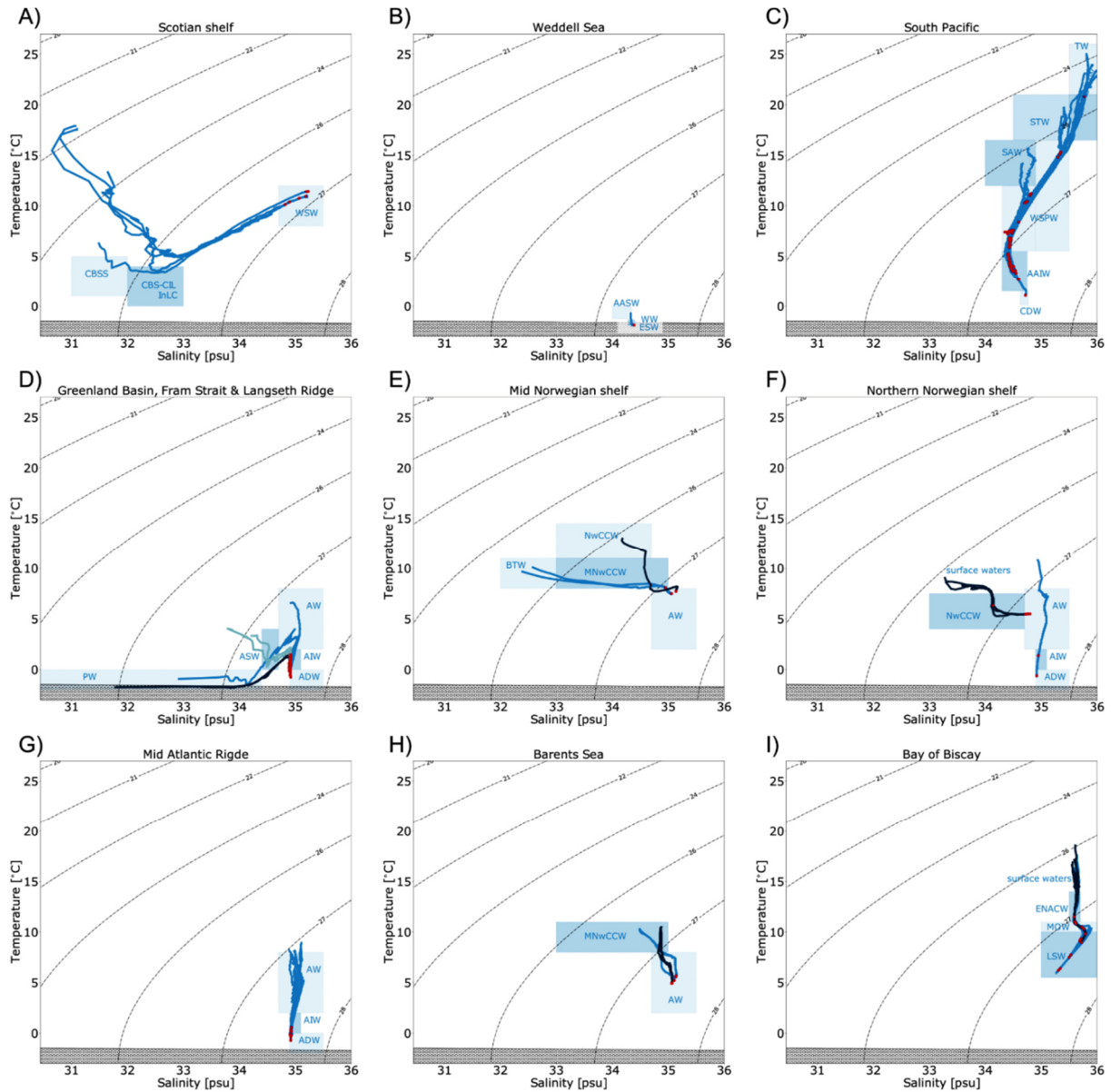


Figure S3 Temperature - salinity (TS) profiles derived from 66 full water conductivity-temperature-depth (CTD) casts in different ocean regions. Presented records start 20 m below the ocean surface and reach down to 5 m above the ocean floor. Blue boxes indicate characteristic T-S - ranges of prevailing watermasses. These watermasses were classified manually based on literature (see methods **Chapter 7**). The full names of all watermass abbreviations are listed in **Table S10**. Red points indicate sampling depths of sponges. Dashed lines represent density lines (σ_{θ} = [kg/m³]) and the lines which are filled with circles at the bottom of each graph, represent freezing point lines.

Table S10 List of watermass acronyms and respective full names.

Acronym	Watermass
AAIW	Antarctic Intermediate Water
AASW	Antarctic Surface Water
ADW	Arctic Deep Water
AIW	Arctic Intermediate Water
ASW	Arctic Surface Water
AW	Atlantic Water
BTW	Brackish Top Water
CBS-CIL	Cabot Strait Surface - Cold Intermediate Layer
CBSS	Cabot Strait Sub-Surface Water
CDW	Circumpolar Deep Water
ENACW	Eastern North Atlantic Central Water
ESW	Eastern Shelf Water
InLC	Inshore Labrador Current
LSW	Labrador Sea Water
MNwCCW	Modified Norwegian Coastal Current Water
MOW	Mediterranean Outflow Water
NwCCW	Norwegian Coastal Current Water
PW	Polar Water
SAW	Sub-Antarctic Water
SPSTMW	South Pacific Sub-Tropical Mode Water
STW	Sub-Tropical Water
TW	Tropical surface Water
WSPW	Western South Pacific Water
WSW	Warm Slope Water
WW	Winter water

Table S11 Abbreviations of the 24 continuous environmental parameters and their respective full names and units. Method refers to the approach with which the data were gathered. *In situ* means that the respective data entries were measured during the 20 conducted cruises. WOA, MODIS, and GLODAP indicate that entries were retrieved from publicly available climatologies. In particular the World Ocean Atlas (WOA; version WOA18), the Global Ocean Data Analysis Project (Glodap; v2 2020), and satellite data (MODIS; see methods **Chapter 7**) were used.

Abbreviation	Full parameter name [unit]	Method
AOU	Apparent oxygen utilization at sampling depth [mL L ⁻¹]	WOA
Chla	Chlorophyll a concentration at ocean surface [mg m ⁻³]	MODIS
cond	Conductivity of seawater at sampling depth [mS cm ⁻¹]	<i>in situ</i>
dens	Density of seawater at sampling depth [kg m ⁻³]	<i>in situ</i>
depth	Sampling depth [m]	<i>in situ</i>
DOC	Dissolved organic carbon concentration at sampling depth [μmol kg ⁻¹]	GLODAP
lat	Sampling position, latitude [dd]	<i>in situ</i>
lon	Sampling position, longitude [dd]	<i>in situ</i>
MLD	Distance from sampling depth to mixed layer depth [m]	calculated
N	Nitrate concentration at sampling depth [μmol L ⁻¹]	WOA
N:P	N:P ratio at sampling depth	calculated
O2	Oxygen concentration at sampling depth [mL L ⁻¹]	WOA
O2sat	Oxygen saturation at sampling depth [%]	WOA
P	Phosphate concentration at sampling depth [μmol L ⁻¹]	WOA
pH	pH at sampling depth	GLODAP
PIC	Particulate inorganic carbon concentration at ocean surface [mol m ⁻³]	MODIS
POC	Particulate organic carbon concentration at ocean surface [mg m ⁻³]	MODIS
pres	Pressure at sampling depth [dbar]	<i>in situ</i>
S	Salinity at sampling depth [psu]	<i>in situ</i>
Si	Silicate concentration at sampling depth [μmol L ⁻¹]	WOA
Si:P	Si:P ratio at sampling depth	calculated
T	Temperature at sampling depth [°C]	<i>in situ</i>
tAlk	Total alkalinity of seawater at sampling depth [μmol kg ⁻¹]	GLODAP
tCO2	Total dissolved carbon in seawater at sampling depth [μmol kg ⁻¹]	GLODAP

Table S18 Results of statistical testing (Mantel tests) conducted to assess correlations between environmental parameters (euclidean distances) and microbial community composition (weighted UniFrac distances) for the three sponge types. Group indicates to which category in the variation partitioning model each of the listed environmental parameters belongs. Parameters belonging to the same group are highly correlated with each other. Only those parameters which turned out to be the most relevant ones in the final variation partitioning models are shown.

Env. parameter	Microbiome	Group	Permutations	Mantel statistic r	p-value
AOU	HMA	nutrients & oxygen	999	0.4	0.001
AOU	LMA_demo	nutrients & oxygen	999	0.1	0.001
AOU	LMA_glass	nutrients & oxygen	999	0.3	0.001
N	HMA	nutrients & oxygen	999	0.4	0.001
N	LMA_demo	nutrients & oxygen	999	0.2	0.001
N	LMA_glass	nutrients & oxygen	999	0.4	0.001
O2	HMA	nutrients & oxygen	999	0.5	0.001
O2	LMA_demo	nutrients & oxygen	999	0.2	0.001
O2	LMA_glass	nutrients & oxygen	999	0.3	0.001
O2sat	HMA	nutrients & oxygen	999	0.4	0.001
O2sat	LMA_demo	nutrients & oxygen	999	0.1	0.001
O2sat	LMA_glass	nutrients & oxygen	999	0.3	0.001
P	HMA	nutrients & oxygen	999	0.4	0.001
P	LMA_demo	nutrients & oxygen	999	0.2	0.001
P	LMA_glass	nutrients & oxygen	999	0.4	0.001
Si	HMA	nutrients & oxygen	999	0.4	0.001
Si	LMA_demo	nutrients & oxygen	999	0.1	0.001
Si	LMA_glass	nutrients & oxygen	999	0.4	0.001
Si:P	HMA	nutrients & oxygen	999	0.4	0.001
Si:P	LMA_demo	nutrients & oxygen	999	0.2	0.001
Si:P	LMA_glass	nutrients & oxygen	999	0.5	0.001
dens	HMA	depth-related	999	0.2	0.001
dens	LMA_demo	depth-related	999	0.1	0.001
dens	LMA_glass	depth-related	999	0.2	0.001
depth	HMA	depth-related	999	0.2	0.001
depth	LMA_demo	depth-related	999	0.2	0.001
depth	LMA_glass	depth-related	999	0.2	0.001
MLD	HMA	depth-related	999	0.2	0.001
MLD	LMA_demo	depth-related	999	0.1	0.010
MLD	LMA_glass	depth-related	999	0.1	0.001
pres	HMA	depth-related	999	0.2	0.001
pres	LMA_demo	depth-related	999	0.2	0.001
pres	LMA_glass	depth-related	999	0.2	0.001
cond	HMA	temp & conductivity	999	0.4	0.001
cond	LMA_demo	temp & conductivity	999	0.1	0.002
cond	LMA_glass	temp & conductivity	999	0.3	0.001

T	HMA	temp & conductivity	999	0.5	0.001
T	LMA_demo	temp & conductivity	999	0.1	0.011
T	LMA_glass	temp & conductivity	999	0.2	0.001
S	HMA	salinity	999	0.4	0.001
S	LMA_demo	salinity	999	0.2	0.001
S	LMA_glass	salinity	999	0.3	0.001

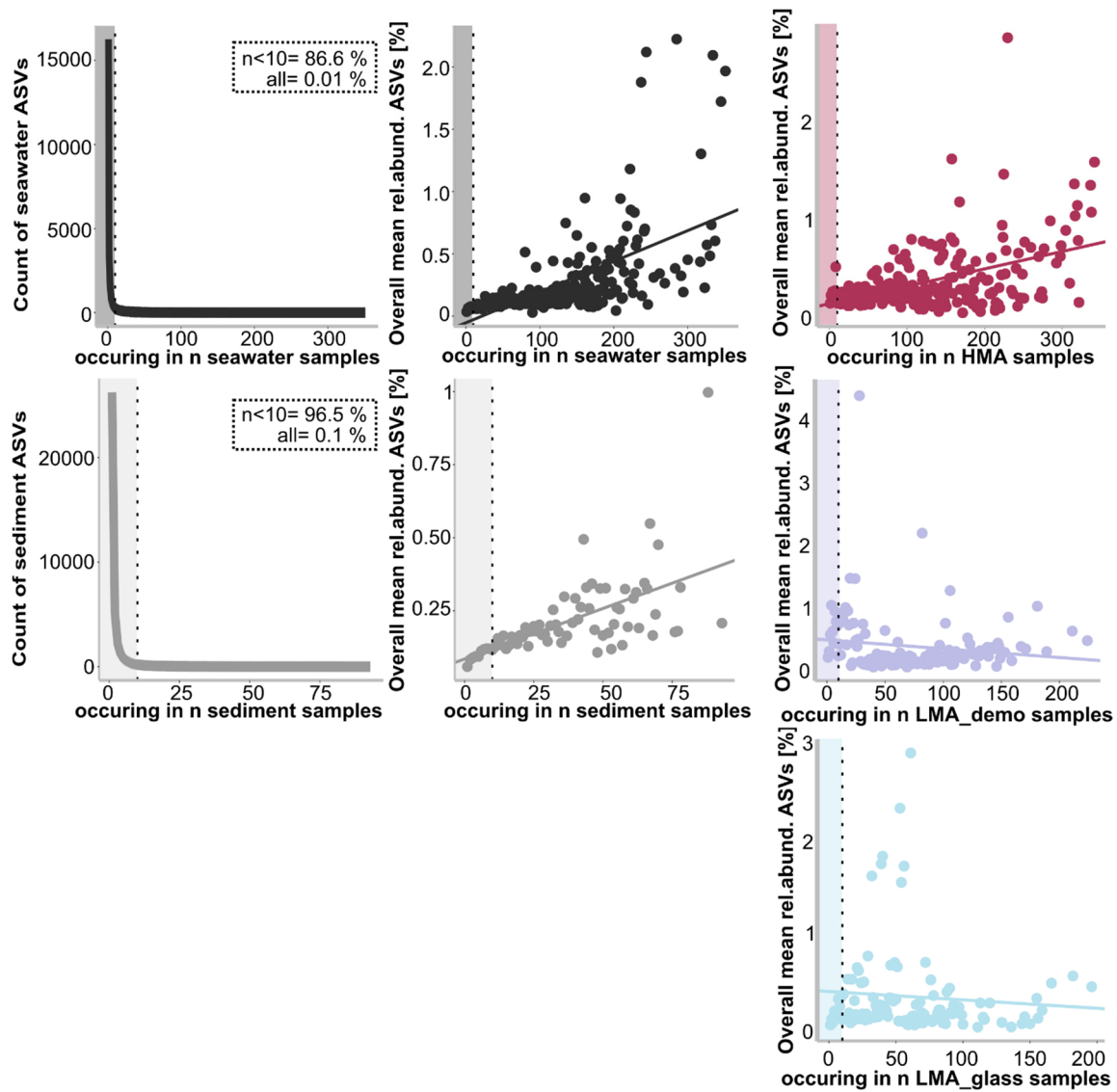


Figure S4 Distribution of ASVs across samples for environmental references (seawater and sediment). In the middle panel and in the right panel, overall mean relative abundances of ASVs across samples is shown.

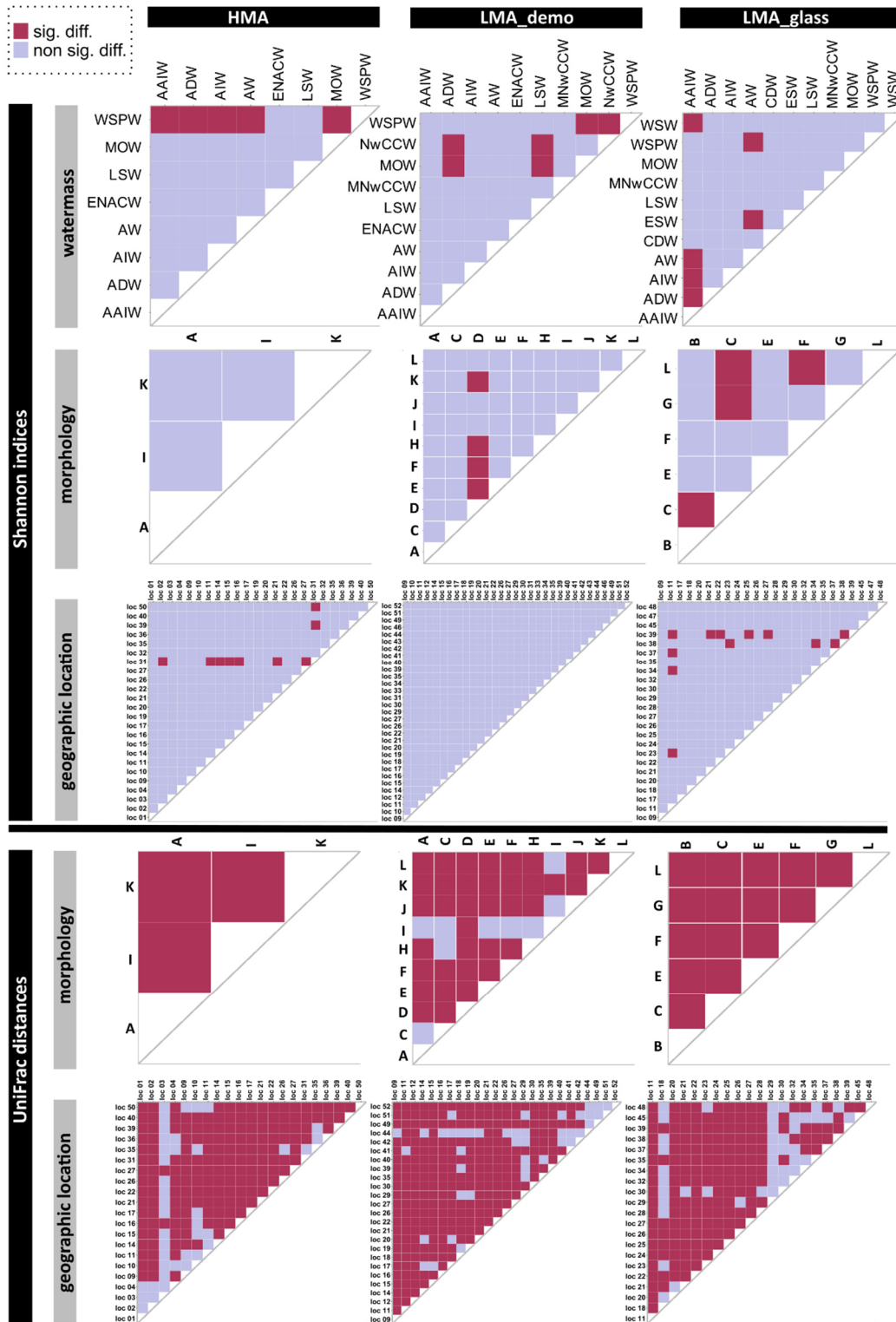


Figure S5 Visual representation of results of statistical testing (Dunn's tests) conducted to assess variations in microbial alpha diversity (Shannon index) between watermasses, between host morphology categories, and between geographic locations (top part of plot). In the lower part of the plot results of statistical testing (PERMANOVAs based on weighted UniFrac distances) conducted to assess differences in the microbial community composition between host morphology categories, and between geographic locations are shown.

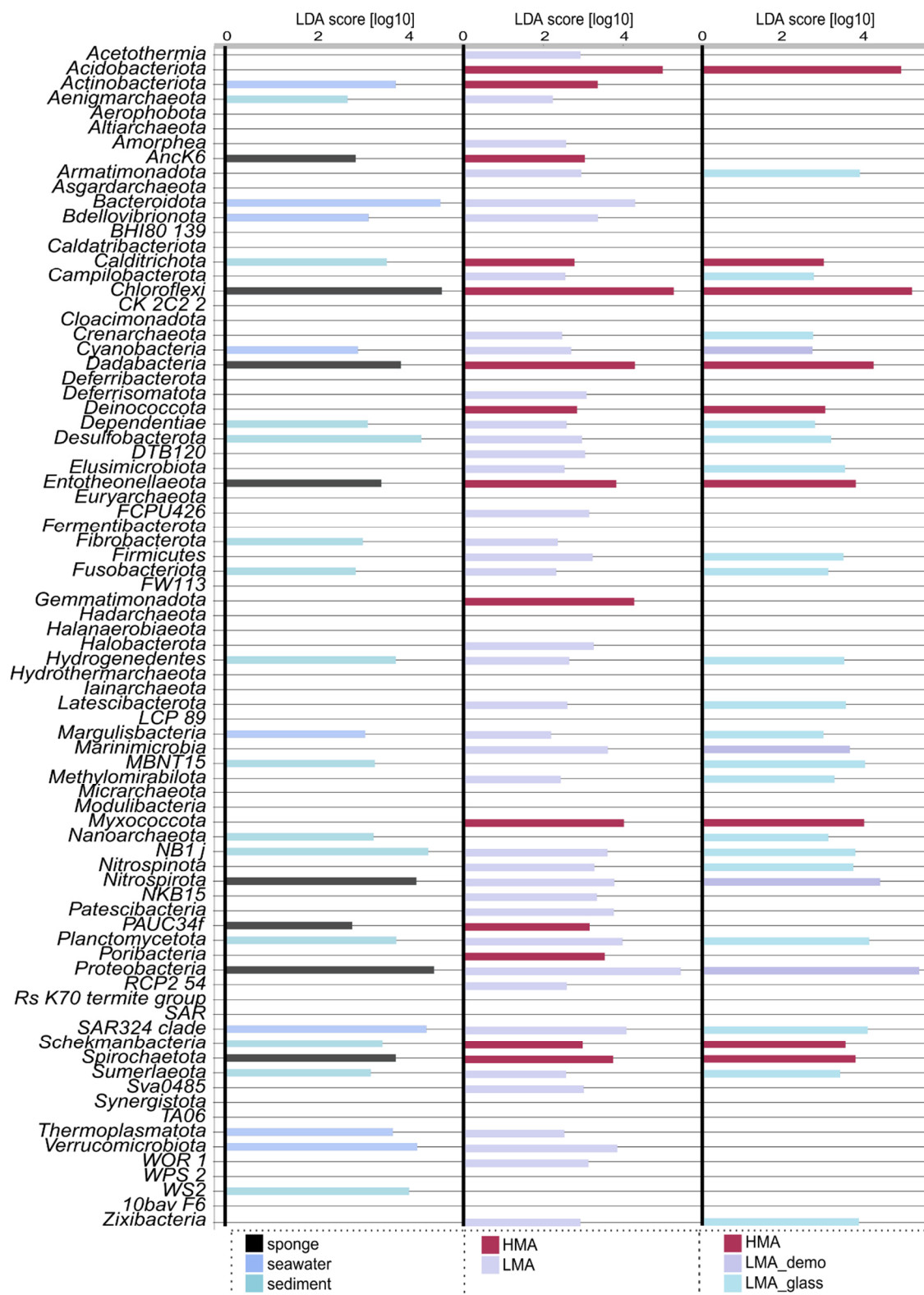


Figure S6 Linear discriminant analysis (LDA) Effect Size (LEfSe) plots. On the y-axis all microbial phyla which were detected in the dataset are shown in alphabetical order. Colored bars are shown if phyla were significantly enriched in one sample category. Left most panel shows variations between sponge, seawater, and sediment. Middle panel shows variations between HMA and LMA sponges, and the right most panel shows variations between HMA, LMA_demo, and LMA_glass sponges.

Table S12 Results of statistical testing (PERMANOVAs based on weighted UniFrac distances) conducted to assess differences in the microbial community composition between host morphology categories for the different sponge types.

Group1	Group2	Sample numbers (Group1 : Group2)	Permuta- tions	pseudo F	p-value
HMA_A	HMA_I	205 : 11	999	7.0	0.001
HMA_A	HMA_K	205 : 104	999	26.1	0.001
HMA_I	HMA_K	11 : 104	999	2.4	0.032
LMA_demo_A	LMA_demo_C	13 : 5	999	1.8	0.071
LMA_demo_A	LMA_demo_D	13 : 117	999	9.6	0.001
LMA_demo_A	LMA_demo_F	13 : 15	999	19.8	0.001
LMA_demo_A	LMA_demo_E	13 : 9	999	5.6	0.001
LMA_demo_A	LMA_demo_I	13 : 1	999	3.0	0.080
LMA_demo_A	LMA_demo_J	13 : 17	999	3.6	0.001
LMA_demo_A	LMA_demo_L	13 : 10	999	6.5	0.001
LMA_demo_A	LMA_demo_H	13 : 14	999	2.6	0.034
LMA_demo_A	LMA_demo_K	13 : 31	999	5.2	0.001
LMA_demo_C	LMA_demo_F	5 : 15	999	12.0	0.001
LMA_demo_C	LMA_demo_D	5 : 117	999	5.1	0.001
LMA_demo_C	LMA_demo_H	5 : 14	999	1.8	0.123
LMA_demo_C	LMA_demo_K	5 : 31	999	4.1	0.007
LMA_demo_C	LMA_demo_E	5 : 9	999	4.5	0.001
LMA_demo_C	LMA_demo_I	5 : 1	999	1.9	0.178
LMA_demo_C	LMA_demo_J	5 : 17	999	3.6	0.002
LMA_demo_C	LMA_demo_L	5 : 10	999	3.6	0.006
LMA_demo_D	LMA_demo_I	117 : 1	999	3.3	0.010
LMA_demo_D	LMA_demo_J	117 : 17	999	15.5	0.001
LMA_demo_D	LMA_demo_L	117 : 10	999	14.9	0.001
LMA_demo_D	LMA_demo_H	117 : 14	999	13.6	0.001
LMA_demo_D	LMA_demo_F	117 : 15	999	35.1	0.001
LMA_demo_D	LMA_demo_K	117 : 31	999	14.8	0.001
LMA_demo_D	LMA_demo_E	117 : 9	999	15.0	0.001
LMA_demo_E	LMA_demo_L	9 : 10	999	11.0	0.001
LMA_demo_E	LMA_demo_F	9 : 15	999	33.0	0.001
LMA_demo_E	LMA_demo_K	9 : 31	999	10.5	0.001
LMA_demo_E	LMA_demo_I	9 : 1	999	7.4	0.096
LMA_demo_E	LMA_demo_J	9 : 17	999	9.4	0.001
LMA_demo_E	LMA_demo_H	9 : 14	999	4.8	0.003
LMA_demo_F	LMA_demo_L	15 : 10	999	20.6	0.001
LMA_demo_F	LMA_demo_J	15 : 17	999	27.0	0.001
LMA_demo_F	LMA_demo_I	15 : 1	999	7.1	0.076
LMA_demo_F	LMA_demo_H	15 : 14	999	13.7	0.001
LMA_demo_F	LMA_demo_K	15 : 31	999	29.8	0.001
LMA_demo_H	LMA_demo_I	14 : 1	999	1.8	0.070
LMA_demo_H	LMA_demo_K	14 : 31	999	6.6	0.001
LMA_demo_H	LMA_demo_J	14 : 17	999	6.2	0.001
LMA_demo_H	LMA_demo_L	14 : 10	999	3.3	0.003
LMA_demo_I	LMA_demo_L	1 : 10	999	1.1	0.439
LMA_demo_I	LMA_demo_J	1 : 17	999	3.9	0.053
LMA_demo_I	LMA_demo_K	1 : 31	999	3.7	0.035

LMA_demo_J	LMA_demo_K	17 : 31	999	9.5	0.001
LMA_demo_J	LMA_demo_L	17 : 10	999	11.9	0.001
LMA_demo_K	LMA_demo_L	31 : 10	999	10.7	0.001
LMA_glass_B	LMA_glass_C	139 : 14	999	12.6	0.001
LMA_glass_B	LMA_glass_E	139 : 15	999	5.0	0.002
LMA_glass_B	LMA_glass_G	139 : 21	999	4.7	0.001
LMA_glass_B	LMA_glass_F	139 : 2	999	2.1	0.027
LMA_glass_B	LMA_glass_L	139 : 5	999	2.2	0.027
LMA_glass_C	LMA_glass_F	14 : 2	999	30.7	0.010
LMA_glass_C	LMA_glass_G	14 : 21	999	20.9	0.001
LMA_glass_C	LMA_glass_E	14 : 15	999	13.5	0.001
LMA_glass_C	LMA_glass_L	14 : 5	999	17.6	0.001
LMA_glass_E	LMA_glass_F	15 : 2	999	2.6	0.028
LMA_glass_E	LMA_glass_L	15 : 5	999	3.2	0.003
LMA_glass_E	LMA_glass_G	15 : 21	999	5.8	0.001
LMA_glass_F	LMA_glass_G	2 : 21	999	2.3	0.030
LMA_glass_F	LMA_glass_L	2 : 5	999	2.9	0.048
LMA_glass_G	LMA_glass_L	21 : 5	999	1.9	0.050

Table S13 Results of statistical testing (Dunn's tests) conducted to assess variations in microbial alpha diversity (Shannon index) between host morphology categories for the different sponge types.

Group1	Group2	Sample number (Group1 : Group2)	Dunns Z	p-value
HMA_A	HMA_I	205 : 11	-0.2	0.862
HMA_A	HMA_K	205 : 104	-0.9	0.356
HMA_I	HMA_K	11 : 104	-0.2	0.857
LMA_demo_A	LMA_demo_C	13 : 5	0.1	0.954
LMA_demo_A	LMA_demo_D	13 : 117	-0.9	0.347
LMA_demo_A	LMA_demo_I	13 : 1	0.4	0.684
LMA_demo_A	LMA_demo_E	13 : 9	1.3	0.203
LMA_demo_A	LMA_demo_J	13 : 17	-0.3	0.779
LMA_demo_A	LMA_demo_F	13 : 15	1.1	0.283
LMA_demo_A	LMA_demo_K	13 : 31	0.5	0.628
LMA_demo_A	LMA_demo_H	13 : 14	1.0	0.313
LMA_demo_A	LMA_demo_L	13 : 10	0.7	0.476
LMA_demo_C	LMA_demo_J	5 : 17	-0.3	0.793
LMA_demo_C	LMA_demo_E	5 : 9	0.9	0.350
LMA_demo_C	LMA_demo_K	5 : 31	0.3	0.787
LMA_demo_C	LMA_demo_I	5 : 1	0.4	0.720
LMA_demo_C	LMA_demo_F	5 : 15	0.7	0.466
LMA_demo_C	LMA_demo_L	5 : 10	0.5	0.623
LMA_demo_C	LMA_demo_H	5 : 14	0.7	0.491
LMA_demo_C	LMA_demo_D	5 : 117	-0.7	0.504
LMA_demo_D	LMA_demo_E	117 : 9	2.4	0.017
LMA_demo_D	LMA_demo_K	117 : 31	2.2	0.031
LMA_demo_D	LMA_demo_I	117 : 1	0.7	0.487
LMA_demo_D	LMA_demo_F	117 : 15	2.5	0.013
LMA_demo_D	LMA_demo_H	117 : 14	2.3	0.019
LMA_demo_D	LMA_demo_L	117 : 10	1.7	0.081
LMA_demo_D	LMA_demo_J	117 : 17	0.7	0.509
LMA_demo_E	LMA_demo_I	9 : 1	-0.1	0.903
LMA_demo_E	LMA_demo_F	9 : 15	-0.3	0.731
LMA_demo_E	LMA_demo_K	9 : 31	-1.0	0.301
LMA_demo_E	LMA_demo_L	9 : 10	-0.5	0.583
LMA_demo_E	LMA_demo_H	9 : 14	-0.4	0.703
LMA_demo_E	LMA_demo_J	9 : 17	-1.6	0.112
LMA_demo_F	LMA_demo_I	15 : 1	0.0	0.988
LMA_demo_F	LMA_demo_H	15 : 14	0.0	0.961
LMA_demo_F	LMA_demo_L	15 : 10	-0.3	0.793
LMA_demo_F	LMA_demo_K	15 : 31	-0.8	0.433
LMA_demo_F	LMA_demo_J	15 : 17	-1.4	0.150
LMA_demo_H	LMA_demo_J	14 : 17	-1.4	0.173
LMA_demo_H	LMA_demo_I	14 : 1	0.0	0.974
LMA_demo_H	LMA_demo_K	14 : 31	-0.7	0.478
LMA_demo_H	LMA_demo_L	14 : 10	-0.2	0.830
LMA_demo_I	LMA_demo_K	1 : 31	-0.3	0.796
LMA_demo_I	LMA_demo_J	1 : 17	-0.5	0.609
LMA_demo_I	LMA_demo_L	1 : 10	-0.1	0.907
LMA_demo_J	LMA_demo_K	17 : 31	0.9	0.382

LMA_demo_J	LMA_demo_L	17 : 10	1.0	0.312
LMA_demo_K	LMA_demo_L	31 : 10	0.4	0.702
LMA_glass_B	LMA_glass_C	139 : 14	3.0	0.002
LMA_glass_B	LMA_glass_E	139 : 15	3.5	0.000
LMA_glass_B	LMA_glass_G	139 : 21	-1.0	0.326
LMA_glass_B	LMA_glass_F	139 : 2	1.5	0.134
LMA_glass_B	LMA_glass_L	139 : 5	-1.9	0.056
LMA_glass_C	LMA_glass_F	14 : 2	0.3	0.777
LMA_glass_C	LMA_glass_G	14 : 21	-3.1	0.002
LMA_glass_C	LMA_glass_E	14 : 15	0.3	0.770
LMA_glass_C	LMA_glass_L	14 : 5	-3.3	0.001
LMA_glass_E	LMA_glass_F	15 : 2	0.1	0.889
LMA_glass_E	LMA_glass_G	15 : 21	-3.5	0.000
LMA_glass_E	LMA_glass_L	15 : 5	-3.5	0.000
LMA_glass_F	LMA_glass_G	2 : 21	-1.8	0.080
LMA_glass_F	LMA_glass_L	2 : 5	-2.3	0.021
LMA_glass_G	LMA_glass_L	21 : 5	-1.3	0.198

Table S14 Results of statistical testing (PERMANOVAs based on weighted UniFrac distances) conducted to assess differences in the microbial community composition between watermasses for the different sampletypes.

Group1	Group2	Sample numb.	Permu-tations	pseudo F	p-value
HMA_AAIW	HMA_ADW	17 : 77	999	24.4	0.001
HMA_AAIW	HMA_AIW	17 : 71	999	15.4	0.001
HMA_AAIW	HMA_LSW	17 : 3	999	4.2	0.004
HMA_AAIW	HMA_AW	17 : 63	999	18.9	0.001
HMA_AAIW	HMA_MOW	17 : 58	999	13.0	0.001
HMA_AAIW	HMA_ENACW	17 : 2	999	3.2	0.009
HMA_AAIW	HMA_WSPW	17 : 14	999	9.5	0.001
HMA_ADW	HMA_ENACW	77 : 2	999	7.0	0.001
HMA_ADW	HMA_AIW	77 : 71	999	9.0	0.001
HMA_ADW	HMA_AW	77 : 63	999	46.7	0.001
HMA_ADW	HMA_LSW	77 : 3	999	10.5	0.001
HMA_ADW	HMA_WSPW	77 : 14	999	28.2	0.001
HMA_ADW	HMA_MOW	77 : 58	999	77.1	0.001
HMA_AIW	HMA_ENACW	71 : 2	999	4.6	0.004
HMA_AIW	HMA_AW	71 : 63	999	23.1	0.001
HMA_AIW	HMA_MOW	71 : 58	999	48.5	0.001
HMA_AIW	HMA_WSPW	71 : 14	999	21.3	0.001
HMA_AIW	HMA_LSW	71 : 3	999	6.5	0.001
HMA_AW	HMA_WSPW	63 : 14	999	18.8	0.001
HMA_AW	HMA_LSW	63 : 3	999	7.0	0.001
HMA_AW	HMA_MOW	63 : 58	999	38.2	0.001
HMA_AW	HMA_ENACW	63 : 2	999	5.2	0.003
HMA_ENACW	HMA_LSW	2 : 3	999	2.0	0.194
HMA_ENACW	HMA_MOW	2 : 58	999	1.3	0.175
HMA_ENACW	HMA_WSPW	2 : 14	999	2.5	0.081
HMA_LSW	HMA_MOW	3 : 58	999	3.3	0.003
HMA_LSW	HMA_WSPW	3 : 14	999	2.6	0.096
HMA_MOW	HMA_WSPW	58 : 14	999	19.1	0.001
LMA_demo_AAIW	LMA_demo_ADW	6 : 36	999	6.8	0.001
LMA_demo_AAIW	LMA_demo_AIW	6 : 27	999	6.6	0.001
LMA_demo_AAIW	LMA_demo_ENACW	6 : 14	999	23.1	0.001
LMA_demo_AAIW	LMA_demo_AW	6 : 53	999	3.8	0.001
LMA_demo_AAIW	LMA_demo_MNwCCW	6 : 6	999	10.3	0.004
LMA_demo_AAIW	LMA_demo_MOW	6 : 36	999	6.2	0.001
LMA_demo_AAIW	LMA_demo_WSPW	6 : 46	999	2.0	0.025
LMA_demo_AAIW	LMA_demo_LSW	6 : 25	999	5.9	0.002
LMA_demo_AAIW	LMA_demo_NwCCW	6 : 9	999	11.2	0.001
LMA_demo_ADW	LMA_demo_ENACW	36 : 14	999	17.4	0.001
LMA_demo_ADW	LMA_demo_AIW	36 : 27	999	7.1	0.001
LMA_demo_ADW	LMA_demo_LSW	36 : 25	999	8.2	0.001

LMA_demo_ADW	LMA_demo_NwCCW	36 : 9	999	14.6	0.001
LMA_demo_ADW	LMA_demo_AW	36 : 53	999	11.3	0.001
LMA_demo_ADW	LMA_demo_MNwCCW	36 : 6	999	13.1	0.001
LMA_demo_ADW	LMA_demo_MOW	36 : 36	999	17.3	0.001
LMA_demo_ADW	LMA_demo_WSPW	36 : 46	999	8.0	0.001
LMA_demo_AIW	LMA_demo_MNwCCW	27 : 6	999	7.1	0.001
LMA_demo_AIW	LMA_demo_MOW	27 : 36	999	9.5	0.001
LMA_demo_AIW	LMA_demo_WSPW	27 : 46	999	7.8	0.001
LMA_demo_AIW	LMA_demo_LSW	27 : 25	999	10.4	0.001
LMA_demo_AIW	LMA_demo_ENACW	27 : 14	999	13.5	0.001
LMA_demo_AIW	LMA_demo_NwCCW	27 : 9	999	8.0	0.001
LMA_demo_AIW	LMA_demo_AW	27 : 53	999	4.7	0.001
LMA_demo_AW	LMA_demo_WSPW	53 : 46	999	5.9	0.001
LMA_demo_AW	LMA_demo_ENACW	53 : 14	999	14.1	0.001
LMA_demo_AW	LMA_demo_NwCCW	53 : 9	999	4.5	0.002
LMA_demo_AW	LMA_demo_MNwCCW	53 : 6	999	5.8	0.001
LMA_demo_AW	LMA_demo_MOW	53 : 36	999	13.0	0.001
LMA_demo_AW	LMA_demo_LSW	53 : 25	999	12.6	0.001
LMA_demo_ENACW	LMA_demo_WSPW	14 : 46	999	12.4	0.001
LMA_demo_ENACW	LMA_demo_MOW	14 : 36	999	8.8	0.002
LMA_demo_ENACW	LMA_demo_MNwCCW	14 : 6	999	34.5	0.001
LMA_demo_ENACW	LMA_demo_LSW	14 : 25	999	18.7	0.001
LMA_demo_ENACW	LMA_demo_NwCCW	14 : 9	999	47.5	0.001
LMA_demo_LSW	LMA_demo_MNwCCW	25 : 6	999	11.5	0.001
LMA_demo_LSW	LMA_demo_NwCCW	25 : 9	999	15.5	0.001
LMA_demo_LSW	LMA_demo_MOW	25 : 36	999	14.8	0.001
LMA_demo_LSW	LMA_demo_WSPW	25 : 46	999	9.0	0.001
LMA_demo_MNwCCW	LMA_demo_WSPW	6 : 46	999	6.9	0.001
LMA_demo_MNwCCW	LMA_demo_MOW	6 : 36	999	7.7	0.001
LMA_demo_MNwCCW	LMA_demo_NwCCW	6 : 9	999	17.6	0.001
LMA_demo_MOW	LMA_demo_NwCCW	36 : 9	999	12.4	0.001
LMA_demo_MOW	LMA_demo_WSPW	36 : 46	999	12.6	0.001
LMA_demo_NwCCW	LMA_demo_WSPW	9 : 46	999	7.8	0.001
LMA_glass_AAIW	LMA_glass_ADW	39 : 45	999	16.3	0.001
LMA_glass_AAIW	LMA_glass_AIW	39 : 19	999	16.9	0.001
LMA_glass_AAIW	LMA_glass_CDW	39 : 8	999	2.1	0.038
LMA_glass_AAIW	LMA_glass_AW	39 : 11	999	8.2	0.001
LMA_glass_AAIW	LMA_glass_LSW	39 : 29	999	9.1	0.001
LMA_glass_AAIW	LMA_glass_MOW	39 : 16	999	12.2	0.001
LMA_glass_AAIW	LMA_glass_WSW	39 : 32	999	27.5	0.001
LMA_glass_AAIW	LMA_glass_ESW	39 : 4	999	5.2	0.001
LMA_glass_AAIW	LMA_glass_WSPW	39 : 34	999	1.5	0.123
LMA_glass_ADW	LMA_glass_CDW	45 : 8	999	5.9	0.001
LMA_glass_ADW	LMA_glass_AIW	45 : 19	999	5.6	0.001
LMA_glass_ADW	LMA_glass_ESW	45 : 4	999	8.8	0.001
LMA_glass_ADW	LMA_glass_WSPW	45 : 34	999	10.2	0.001

LMA_glass_ADW	LMA_glass_AW	45 : 11	999	4.8	0.002
LMA_glass_ADW	LMA_glass_LSW	45 : 29	999	28.7	0.001
LMA_glass_ADW	LMA_glass_MOW	45 : 16	999	33.0	0.001
LMA_glass_ADW	LMA_glass_WSW	45 : 32	999	31.8	0.001
LMA_glass_AIW	LMA_glass_LSW	19 : 29	999	31.2	0.001
LMA_glass_AIW	LMA_glass_MOW	19 : 16	999	50.1	0.001
LMA_glass_AIW	LMA_glass_WSW	19 : 32	999	42.2	0.001
LMA_glass_AIW	LMA_glass_ESW	19 : 4	999	17.4	0.001
LMA_glass_AIW	LMA_glass_CDW	19 : 8	999	10.3	0.001
LMA_glass_AIW	LMA_glass_WSPW	19 : 34	999	12.7	0.001
LMA_glass_AIW	LMA_glass_AW	19 : 11	999	5.7	0.001
LMA_glass_AW	LMA_glass_WSW	11 : 32	999	20.5	0.001
LMA_glass_AW	LMA_glass_CDW	11 : 8	999	5.4	0.001
LMA_glass_AW	LMA_glass_WSPW	11 : 34	999	5.6	0.001
LMA_glass_AW	LMA_glass_LSW	11 : 29	999	16.0	0.001
LMA_glass_AW	LMA_glass_MOW	11 : 16	999	30.9	0.001
LMA_glass_AW	LMA_glass_ESW	11 : 4	999	12.7	0.002
LMA_glass_CDW	LMA_glass_WSW	8 : 32	999	15.2	0.001
LMA_glass_CDW	LMA_glass_MOW	8 : 16	999	8.8	0.001
LMA_glass_CDW	LMA_glass_LSW	8 : 29	999	4.6	0.004
LMA_glass_CDW	LMA_glass_ESW	8 : 4	999	4.7	0.001
LMA_glass_CDW	LMA_glass_WSPW	8 : 34	999	1.8	0.048
LMA_glass_ESW	LMA_glass_LSW	4 : 29	999	10.5	0.001
LMA_glass_ESW	LMA_glass_WSPW	4 : 34	999	4.5	0.001
LMA_glass_ESW	LMA_glass_MOW	4 : 16	999	21.7	0.001
LMA_glass_ESW	LMA_glass_WSW	4 : 32	999	6.5	0.001
LMA_glass_LSW	LMA_glass_WSW	29 : 32	999	39.7	0.001
LMA_glass_LSW	LMA_glass_MOW	29 : 16	999	5.0	0.007
LMA_glass_LSW	LMA_glass_WSPW	29 : 34	999	9.4	0.001
LMA_glass_MOW	LMA_glass_WSPW	16 : 34	999	13.9	0.001
LMA_glass_MOW	LMA_glass_WSW	16 : 32	999	50.5	0.001
LMA_glass_WSPW	LMA_glass_WSW	34 : 32	999	21.9	0.001
seawater_AAIW	seawater_ADW	15 : 82	999	40.9	0.001
seawater_AAIW	seawater_AIW	15 : 31	999	38.2	0.001
seawater_AAIW	seawater_LSW	15 : 22	999	43.9	0.001
seawater_AAIW	seawater_CDW	15 : 3	999	3.0	0.022
seawater_AAIW	seawater_AW	15 : 48	999	39.5	0.001
seawater_AAIW	seawater_WSW	15 : 19	999	45.6	0.001
seawater_AAIW	seawater_MNwCCW	15 : 3	999	13.0	0.002
seawater_AAIW	seawater_ENACW	15 : 6	999	16.0	0.001
seawater_AAIW	seawater_MOW	15 : 18	999	43.8	0.001
seawater_AAIW	seawater_WSPW	15 : 24	999	5.1	0.001
seawater_ADW	seawater_CDW	82 : 3	999	14.3	0.001
seawater_ADW	seawater_AW	82 : 48	999	48.3	0.001
seawater_ADW	seawater_WSPW	82 : 24	999	32.0	0.001
seawater_ADW	seawater_WSW	82 : 19	999	45.0	0.001

seawater_ADW	seawater_ENACW	82 : 6	999	8.2	0.001
seawater_ADW	seawater_AIW	82 : 31	999	16.8	0.001
seawater_ADW	seawater_MOW	82 : 18	999	19.1	0.001
seawater_ADW	seawater_LSW	82 : 22	999	23.7	0.001
seawater_ADW	seawater_MNwCCW	82 : 3	999	5.5	0.002
seawater_AIW	seawater_CDW	31 : 3	999	16.6	0.001
seawater_AIW	seawater_ENACW	31 : 6	999	8.7	0.001
seawater_AIW	seawater_MNwCCW	31 : 3	999	4.7	0.002
seawater_AIW	seawater_WSW	31 : 19	999	28.4	0.001
seawater_AIW	seawater_LSW	31 : 22	999	33.2	0.001
seawater_AIW	seawater_WSPW	31 : 24	999	25.1	0.001
seawater_AIW	seawater_MOW	31 : 18	999	29.3	0.001
seawater_AIW	seawater_AW	31 : 48	999	12.5	0.001
seawater_AW	seawater_WSW	48 : 19	999	19.3	0.001
seawater_AW	seawater_ENACW	48 : 6	999	7.6	0.001
seawater_AW	seawater_MOW	48 : 18	999	33.6	0.001
seawater_AW	seawater_MNwCCW	48 : 3	999	4.1	0.005
seawater_AW	seawater_WSPW	48 : 24	999	29.5	0.001
seawater_AW	seawater_LSW	48 : 22	999	42.0	0.001
seawater_AW	seawater_CDW	48 : 3	999	14.7	0.001
seawater_CDW	seawater_ENACW	3 : 6	999	21.2	0.009
seawater_CDW	seawater_LSW	3 : 22	999	30.0	0.001
seawater_CDW	seawater_MNwCCW	3 : 3	999	44.9	0.106
seawater_CDW	seawater_WSPW	3 : 24	999	4.1	0.002
seawater_CDW	seawater_WSW	3 : 19	999	27.1	0.002
seawater_CDW	seawater_MOW	3 : 18	999	36.0	0.002
seawater_ENACW	seawater_WSPW	6 : 24	999	6.9	0.001
seawater_ENACW	seawater_LSW	6 : 22	999	8.1	0.001
seawater_ENACW	seawater_MOW	6 : 18	999	4.8	0.001
seawater_ENACW	seawater_WSW	6 : 19	999	11.7	0.001
seawater_ENACW	seawater_MNwCCW	6 : 3	999	5.0	0.010
seawater_LSW	seawater_WSPW	22 : 24	999	24.9	0.001
seawater_LSW	seawater_WSW	22 : 19	999	59.1	0.001
seawater_LSW	seawater_MOW	22 : 18	999	4.9	0.001
seawater_LSW	seawater_MNwCCW	22 : 3	999	12.2	0.001
seawater_MNwCCW	seawater_MOW	3 : 18	999	11.7	0.001
seawater_MNwCCW	seawater_WSW	3 : 19	999	7.4	0.001
seawater_MNwCCW	seawater_WSPW	3 : 24	999	6.0	0.001
seawater_MOW	seawater_WSPW	18 : 24	999	22.5	0.001
seawater_MOW	seawater_WSW	18 : 19	999	48.9	0.001
seawater_WSPW	seawater_WSW	24 : 19	999	27.5	0.001
sediment_AAIW	sediment_ADW	6 : 19	999	2.8	0.004
sediment_AAIW	sediment_AIW	6 : 21	999	4.2	0.002
sediment_AAIW	sediment_MOW	6 : 3	999	2.2	0.032
sediment_AAIW	sediment_AW	6 : 7	999	2.3	0.020
sediment_AAIW	sediment_LSW	6 : 9	999	4.3	0.001

sediment_AAIW	sediment_WSPW	6 : 6	999	0.2	1.000
sediment_AAIW	sediment_CDW	6 : 6	999	5.0	0.001
sediment_AAIW	sediment_WSW	6 : 23	999	3.3	0.006
sediment_ADW	sediment_AIW	19 : 21	999	2.8	0.017
sediment_ADW	sediment_AW	19 : 7	999	3.0	0.008
sediment_ADW	sediment_WSPW	19 : 6	999	3.1	0.004
sediment_ADW	sediment_LSW	19 : 9	999	3.9	0.003
sediment_ADW	sediment_MOW	19 : 3	999	2.2	0.059
sediment_ADW	sediment_WSW	19 : 23	999	17.1	0.001
sediment_ADW	sediment_CDW	19 : 6	999	3.4	0.007
sediment_AIW	sediment_AW	21 : 7	999	4.5	0.001
sediment_AIW	sediment_CDW	21 : 6	999	2.9	0.014
sediment_AIW	sediment_WSPW	21 : 6	999	4.7	0.001
sediment_AIW	sediment_LSW	21 : 9	999	5.4	0.001
sediment_AIW	sediment_WSW	21 : 23	999	22.8	0.001
sediment_AIW	sediment_MOW	21 : 3	999	2.6	0.011
sediment_AW	sediment_CDW	7 : 6	999	8.2	0.002
sediment_AW	sediment_MOW	7 : 3	999	2.9	0.027
sediment_AW	sediment_WSPW	7 : 6	999	2.3	0.029
sediment_AW	sediment_WSW	7 : 23	999	10.5	0.001
sediment_AW	sediment_LSW	7 : 9	999	4.3	0.001
sediment_CDW	sediment_WSPW	6 : 6	999	5.6	0.002
sediment_CDW	sediment_LSW	6 : 9	999	11.9	0.001
sediment_CDW	sediment_WSW	6 : 23	999	16.8	0.001
sediment_CDW	sediment_MOW	6 : 3	999	8.1	0.007
sediment_LSW	sediment_MOW	9 : 3	999	1.1	0.338
sediment_LSW	sediment_WSPW	9 : 6	999	4.6	0.001
sediment_LSW	sediment_WSW	9 : 23	999	17.8	0.001
sediment_MOW	sediment_WSPW	3 : 6	999	2.4	0.058
sediment_MOW	sediment_WSW	3 : 23	999	8.2	0.002
sediment_WSPW	sediment_WSW	6 : 23	999	2.7	0.029

Table S15 Results of statistical testing (Dunn's tests) conducted to assess variations in microbial alpha diversity (Shannon index) between watermasses for the different sampletypes.

Group1	Group2	Sample numbers (Group1:Group2)	Dunns Z	p-value
HMA_AAIW	HMA_ADW	17 : 77	0.8	0.408
HMA_AAIW	HMA_AIW	17 : 71	0.8	0.444
HMA_AAIW	HMA_MOW	17 : 58	0.6	0.517
HMA_AAIW	HMA_AW	17 : 63	1.0	0.336
HMA_AAIW	HMA_WSPW	17 : 14	2.7	0.007
HMA_AAIW	HMA_ENACW	17 : 2	0.0	0.989
HMA_AAIW	HMA_LSW	17 : 3	1.4	0.174
HMA_ADW	HMA_WSPW	77 : 14	2.6	0.010
HMA_ADW	HMA_ENACW	77 : 2	-0.3	0.768
HMA_ADW	HMA_AW	77 : 63	0.2	0.807
HMA_ADW	HMA_LSW	77 : 3	1.1	0.285
HMA_ADW	HMA_MOW	77 : 58	-0.2	0.804
HMA_ADW	HMA_AIW	77 : 71	-0.1	0.928
HMA_AIW	HMA_AW	71 : 63	0.3	0.745
HMA_AIW	HMA_MOW	71 : 58	-0.2	0.873
HMA_AIW	HMA_LSW	71 : 3	1.1	0.275
HMA_AIW	HMA_ENACW	71 : 2	-0.3	0.784
HMA_AIW	HMA_WSPW	71 : 14	2.6	0.009
HMA_AW	HMA_MOW	63 : 58	-0.5	0.642
HMA_AW	HMA_ENACW	63 : 2	-0.4	0.725
HMA_AW	HMA_WSPW	63 : 14	2.4	0.017
HMA_AW	HMA_LSW	63 : 3	1.0	0.320
HMA_ENACW	HMA_LSW	2 : 3	0.9	0.357
HMA_ENACW	HMA_MOW	2 : 58	0.2	0.815
HMA_ENACW	HMA_WSPW	2 : 14	1.3	0.204
HMA_LSW	HMA_MOW	3 : 58	-1.1	0.256
HMA_LSW	HMA_WSPW	3 : 14	0.2	0.851
HMA_MOW	HMA_WSPW	58 : 14	2.7	0.008
LMA_demo_AAIW	LMA_demo_ADW	6 : 36	0.5	0.625
LMA_demo_AAIW	LMA_demo_AIW	6 : 27	-0.4	0.692
LMA_demo_AAIW	LMA_demo_MNwCCW	6 : 6	-0.2	0.841
LMA_demo_AAIW	LMA_demo_AW	6 : 53	-0.2	0.872
LMA_demo_AAIW	LMA_demo_MOW	6 : 36	-0.8	0.440
LMA_demo_AAIW	LMA_demo_ENACW	6 : 14	0.3	0.773
LMA_demo_AAIW	LMA_demo_NwCCW	6 : 9	-1.2	0.225
LMA_demo_AAIW	LMA_demo_LSW	6 : 25	0.4	0.692
LMA_demo_AAIW	LMA_demo_WSPW	6 : 46	0.6	0.574
LMA_demo_ADW	LMA_demo_MOW	36 : 36	-2.4	0.018
LMA_demo_ADW	LMA_demo_AW	36 : 53	-1.3	0.187
LMA_demo_ADW	LMA_demo_NwCCW	36 : 9	-2.3	0.022
LMA_demo_ADW	LMA_demo_MNwCCW	36 : 6	-0.8	0.452
LMA_demo_ADW	LMA_demo_ENACW	36 : 14	-0.2	0.812
LMA_demo_ADW	LMA_demo_WSPW	36 : 46	0.1	0.899
LMA_demo_ADW	LMA_demo_LSW	36 : 25	-0.1	0.892
LMA_demo_ADW	LMA_demo_AIW	36 : 27	-1.5	0.121
LMA_demo_AIW	LMA_demo_AW	27 : 53	0.5	0.643

LMA_demo_AIW	LMA_demo_NwCCW	27 : 9	-1.2	0.231
LMA_demo_AIW	LMA_demo_MNwCCW	27 : 6	0.1	0.890
LMA_demo_AIW	LMA_demo_ENACW	27 : 14	1.0	0.332
LMA_demo_AIW	LMA_demo_LSW	27 : 25	1.3	0.196
LMA_demo_AIW	LMA_demo_WSPW	27 : 46	1.7	0.081
LMA_demo_AIW	LMA_demo_MOW	27 : 36	-0.6	0.525
LMA_demo_AW	LMA_demo_MNwCCW	53 : 6	-0.1	0.913
LMA_demo_AW	LMA_demo_ENACW	53 : 14	0.7	0.485
LMA_demo_AW	LMA_demo_NwCCW	53 : 9	-1.6	0.114
LMA_demo_AW	LMA_demo_WSPW	53 : 46	1.6	0.120
LMA_demo_AW	LMA_demo_LSW	53 : 25	1.0	0.304
LMA_demo_AW	LMA_demo_MOW	53 : 36	-1.3	0.209
LMA_demo_ENACW	LMA_demo_MNwCCW	14 : 6	-0.5	0.599
LMA_demo_ENACW	LMA_demo_LSW	14 : 25	0.1	0.905
LMA_demo_ENACW	LMA_demo_WSPW	14 : 46	0.3	0.735
LMA_demo_ENACW	LMA_demo_NwCCW	14 : 9	-1.8	0.068
LMA_demo_ENACW	LMA_demo_MOW	14 : 36	-1.5	0.126
LMA_demo_LSW	LMA_demo_MOW	25 : 36	-2.0	0.045
LMA_demo_LSW	LMA_demo_MNwCCW	25 : 6	-0.7	0.514
LMA_demo_LSW	LMA_demo_NwCCW	25 : 9	-2.1	0.035
LMA_demo_LSW	LMA_demo_WSPW	25 : 46	0.3	0.798
LMA_demo_MNwCCW	LMA_demo_NwCCW	6 : 9	-1.0	0.321
LMA_demo_MNwCCW	LMA_demo_MOW	6 : 36	-0.5	0.610
LMA_demo_MNwCCW	LMA_demo_WSPW	6 : 46	0.8	0.407
LMA_demo_MOW	LMA_demo_NwCCW	36 : 9	-0.8	0.423
LMA_demo_MOW	LMA_demo_WSPW	36 : 46	2.6	0.009
LMA_demo_NwCCW	LMA_demo_WSPW	9 : 46	2.4	0.015
LMA_glass_AAIW	LMA_glass_ADW	39 : 45	-2.4	0.017
LMA_glass_AAIW	LMA_glass_AIW	39 : 19	-2.2	0.025
LMA_glass_AAIW	LMA_glass_LSW	39 : 29	-1.9	0.061
LMA_glass_AAIW	LMA_glass_AW	39 : 11	-2.9	0.004
LMA_glass_AAIW	LMA_glass_MNwCCW	39 : 1	-0.3	0.791
LMA_glass_AAIW	LMA_glass_WSW	39 : 32	-2.6	0.009
LMA_glass_AAIW	LMA_glass_CDW	39 : 8	-1.3	0.190
LMA_glass_AAIW	LMA_glass_MOW	39 : 16	-1.9	0.063
LMA_glass_AAIW	LMA_glass_ESW	39 : 4	0.6	0.542
LMA_glass_AAIW	LMA_glass_WSPW	39 : 34	-1.1	0.268
LMA_glass_ADW	LMA_glass_AW	45 : 11	-1.4	0.175
LMA_glass_ADW	LMA_glass_MNwCCW	45 : 1	0.3	0.801
LMA_glass_ADW	LMA_glass_LSW	45 : 29	0.3	0.789
LMA_glass_ADW	LMA_glass_MOW	45 : 16	-0.1	0.918
LMA_glass_ADW	LMA_glass_CDW	45 : 8	0.0	0.971
LMA_glass_ADW	LMA_glass_WSW	45 : 32	-0.5	0.653
LMA_glass_ADW	LMA_glass_ESW	45 : 4	1.6	0.106
LMA_glass_ADW	LMA_glass_WSPW	45 : 34	1.2	0.247
LMA_glass_ADW	LMA_glass_AIW	45 : 19	-0.4	0.698
LMA_glass_AIW	LMA_glass_MOW	19 : 16	0.2	0.823
LMA_glass_AIW	LMA_glass_LSW	19 : 29	0.6	0.565
LMA_glass_AIW	LMA_glass_WSW	19 : 32	0.0	0.995
LMA_glass_AIW	LMA_glass_CDW	19 : 8	0.3	0.776
LMA_glass_AIW	LMA_glass_AW	19 : 11	-0.9	0.355
LMA_glass_AIW	LMA_glass_WSPW	19 : 34	1.3	0.198

LMA_glass_AIW	LMA_glass_ESW	19 : 4	1.7	0.085
LMA_glass_AIW	LMA_glass_MNwCCW	19 : 1	0.4	0.725
LMA_glass_AW	LMA_glass_LSW	11 : 29	1.5	0.142
LMA_glass_AW	LMA_glass_CDW	11 : 8	1.0	0.311
LMA_glass_AW	LMA_glass_MOW	11 : 16	1.1	0.276
LMA_glass_AW	LMA_glass_ESW	11 : 4	2.2	0.026
LMA_glass_AW	LMA_glass_WSPW	11 : 34	2.1	0.038
LMA_glass_AW	LMA_glass_WSW	11 : 32	1.0	0.314
LMA_glass_AW	LMA_glass_MNwCCW	11 : 1	0.7	0.496
LMA_glass_CDW	LMA_glass_ESW	8 : 4	1.4	0.176
LMA_glass_CDW	LMA_glass_WSPW	8 : 34	0.6	0.526
LMA_glass_CDW	LMA_glass_LSW	8 : 29	0.1	0.901
LMA_glass_CDW	LMA_glass_MOW	8 : 16	-0.1	0.919
LMA_glass_CDW	LMA_glass_MNwCCW	8 : 1	0.2	0.821
LMA_glass_CDW	LMA_glass_WSW	8 : 32	-0.3	0.765
LMA_glass_ESW	LMA_glass_MNwCCW	4 : 1	-0.5	0.599
LMA_glass_ESW	LMA_glass_MOW	4 : 16	-1.6	0.119
LMA_glass_ESW	LMA_glass_LSW	4 : 29	-1.5	0.144
LMA_glass_ESW	LMA_glass_WSW	4 : 32	-1.8	0.074
LMA_glass_ESW	LMA_glass_WSPW	4 : 34	-1.1	0.273
LMA_glass_LSW	LMA_glass_MNwCCW	29 : 1	0.2	0.851
LMA_glass_LSW	LMA_glass_MOW	29 : 16	-0.3	0.763
LMA_glass_LSW	LMA_glass_WSW	29 : 32	-0.7	0.513
LMA_glass_LSW	LMA_glass_WSPW	29 : 34	0.8	0.431
LMA_glass_MNwCCW	LMA_glass_MOW	1 : 16	-0.3	0.783
LMA_glass_MNwCCW	LMA_glass_WSW	1 : 32	-0.4	0.724
LMA_glass_MNwCCW	LMA_glass_WSPW	1 : 34	0.0	0.993
LMA_glass_MOW	LMA_glass_WSPW	16 : 34	1.0	0.334
LMA_glass_MOW	LMA_glass_WSW	16 : 32	-0.2	0.809
LMA_glass_WSPW	LMA_glass_WSW	34 : 32	-1.5	0.136
seawater_AAIW	seawater_ADW	15 : 82	-3.9	0.000
seawater_AAIW	seawater_AIW	15 : 31	-3.9	0.000
seawater_AAIW	seawater_ESW	15 : 1	0.5	0.643
seawater_AAIW	seawater_AW	15 : 48	-4.9	0.000
seawater_AAIW	seawater_LSW	15 : 22	-5.0	0.000
seawater_AAIW	seawater_WSPW	15 : 24	-1.6	0.102
seawater_AAIW	seawater_CDW	15 : 3	0.9	0.392
seawater_AAIW	seawater_MNwCCW	15 : 3	-2.7	0.007
seawater_AAIW	seawater_WSW	15 : 19	-3.9	0.000
seawater_AAIW	seawater_ENACW	15 : 6	-3.4	0.001
seawater_AAIW	seawater_MOW	15 : 18	-4.8	0.000
seawater_ADW	seawater_LSW	82 : 22	-2.4	0.018
seawater_ADW	seawater_AW	82 : 48	-1.9	0.062
seawater_ADW	seawater_WSW	82 : 19	-1.0	0.333
seawater_ADW	seawater_MNwCCW	82 : 3	-1.0	0.301
seawater_ADW	seawater_CDW	82 : 3	2.8	0.005
seawater_ADW	seawater_WSPW	82 : 24	2.4	0.015
seawater_ADW	seawater_ESW	82 : 1	1.6	0.116
seawater_ADW	seawater_ENACW	82 : 6	-1.3	0.185
seawater_ADW	seawater_MOW	82 : 18	-2.3	0.023
seawater_ADW	seawater_AIW	82 : 31	-0.6	0.535
seawater_AIW	seawater_WSW	31 : 19	-0.4	0.692

seawater_AIW	seawater_CDW	31 : 3	2.9	0.003
seawater_AIW	seawater_ESW	31 : 1	1.7	0.092
seawater_AIW	seawater_AW	31 : 48	-0.9	0.367
seawater_AIW	seawater_MOW	31 : 18	-1.6	0.120
seawater_AIW	seawater_WSPW	31 : 24	2.6	0.011
seawater_AIW	seawater_MNwCCW	31 : 3	-0.8	0.431
seawater_AIW	seawater_ENACW	31 : 6	-1.0	0.335
seawater_AIW	seawater_LSW	31 : 22	-1.6	0.116
seawater_AW	seawater_ENACW	48 : 6	-0.5	0.608
seawater_AW	seawater_CDW	48 : 3	3.3	0.001
seawater_AW	seawater_ESW	48 : 1	1.9	0.057
seawater_AW	seawater_LSW	48 : 22	-0.9	0.372
seawater_AW	seawater_MOW	48 : 18	-0.9	0.360
seawater_AW	seawater_WSPW	48 : 24	3.6	0.000
seawater_AW	seawater_WSW	48 : 19	0.3	0.733
seawater_AW	seawater_MNwCCW	48 : 3	-0.5	0.652
seawater_CDW	seawater_WSW	3 : 19	-3.0	0.002
seawater_CDW	seawater_ENACW	3 : 6	-3.1	0.002
seawater_CDW	seawater_MOW	3 : 18	-3.6	0.000
seawater_CDW	seawater_ESW	3 : 1	-0.1	0.957
seawater_CDW	seawater_LSW	3 : 22	-3.6	0.000
seawater_CDW	seawater_WSPW	3 : 24	-1.8	0.078
seawater_CDW	seawater_MNwCCW	3 : 3	-2.8	0.006
seawater_ENACW	seawater_LSW	6 : 22	0.0	0.987
seawater_ENACW	seawater_WSW	6 : 19	0.7	0.501
seawater_ENACW	seawater_MOW	6 : 18	-0.1	0.948
seawater_ENACW	seawater_ESW	6 : 1	2.0	0.047
seawater_ENACW	seawater_MNwCCW	6 : 3	-0.1	0.948
seawater_ENACW	seawater_WSPW	6 : 24	2.5	0.014
seawater_ESW	seawater_LSW	1 : 22	-2.1	0.036
seawater_ESW	seawater_WSW	1 : 19	-1.8	0.075
seawater_ESW	seawater_MNwCCW	1 : 3	-1.9	0.058
seawater_ESW	seawater_WSPW	1 : 24	-1.0	0.319
seawater_ESW	seawater_MOW	1 : 18	-2.1	0.034
seawater_LSW	seawater_MNwCCW	22 : 3	-0.1	0.950
seawater_LSW	seawater_WSPW	22 : 24	3.8	0.000
seawater_LSW	seawater_WSW	22 : 19	1.0	0.304
seawater_LSW	seawater_MOW	22 : 18	-0.1	0.942
seawater_MNwCCW	seawater_WSPW	3 : 24	1.9	0.056
seawater_MNwCCW	seawater_MOW	3 : 18	0.0	0.980
seawater_MNwCCW	seawater_WSW	3 : 19	0.6	0.561
seawater_MOW	seawater_WSPW	18 : 24	3.7	0.000
seawater_MOW	seawater_WSW	18 : 19	1.1	0.294
seawater_WSPW	seawater_WSW	24 : 19	-2.6	0.008
sediment_AAIW	sediment_ADW	6 : 19	0.3	0.783
sediment_AAIW	sediment_AIW	6 : 21	0.1	0.959
sediment_AAIW	sediment_MOW	6 : 3	-0.1	0.909
sediment_AAIW	sediment_AW	6 : 7	-0.2	0.880
sediment_AAIW	sediment_WSPW	6 : 6	-0.1	0.949
sediment_AAIW	sediment_CDW	6 : 6	0.1	0.933
sediment_AAIW	sediment_WSW	6 : 23	0.0	0.976
sediment_AAIW	sediment_LSW	6 : 9	-0.2	0.847

sediment_ADW	sediment_WSPW	19 : 6	-0.4	0.723
sediment_ADW	sediment_AW	19 : 7	-0.5	0.630
sediment_ADW	sediment_WSW	19 : 23	-0.5	0.646
sediment_ADW	sediment_CDW	19 : 6	-0.2	0.863
sediment_ADW	sediment_MOW	19 : 3	-0.3	0.736
sediment_ADW	sediment_LSW	19 : 9	-0.6	0.569
sediment_ADW	sediment_AIW	19 : 21	-0.3	0.740
sediment_AIW	sediment_AW	21 : 7	-0.2	0.804
sediment_AIW	sediment_WSW	21 : 23	-0.1	0.901
sediment_AIW	sediment_MOW	21 : 3	-0.2	0.865
sediment_AIW	sediment_WSPW	21 : 6	-0.1	0.895
sediment_AIW	sediment_LSW	21 : 9	-0.3	0.752
sediment_AIW	sediment_CDW	21 : 6	0.1	0.958
sediment_AW	sediment_MOW	7 : 3	0.0	0.996
sediment_AW	sediment_CDW	7 : 6	0.2	0.812
sediment_AW	sediment_WSW	7 : 23	0.2	0.870
sediment_AW	sediment_WSPW	7 : 6	0.1	0.932
sediment_AW	sediment_LSW	7 : 9	0.0	0.973
sediment_CDW	sediment_LSW	6 : 9	-0.3	0.776
sediment_CDW	sediment_MOW	6 : 3	-0.2	0.855
sediment_CDW	sediment_WSW	6 : 23	-0.1	0.893
sediment_CDW	sediment_WSPW	6 : 6	-0.1	0.883
sediment_LSW	sediment_WSPW	9 : 6	0.1	0.902
sediment_LSW	sediment_MOW	9 : 3	0.0	0.975
sediment_LSW	sediment_WSW	9 : 23	0.2	0.823
sediment_MOW	sediment_WSPW	3 : 6	0.1	0.951
sediment_MOW	sediment_WSW	3 : 23	0.1	0.913
sediment_WSPW	sediment_WSW	6 : 23	0.1	0.959

Table S16 Results of statistical testing (Dunn’s tests) conducted to assess variations in microbial alpha diversity (Shannon index) between geographic locations for the different samplotypes.

Group1	Group2	Sample no.	Dun ns Z	p-value
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	23 : 11	0.1	0.885
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	23 : 52	-0.5	0.632
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Finnmark (MRGID:8962) Stjærnsund	23 : 17	0.3	0.734
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Colville (MRGID:5846) volcano	23 : 11	2.1	0.034
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Tromsøflaket (MRGID:5034) C	23 : 8	-0.6	0.528
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Gakkell Ridge (GAZ:00002436) Saddle between Seamounts	23 : 2	-0.3	0.727
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Davis Strait (MRGID:4250)	23 : 15	0.6	0.570
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Vesteris Seamount (MRGID:5046)	23 : 46	0.6	0.556
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Gakkell Ridge (GAZ:00002436) Central Seamount	23 : 1	-0.6	0.538
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Tromsøflaket (MRGID:5034) B	23 : 12	-0.2	0.877
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Devonport Seamount Chain (MRGID:26630) seamount114	23 : 1	-0.6	0.537
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Gakkell Ridge (GAZ:00002436) Karasik Seamount	23 : 17	-0.7	0.507
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA off Vesterålen (MRGID:18663) Bleiksdjupet	23 : 1	0.1	0.888
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Kermadec Ridge (MRGID:7213)	23 : 16	-0.9	0.380
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Otago (MRGID:26502) Slope	23 : 1	0.3	0.777
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Schulz Bank (MRGID:5039)	23 : 76	-0.8	0.436
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Raoul Island (MRGID:32903) Denham Bay (MRGID:58148)	23 : 2	0.7	0.509
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA off Hawke’s Bay (MRGID:32521) seamount986	23 : 2	0.5	0.592
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA off Troms (MRGID:23319) Malangsrunden	23 : 4	-0.4	0.686
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Gakkell Ridge (GAZ:00002436) Northern Seamount	23 : 5	0.0	0.969
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Sularevet (WDPaid:555557185)	23 : 9	0.5	0.599
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Tromsøflaket (MRGID:5034) A	23 : 17	-0.5	0.611
HMA Bømlø-Sotra archipelago (MRGID:5285)	HMA Hordaland (MRGID:18643) Korsfjorden	23 : 10	-0.6	0.563
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Davis Strait (MRGID:4250)	11 : 15	0.3	0.733
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Schulz Bank (MRGID:5039)	11 : 76	-0.7	0.459
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Colville (MRGID:5846) volcano	11 : 11	1.7	0.090
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Tromsøflaket (MRGID:5034) C	11 : 8	-0.7	0.502
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Kermadec Ridge (MRGID:7213)	11 : 16	-0.9	0.387
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	11 : 52	-0.5	0.603
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Gakkell Ridge (GAZ:00002436) Saddle between Seamounts	11 : 2	-0.4	0.686
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Gakkell Ridge (GAZ:00002436) Central Seamount	11 : 1	-0.7	0.514
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Tromsøflaket (MRGID:5034) A	11 : 17	-0.6	0.577
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA off Vesterålen (MRGID:18663) Bleiksdjupet	11 : 1	0.1	0.931
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Gakkell Ridge (GAZ:00002436) Northern Seamount	11 : 5	-0.1	0.893
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA off Hawke’s Bay (MRGID:32521) seamount986	11 : 2	0.4	0.656
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA off Troms (MRGID:23319) Malangsrunden	11 : 4	-0.5	0.641
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Hordaland (MRGID:18643) Korsfjorden	11 : 10	-0.6	0.533
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Otago (MRGID:26502) Slope	11 : 1	0.2	0.821
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Raoul Island (MRGID:32903) Denham Bay (MRGID:58148)	11 : 2	0.6	0.572
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Finnmark (MRGID:8962) Stjærnsund	11 : 17	0.1	0.886
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Sularevet (WDPaid:555557185)	11 : 9	0.3	0.732
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Devonport Seamount Chain (MRGID:26630) seamount114	11 : 1	-0.7	0.512
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Vesteris Seamount (MRGID:5046)	11 : 46	0.3	0.772
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Gakkell Ridge (GAZ:00002436) Karasik Seamount	11 : 17	-0.7	0.493
HMA Cantabrian (MRGID:3758) Sea Avilés Canyon (MRGID:36126)	HMA Tromsøflaket (MRGID:5034) B	11 : 12	-0.3	0.795
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Schulz Bank (MRGID:5039)	52 : 76	-0.4	0.715
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Davis Strait (MRGID:4250)	52 : 15	1.1	0.293
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Tromsøflaket (MRGID:5034) C	52 : 8	-0.4	0.714
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Gakkell Ridge (GAZ:00002436) Saddle between Seamounts	52 : 2	-0.2	0.848
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Tromsøflaket (MRGID:5034) A	52 : 17	-0.2	0.877
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA off Vesterålen (MRGID:18663) Bleiksdjupet	52 : 1	0.3	0.794
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Otago (MRGID:26502) Slope	52 : 1	0.4	0.685
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Colville (MRGID:5846) volcano	52 : 11	2.7	0.007
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Raoul Island (MRGID:32903) Denham Bay (MRGID:58148)	52 : 2	0.8	0.399
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Hordaland (MRGID:18643) Korsfjorden	52 : 10	-0.3	0.773
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA off Hawke’s Bay (MRGID:32521) seamount986	52 : 2	0.7	0.475
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Gakkell Ridge (GAZ:00002436) Central Seamount	52 : 1	-0.5	0.614
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Vesteris Seamount (MRGID:5046)	52 : 46	1.3	0.182
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA off Troms (MRGID:23319) Malangsrunden	52 : 4	-0.2	0.848
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Gakkell Ridge (GAZ:00002436) Karasik Seamount	52 : 17	-0.3	0.741
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Kermadec Ridge (MRGID:7213)	52 : 16	-0.6	0.561
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Tromsøflaket (MRGID:5034) B	52 : 12	0.2	0.840
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Finnmark (MRGID:8962) Stjærnsund	52 : 17	0.8	0.414
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Devonport Seamount Chain (MRGID:26630) seamount114	52 : 1	-0.5	0.612
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Sularevet (WDPaid:555557185)	52 : 9	0.9	0.366
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Gakkell Ridge (GAZ:00002436) Northern Seamount	52 : 5	0.2	0.830
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Schulz Bank (MRGID:5039)	11 : 76	-3.0	0.003
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Hordaland (MRGID:18643) Korsfjorden	11 : 10	-2.3	0.023
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Gakkell Ridge (GAZ:00002436) Saddle between Seamounts	11 : 2	-1.3	0.178
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Davis Strait (MRGID:4250)	11 : 15	-1.5	0.138
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Gakkell Ridge (GAZ:00002436) Central Seamount	11 : 1	-1.3	0.178
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Vesteris Seamount (MRGID:5046)	11 : 46	-1.9	0.062
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Tromsøflaket (MRGID:5034) C	11 : 8	-2.2	0.026
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Otago (MRGID:26502) Slope	11 : 1	-0.5	0.641
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA off Vesterålen (MRGID:18663) Bleiksdjupet	11 : 1	-0.6	0.544
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Tromsøflaket (MRGID:5034) B	11 : 12	-2.0	0.046
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Gakkell Ridge (GAZ:00002436) Karasik Seamount	11 : 17	-2.6	0.011
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Tromsøflaket (MRGID:5034) A	11 : 17	-2.4	0.015
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA off Hawke’s Bay (MRGID:32521) seamount986	11 : 2	-0.5	0.620
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Devonport Seamount Chain (MRGID:26630) seamount114	11 : 1	-1.3	0.178
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Raoul Island (MRGID:32903) Denham Bay (MRGID:58148)	11 : 2	-0.4	0.706
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Kermadec Ridge (MRGID:7213)	11 : 16	-2.7	0.007
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA off Troms (MRGID:23319) Malangsrunden	11 : 4	-1.7	0.088
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Gakkell Ridge (GAZ:00002436) Northern Seamount	11 : 5	-1.5	0.140
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Finnmark (MRGID:8962) Stjærnsund	11 : 17	-1.7	0.084
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Sularevet (WDPaid:555557185)	11 : 9	-1.3	0.205
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Gakkell Ridge (GAZ:00002436) Saddle between Seamounts	15 : 2	-0.6	0.553
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Gakkell Ridge (GAZ:00002436) Karasik Seamount	15 : 17	-1.1	0.258
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Otago (MRGID:26502) Slope	15 : 1	0.1	0.922
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Tromsøflaket (MRGID:5034) C	15 : 8	-1.0	0.306
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Tromsøflaket (MRGID:5034) A	15 : 17	-1.0	0.321
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Gakkell Ridge (GAZ:00002436) Central Seamount	15 : 1	-0.8	0.429
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA off Vesterålen (MRGID:18663) Bleiksdjupet	15 : 1	0.0	0.965
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Schulz Bank (MRGID:5039)	15 : 76	-1.3	0.185
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Hordaland (MRGID:18643) Korsfjorden	15 : 10	-1.0	0.318
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Kermadec Ridge (MRGID:7213)	15 : 16	-1.3	0.187
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA off Hawke’s Bay (MRGID:32521) seamount986	15 : 2	0.3	0.783
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Devonport Seamount Chain (MRGID:26630) seamount114	15 : 1	-0.8	0.427
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Raoul Island (MRGID:32903) Denham Bay (MRGID:58148)	15 : 2	0.4	0.691
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Finnmark (MRGID:8962) Stjærnsund	15 : 17	-0.2	0.821
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Vesteris Seamount (MRGID:5046)	15 : 46	-0.1	0.897
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA off Troms (MRGID:23319) Malangsrunden	15 : 4	-0.7	0.469
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Tromsøflaket (MRGID:5034) B	15 : 12	-0.6	0.529
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Gakkell Ridge (GAZ:00002436) Northern Seamount	15 : 5	-0.4	0.687
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Sularevet (WDPaid:555557185)	15 : 9	0.0	0.966
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Tromsøflaket (MRGID:5034) C	1 : 8	0.4	0.726
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Gakkell Ridge (GAZ:00002436) Karasik Seamount	1 : 17	0.4	0.684

LMA demo Kermadec Ridge (MRGID:7213)	LMA demo Sognefjorden (MRGID:3376)	5 : 6	-0.1	0.933
LMA demo Kermadec Ridge (MRGID:7213)	LMA demo Vesteris Seamount (MRGID:5046)	5 : 26	0.0	0.961
LMA demo Kermadec Ridge (MRGID:7213)	LMA demo Tromsøflaket (MRGID:5034) C	5 : 10	0.4	0.699
LMA demo Kermadec Ridge (MRGID:7213)	LMA demo off East Cape (MRGID:24347)	5 : 1	-0.7	0.499
LMA demo Kermadec Ridge (MRGID:7213)	LMA demo Kiwi Seamount (MRGID:7051)	5 : 3	0.3	0.768
LMA demo Kermadec Ridge (MRGID:7213)	LMA demo Rockall Bank (MRGID:4207)	5 : 4	0.6	0.572
LMA demo Kermadec Ridge (MRGID:7213)	LMA demo off Vesterålen (MRGID:18663) Bleiksdjupet	5 : 11	0.1	0.923
LMA demo Kermadec Ridge (MRGID:7213)	LMA demo Sularevet (WDPaid:555557185)	5 : 8	0.1	0.955
LMA demo Kermadec Ridge (MRGID:7213)	LMA demo North West of Orkney (MRGID:4520)	5 : 1	-0.2	0.834
LMA demo Kermadec Ridge (MRGID:7213)	LMA demo off Hawke's Bay (MRGID:32521) seamount986	5 : 6	0.9	0.389
LMA demo Kermadec Ridge (MRGID:7213)	LMA demo Schulz Bank (MRGID:5039)	5 : 26	0.4	0.664
LMA demo Kermadec Ridge (MRGID:7213)	LMA demo off East Cape (MRGID:24347) seamount1247	5 : 1	0.4	0.688
LMA demo Kermadec Ridge (MRGID:7213)	LMA demo Tromsøflaket (MRGID:5034) A	5 : 6	-0.1	0.939
LMA demo Kermadec Ridge (MRGID:7213)	LMA demo Otago (MRGID:26502) Slope	5 : 25	0.6	0.548
LMA demo Kermadec Ridge (MRGID:7213)	LMA demo Weddell Sea (MRGID:4330) Prince Gustav Channel (MRGID:9417)	5 : 9	-0.2	0.849
LMA demo Kermadec Ridge (MRGID:7213)	LMA demo Kerry Head (MRGID:28479) Reefs	5 : 2	0.3	0.728
LMA demo Kermadec Ridge (MRGID:7213)	LMA demo Tromsøflaket (MRGID:5034) B	5 : 10	-0.6	0.532
LMA demo Kerry Head (MRGID:28479) Reefs	LMA demo Tromsøflaket (MRGID:5034) C	2 : 10	-0.1	0.919
LMA demo Kerry Head (MRGID:28479) Reefs	LMA demo Kiwi Seamount (MRGID:7051)	2 : 3	-0.1	0.934
LMA demo Kerry Head (MRGID:28479) Reefs	LMA demo Rockall Bank (MRGID:4207)	2 : 4	0.1	0.919
LMA demo Kerry Head (MRGID:28479) Reefs	LMA demo Schulz Bank (MRGID:5039)	2 : 26	-0.1	0.914
LMA demo Kerry Head (MRGID:28479) Reefs	LMA demo off East Cape (MRGID:24347)	2 : 1	-0.8	0.400
LMA demo Kerry Head (MRGID:28479) Reefs	LMA demo North West of Orkney (MRGID:4520)	2 : 1	-0.4	0.671
LMA demo Kerry Head (MRGID:28479) Reefs	LMA demo Sognefjorden (MRGID:3376)	2 : 6	-0.4	0.675
LMA demo Kerry Head (MRGID:28479) Reefs	LMA demo off Hawke's Bay (MRGID:32521) seamount986	2 : 6	0.3	0.778
LMA demo Kerry Head (MRGID:28479) Reefs	LMA demo Vesteris Seamount (MRGID:5046)	2 : 26	-0.4	0.716
LMA demo Kerry Head (MRGID:28479) Reefs	LMA demo Tromsøflaket (MRGID:5034) B	2 : 10	-0.8	0.414
LMA demo Kerry Head (MRGID:28479) Reefs	LMA demo Sularevet (WDPaid:555557185)	2 : 8	-0.3	0.744
LMA demo Kerry Head (MRGID:28479) Reefs	LMA demo off Troms (MRGID:23319) Malanggrunnen	2 : 1	-0.7	0.488
LMA demo Kerry Head (MRGID:28479) Reefs	LMA demo off East Cape (MRGID:24347) seamount1247	2 : 1	0.1	0.903
LMA demo Kerry Head (MRGID:28479) Reefs	LMA demo off Vesterålen (MRGID:18663) Bleiksdjupet	2 : 11	-0.3	0.757
LMA demo Kerry Head (MRGID:28479) Reefs	LMA demo Otago (MRGID:26502) Slope	2 : 25	0.0	0.996
LMA demo Kerry Head (MRGID:28479) Reefs	LMA demo Tromsøflaket (MRGID:5034) A	2 : 6	-0.4	0.680
LMA demo Kerry Head (MRGID:28479) Reefs	LMA demo Weddell Sea (MRGID:4330) Prince Gustav Channel (MRGID:9417)	2 : 9	-0.5	0.611
LMA demo Kiwi Seamount (MRGID:7051)	LMA demo off East Cape (MRGID:24347) seamount1247	3 : 1	0.2	0.846
LMA demo Kiwi Seamount (MRGID:7051)	LMA demo off East Cape (MRGID:24347)	3 : 1	-0.8	0.408
LMA demo Kiwi Seamount (MRGID:7051)	LMA demo Sognefjorden (MRGID:3376)	3 : 6	-0.4	0.706
LMA demo Kiwi Seamount (MRGID:7051)	LMA demo off Hawke's Bay (MRGID:32521) seamount986	3 : 6	0.4	0.665
LMA demo Kiwi Seamount (MRGID:7051)	LMA demo Sularevet (WDPaid:555557185)	3 : 8	-0.3	0.787
LMA demo Kiwi Seamount (MRGID:7051)	LMA demo Rockall Bank (MRGID:4207)	3 : 4	0.2	0.831
LMA demo Kiwi Seamount (MRGID:7051)	LMA demo Tromsøflaket (MRGID:5034) C	3 : 10	0.0	0.995
LMA demo Kiwi Seamount (MRGID:7051)	LMA demo off Troms (MRGID:23319) Malanggrunnen	3 : 1	-0.7	0.503
LMA demo Kiwi Seamount (MRGID:7051)	LMA demo Schulz Bank (MRGID:5039)	3 : 26	0.0	0.995
LMA demo Kiwi Seamount (MRGID:7051)	LMA demo Tromsøflaket (MRGID:5034) A	3 : 6	-0.4	0.711
LMA demo Kiwi Seamount (MRGID:7051)	LMA demo North West of Orkney (MRGID:4520)	3 : 1	-0.4	0.700
LMA demo Kiwi Seamount (MRGID:7051)	LMA demo Otago (MRGID:26502) Slope	3 : 25	0.1	0.897
LMA demo Kiwi Seamount (MRGID:7051)	LMA demo Weddell Sea (MRGID:4330) Prince Gustav Channel (MRGID:9417)	3 : 9	-0.5	0.629
LMA demo Kiwi Seamount (MRGID:7051)	LMA demo Vesteris Seamount (MRGID:5046)	3 : 26	-0.3	0.753
LMA demo Kiwi Seamount (MRGID:7051)	LMA demo off Vesterålen (MRGID:18663) Bleiksdjupet	3 : 11	-0.3	0.802
LMA demo Kiwi Seamount (MRGID:7051)	LMA demo Tromsøflaket (MRGID:5034) B	3 : 10	-0.8	0.397
LMA demo North West of Orkney (MRGID:4520)	LMA demo Schulz Bank (MRGID:5039)	1 : 26	0.4	0.665
LMA demo North West of Orkney (MRGID:4520)	LMA demo Sognefjorden (MRGID:3376)	1 : 6	0.2	0.869
LMA demo North West of Orkney (MRGID:4520)	LMA demo off Hawke's Bay (MRGID:32521) seamount986	1 : 6	0.7	0.447
LMA demo North West of Orkney (MRGID:4520)	LMA demo Tromsøflaket (MRGID:5034) C	1 : 10	0.4	0.674
LMA demo North West of Orkney (MRGID:4520)	LMA demo off East Cape (MRGID:24347)	1 : 1	-0.4	0.718
LMA demo North West of Orkney (MRGID:4520)	LMA demo off Troms (MRGID:23319) Malanggrunnen	1 : 1	-0.2	0.816
LMA demo North West of Orkney (MRGID:4520)	LMA demo Weddell Sea (MRGID:4330) Prince Gustav Channel (MRGID:9417)	1 : 9	0.1	0.908
LMA demo North West of Orkney (MRGID:4520)	LMA demo Vesteris Seamount (MRGID:5046)	1 : 26	0.2	0.804
LMA demo North West of Orkney (MRGID:4520)	LMA demo Otago (MRGID:26502) Slope	1 : 25	0.5	0.608
LMA demo North West of Orkney (MRGID:4520)	LMA demo off East Cape (MRGID:24347) seamount1247	1 : 1	0.5	0.636
LMA demo North West of Orkney (MRGID:4520)	LMA demo Tromsøflaket (MRGID:5034) A	1 : 6	0.2	0.866
LMA demo North West of Orkney (MRGID:4520)	LMA demo Rockall Bank (MRGID:4207)	1 : 4	0.5	0.587
LMA demo North West of Orkney (MRGID:4520)	LMA demo Sularevet (WDPaid:555557185)	1 : 8	0.2	0.805
LMA demo North West of Orkney (MRGID:4520)	LMA demo off Vesterålen (MRGID:18663) Bleiksdjupet	1 : 11	0.3	0.788
LMA demo North West of Orkney (MRGID:4520)	LMA demo Tromsøflaket (MRGID:5034) B	1 : 10	-0.1	0.914
LMA demo off East Cape (MRGID:24347)	LMA demo off Troms (MRGID:23319) Malanggrunnen	1 : 1	0.1	0.898
LMA demo off East Cape (MRGID:24347)	LMA demo Weddell Sea (MRGID:4330) Prince Gustav Channel (MRGID:9417)	1 : 9	0.6	0.548
LMA demo off East Cape (MRGID:24347)	LMA demo Sularevet (WDPaid:555557185)	1 : 8	0.7	0.467
LMA demo off East Cape (MRGID:24347)	LMA demo Otago (MRGID:26502) Slope	1 : 25	1.0	0.310
LMA demo off East Cape (MRGID:24347)	LMA demo Sognefjorden (MRGID:3376)	1 : 6	0.6	0.524
LMA demo off East Cape (MRGID:24347)	LMA demo Schulz Bank (MRGID:5039)	1 : 26	0.9	0.350
LMA demo off East Cape (MRGID:24347)	LMA demo off Hawke's Bay (MRGID:32521) seamount986	1 : 6	1.2	0.243
LMA demo off East Cape (MRGID:24347)	LMA demo Tromsøflaket (MRGID:5034) C	1 : 10	0.9	0.364
LMA demo off East Cape (MRGID:24347)	LMA demo Vesteris Seamount (MRGID:5046)	1 : 26	0.7	0.453
LMA demo off East Cape (MRGID:24347)	LMA demo off East Cape (MRGID:24347) seamount1247	1 : 1	0.8	0.404
LMA demo off East Cape (MRGID:24347)	LMA demo Tromsøflaket (MRGID:5034) B	1 : 10	0.4	0.705
LMA demo off East Cape (MRGID:24347)	LMA demo Tromsøflaket (MRGID:5034) A	1 : 6	0.6	0.521
LMA demo off East Cape (MRGID:24347)	LMA demo Rockall Bank (MRGID:4207)	1 : 4	1.0	0.317
LMA demo off East Cape (MRGID:24347)	LMA demo off Vesterålen (MRGID:18663) Bleiksdjupet	1 : 11	0.8	0.448
LMA demo off East Cape (MRGID:24347) seamount1247	LMA demo off Troms (MRGID:23319) Malanggrunnen	1 : 1	-0.7	0.480
LMA demo off East Cape (MRGID:24347) seamount1247	LMA demo Schulz Bank (MRGID:5039)	1 : 26	-0.2	0.823
LMA demo off East Cape (MRGID:24347) seamount1247	LMA demo Sognefjorden (MRGID:3376)	1 : 6	-0.5	0.649
LMA demo off East Cape (MRGID:24347) seamount1247	LMA demo Weddell Sea (MRGID:4330) Prince Gustav Channel (MRGID:9417)	1 : 9	-0.5	0.604
LMA demo off East Cape (MRGID:24347) seamount1247	LMA demo Otago (MRGID:26502) Slope	1 : 25	-0.1	0.886
LMA demo off East Cape (MRGID:24347) seamount1247	LMA demo off Hawke's Bay (MRGID:32521) seamount986	1 : 6	0.1	0.940
LMA demo off East Cape (MRGID:24347) seamount1247	LMA demo Tromsøflaket (MRGID:5034) A	1 : 6	-0.5	0.652
LMA demo off East Cape (MRGID:24347) seamount1247	LMA demo Sularevet (WDPaid:555557185)	1 : 8	-0.4	0.700
LMA demo off East Cape (MRGID:24347) seamount1247	LMA demo Vesteris Seamount (MRGID:5046)	1 : 26	-0.4	0.683
LMA demo off East Cape (MRGID:24347) seamount1247	LMA demo off Vesterålen (MRGID:18663) Bleiksdjupet	1 : 11	-0.4	0.710
LMA demo off East Cape (MRGID:24347) seamount1247	LMA demo Tromsøflaket (MRGID:5034) C	1 : 10	-0.2	0.827
LMA demo off East Cape (MRGID:24347) seamount1247	LMA demo Tromsøflaket (MRGID:5034) B	1 : 10	-0.7	0.456
LMA demo off East Cape (MRGID:24347) seamount1247	LMA demo Rockall Bank (MRGID:4207)	1 : 4	-0.1	0.956
LMA demo off Hawke's Bay (MRGID:32521) seamount986	LMA demo Sognefjorden (MRGID:3376)	6 : 6	-1.0	0.321
LMA demo off Hawke's Bay (MRGID:32521) seamount986	LMA demo Otago (MRGID:26502) Slope	6 : 25	-0.5	0.618
LMA demo off Hawke's Bay (MRGID:32521) seamount986	LMA demo off Troms (MRGID:23319) Malanggrunnen	6 : 1	-1.0	0.317
LMA demo off Hawke's Bay (MRGID:32521) seamount986	LMA demo Vesteris Seamount (MRGID:5046)	6 : 26	-1.1	0.272
LMA demo off Hawke's Bay (MRGID:32521) seamount986	LMA demo Sularevet (WDPaid:555557185)	6 : 8	-0.9	0.365
LMA demo off Hawke's Bay (MRGID:32521) seamount986	LMA demo Weddell Sea (MRGID:4330) Prince Gustav Channel (MRGID:9417)	6 : 9	-1.2	0.234
LMA demo off Hawke's Bay (MRGID:32521) seamount986	LMA demo Schulz Bank (MRGID:5039)	6 : 26	-0.7	0.494
LMA demo off Hawke's Bay (MRGID:32521) seamount986	LMA demo Tromsøflaket (MRGID:5034) B	6 : 10	-1.7	0.094
LMA demo off Hawke's Bay (MRGID:32521) seamount986	LMA demo Rockall Bank (MRGID:4207)	6 : 4	-0.2	0.825
LMA demo off Hawke's Bay (MRGID:32521) seamount986	LMA demo Tromsøflaket (MRGID:5034) A	6 : 6	-1.0	0.325
LMA demo off Hawke's Bay (MRGID:32521) seamount986	LMA demo off Vesterålen (MRGID:18663) Bleiksdjupet	6 : 11	-0.9	0.355
LMA demo off Hawke's Bay (MRGID:32521) seamount986	LMA demo Tromsøflaket (MRGID:5034) C	6 : 10	-0.6	0.549
LMA demo off Troms (MRGID:23319) Malanggrunnen	LMA demo Sognefjorden (MRGID:3376)	1 : 6	0.5	0.638
LMA demo off Troms (MRGID:23319) Malanggrunnen	LMA demo Otago (MRGID:26502) Slope	1 : 25	0.8	0.403
LMA demo off Troms (MRGID:23319) Malanggrunnen	LMA demo Weddell Sea (MRGID:4330) Prince Gustav Channel (MRGID:9417)	1 : 9	0.4	0.668
LMA demo off Troms (MRGID:23319) Malanggrunnen	LMA demo Tromsøflaket (MRGID:5034) B	1 : 10	0.2	0.837
LMA demo off Troms (MRGID:23319) Malanggrunnen	LMA demo Tromsøflaket (MRGID:5034) A	1 : 6	0.5	0.635
LMA demo off Troms (MRGID:23319) Malanggrunnen	LMA demo Sularevet (WDPaid:555557185)	1 : 8	0.6	0.577
LMA demo off Troms (MRGID:23319) Malanggrunnen	LMA demo Vesteris Seamount (MRGID:5046)	1 : 26	0.6	0.567
LMA demo off Troms (MRGID:23319) Malanggrunnen	LMA demo off Vesterålen (MRGID:18663) Bleiksdjupet	1 : 11	0.6	0.559
LMA demo off Troms (MRGID:23319) Malanggrunnen	LMA demo Tromsøflaket (MRGID:5034) C	1 : 10	0.7	0.463
LMA demo off Troms (MRGID:23319) Malanggrunnen	LMA demo Schulz Bank (MRGID:5039)	1 : 26	0.8	0.450
LMA demo off Troms (MRGID:23319) Malanggrunnen	LMA demo Rockall Bank (MRGID:4207)	1 : 4	0.8	0.402
LMA demo off Vesterålen (MRGID:18663) Bleiksdjupet	LMA demo Sognefjorden (MRGID:3376)	11 : 6	-0.2	0.838

LMA glass Kermadec Trench (MRGID:7571)	LMA glass Tromsøflaket (MRGID:5034) C	3	8	0.5	0.653
LMA glass Kermadec Trench (MRGID:7571)	LMA glass Sognefjorden (MRGID:3376)	3	2	0.8	0.412
LMA glass Kermadec Trench (MRGID:7571)	LMA glass Kiwi Seamount (MRGID:7051)	3	4	1.8	0.078
LMA glass Kermadec Trench (MRGID:7571)	LMA glass off Vesterålen (MRGID:18663) Bleiksdjupet	3	5	1.1	0.266
LMA glass Kermadec Trench (MRGID:7571)	LMA glass Sularevet (WDPaid:55557185)	3	1	0.7	0.490
LMA glass Kermadec Trench (MRGID:7571)	LMA glass Vesteris Seamount (MRGID:5046)	3	23	1.2	0.226
LMA glass Kermadec Trench (MRGID:7571)	LMA glass off Pico (MRGID:2460)	3	5	1.6	0.118
LMA glass Kiwi Seamount (MRGID:7051)	LMA glass Schulz Bank (MRGID:5039)	4	33	-1.1	0.275
LMA glass Kiwi Seamount (MRGID:7051)	LMA glass off East Cape (MRGID:24347) seamount1247	4	5	0.2	0.826
LMA glass Kiwi Seamount (MRGID:7051)	LMA glass off Hawke's Bay (MRGID:32521) seamount986	4	27	-0.7	0.514
LMA glass Kiwi Seamount (MRGID:7051)	LMA glass rift valley deep volcano off Schulz Bank (MRGID:5039)	4	3	-0.3	0.760
LMA glass Kiwi Seamount (MRGID:7051)	LMA glass Sambro Bank (MRGID:34408)	4	24	-1.4	0.171
LMA glass Kiwi Seamount (MRGID:7051)	LMA glass Weddell Sea (MRGID:4330) Filchner-Ronne (MRGID:17335) Outflow System	4	4	0.4	0.688
LMA glass Kiwi Seamount (MRGID:7051)	LMA glass Tromsøflaket (MRGID:5034) C	4	8	-1.7	0.089
LMA glass Kiwi Seamount (MRGID:7051)	LMA glass Wairapa (MRGID:34834) Slope	4	5	0.1	0.891
LMA glass Kiwi Seamount (MRGID:7051)	LMA glass Sognefjorden (MRGID:3376)	4	2	-0.7	0.489
LMA glass Kiwi Seamount (MRGID:7051)	LMA glass off Pico (MRGID:2460)	4	5	-0.3	0.760
LMA glass Kiwi Seamount (MRGID:7051)	LMA glass Sularevet (WDPaid:55557185)	4	1	-0.5	0.623
LMA glass Kiwi Seamount (MRGID:7051)	LMA glass Vesteris Seamount (MRGID:5046)	4	23	-1.1	0.264
LMA glass Kiwi Seamount (MRGID:7051)	LMA glass off Vesterålen (MRGID:18663) Bleiksdjupet	4	5	-0.8	0.424
LMA glass off East Cape (MRGID:24347) seamount1247	LMA glass off Hawke's Bay (MRGID:32521) seamount986	5	27	-1.0	0.307
LMA glass off East Cape (MRGID:24347) seamount1247	LMA glass Sognefjorden (MRGID:3376)	5	2	-0.9	0.372
LMA glass off East Cape (MRGID:24347) seamount1247	LMA glass Weddell Sea (MRGID:4330) Filchner-Ronne (MRGID:17335) Outflow System	5	4	0.2	0.838
LMA glass off East Cape (MRGID:24347) seamount1247	LMA glass Sambro Bank (MRGID:34408)	5	24	-1.8	0.071
LMA glass off East Cape (MRGID:24347) seamount1247	LMA glass rift valley deep volcano off Schulz Bank (MRGID:5039)	5	3	-0.5	0.602
LMA glass off East Cape (MRGID:24347) seamount1247	LMA glass Schulz Bank (MRGID:5039)	5	33	-1.5	0.131
LMA glass off East Cape (MRGID:24347) seamount1247	LMA glass off Pico (MRGID:2460)	5	5	-0.6	0.578
LMA glass off East Cape (MRGID:24347) seamount1247	LMA glass Sularevet (WDPaid:55557185)	5	1	-0.6	0.525
LMA glass off East Cape (MRGID:24347) seamount1247	LMA glass Wairapa (MRGID:34834) Slope	5	5	-0.1	0.930
LMA glass off East Cape (MRGID:24347) seamount1247	LMA glass Tromsøflaket (MRGID:5034) C	5	8	-2.1	0.037
LMA glass off East Cape (MRGID:24347) seamount1247	LMA glass off Vesterålen (MRGID:18663) Bleiksdjupet	5	5	-1.1	0.280
LMA glass off East Cape (MRGID:24347) seamount1247	LMA glass Vesteris Seamount (MRGID:5046)	5	23	-1.5	0.128
LMA glass off Hawke's Bay (MRGID:32521) seamount986	LMA glass Schulz Bank (MRGID:5039)	27	33	-0.9	0.379
LMA glass off Hawke's Bay (MRGID:32521) seamount986	LMA glass Wairapa (MRGID:34834) Slope	27	5	0.9	0.364
LMA glass off Hawke's Bay (MRGID:32521) seamount986	LMA glass Sognefjorden (MRGID:3376)	27	2	-0.3	0.734
LMA glass off Hawke's Bay (MRGID:32521) seamount986	LMA glass Sambro Bank (MRGID:34408)	27	24	-1.4	0.164
LMA glass off Hawke's Bay (MRGID:32521) seamount986	LMA glass Sularevet (WDPaid:55557185)	27	1	-0.2	0.844
LMA glass off Hawke's Bay (MRGID:32521) seamount986	LMA glass Weddell Sea (MRGID:4330) Filchner-Ronne (MRGID:17335) Outflow System	27	4	1.2	0.236
LMA glass off Hawke's Bay (MRGID:32521) seamount986	LMA glass rift valley deep volcano off Schulz Bank (MRGID:5039)	27	3	0.2	0.848
LMA glass off Hawke's Bay (MRGID:32521) seamount986	LMA glass Vesteris Seamount (MRGID:5046)	27	23	-0.9	0.369
LMA glass off Hawke's Bay (MRGID:32521) seamount986	LMA glass Tromsøflaket (MRGID:5034) C	27	8	-1.7	0.085
LMA glass off Hawke's Bay (MRGID:32521) seamount986	LMA glass off Vesterålen (MRGID:18663) Bleiksdjupet	27	5	-0.4	0.703
LMA glass off Hawke's Bay (MRGID:32521) seamount986	LMA glass off Pico (MRGID:2460)	27	5	0.3	0.765
LMA glass off Pico (MRGID:2460)	LMA glass rift valley deep volcano off Schulz Bank (MRGID:5039)	5	3	0.0	0.968
LMA glass off Pico (MRGID:2460)	LMA glass Schulz Bank (MRGID:5039)	5	33	-0.8	0.436
LMA glass off Pico (MRGID:2460)	LMA glass Wairapa (MRGID:34834) Slope	5	5	0.5	0.639
LMA glass off Pico (MRGID:2460)	LMA glass Sognefjorden (MRGID:3376)	5	2	-0.5	0.667
LMA glass off Pico (MRGID:2460)	LMA glass Weddell Sea (MRGID:4330) Filchner-Ronne (MRGID:17335) Outflow System	5	4	0.7	0.466
LMA glass off Pico (MRGID:2460)	LMA glass Tromsøflaket (MRGID:5034) C	5	8	-1.5	0.141
LMA glass off Pico (MRGID:2460)	LMA glass Sambro Bank (MRGID:34408)	5	24	-1.1	0.276
LMA glass off Pico (MRGID:2460)	LMA glass Sularevet (WDPaid:55557185)	5	1	-0.3	0.753
LMA glass off Pico (MRGID:2460)	LMA glass Vesteris Seamount (MRGID:5046)	5	23	-0.8	0.417
LMA glass off Pico (MRGID:2460)	LMA glass off Vesterålen (MRGID:18663) Bleiksdjupet	5	5	-0.5	0.601
LMA glass off Pico (MRGID:2460)	LMA glass Sognefjorden (MRGID:3376)	5	2	-0.1	0.940
LMA glass off Vesterålen (MRGID:18663) Bleiksdjupet	LMA glass rift valley deep volcano off Schulz Bank (MRGID:5039)	5	3	0.4	0.679
LMA glass off Vesterålen (MRGID:18663) Bleiksdjupet	LMA glass Schulz Bank (MRGID:5039)	5	33	-0.1	0.930
LMA glass off Vesterålen (MRGID:18663) Bleiksdjupet	LMA glass Weddell Sea (MRGID:4330) Filchner-Ronne (MRGID:17335) Outflow System	5	4	1.2	0.221
LMA glass off Vesterålen (MRGID:18663) Bleiksdjupet	LMA glass Wairapa (MRGID:34834) Slope	5	5	1.0	0.321
LMA glass off Vesterålen (MRGID:18663) Bleiksdjupet	LMA glass Tromsøflaket (MRGID:5034) C	5	8	-0.9	0.374
LMA glass off Vesterålen (MRGID:18663) Bleiksdjupet	LMA glass Sularevet (WDPaid:55557185)	5	1	0.0	0.990
LMA glass off Vesterålen (MRGID:18663) Bleiksdjupet	LMA glass Vesteris Seamount (MRGID:5046)	5	23	-0.1	0.889
LMA glass off Vesterålen (MRGID:18663) Bleiksdjupet	LMA glass Sambro Bank (MRGID:34408)	5	24	-0.4	0.678
LMA glass rift valley deep volcano off Schulz Bank (MRGID:5039)	LMA glass Sognefjorden (MRGID:3376)	3	2	-0.4	0.689
LMA glass rift valley deep volcano off Schulz Bank (MRGID:5039)	LMA glass Wairapa (MRGID:34834) Slope	3	5	0.4	0.656
LMA glass rift valley deep volcano off Schulz Bank (MRGID:5039)	LMA glass Tromsøflaket (MRGID:5034) C	3	8	-1.2	0.232
LMA glass rift valley deep volcano off Schulz Bank (MRGID:5039)	LMA glass Sularevet (WDPaid:55557185)	3	1	-0.3	0.784
LMA glass rift valley deep volcano off Schulz Bank (MRGID:5039)	LMA glass Weddell Sea (MRGID:4330) Filchner-Ronne (MRGID:17335) Outflow System	3	4	0.7	0.498
LMA glass rift valley deep volcano off Schulz Bank (MRGID:5039)	LMA glass Schulz Bank (MRGID:5039)	3	33	-0.6	0.568
LMA glass rift valley deep volcano off Schulz Bank (MRGID:5039)	LMA glass Sambro Bank (MRGID:34408)	3	24	-0.8	0.408
LMA glass rift valley deep volcano off Schulz Bank (MRGID:5039)	LMA glass Vesteris Seamount (MRGID:5046)	3	23	-0.6	0.546
LMA glass Sambro Bank (MRGID:34408)	LMA glass Sognefjorden (MRGID:3376)	24	2	0.2	0.848
LMA glass Sambro Bank (MRGID:34408)	LMA glass Weddell Sea (MRGID:4330) Filchner-Ronne (MRGID:17335) Outflow System	24	4	1.9	0.058
LMA glass Sambro Bank (MRGID:34408)	LMA glass Wairapa (MRGID:34834) Slope	24	5	1.7	0.091
LMA glass Sambro Bank (MRGID:34408)	LMA glass Tromsøflaket (MRGID:5034) C	24	8	-0.7	0.458
LMA glass Sambro Bank (MRGID:34408)	LMA glass Sularevet (WDPaid:55557185)	24	1	0.2	0.852
LMA glass Sambro Bank (MRGID:34408)	LMA glass Schulz Bank (MRGID:5039)	24	33	0.6	0.546
LMA glass Sambro Bank (MRGID:34408)	LMA glass Vesteris Seamount (MRGID:5046)	24	23	0.5	0.643
LMA glass Schulz Bank (MRGID:5039)	LMA glass Weddell Sea (MRGID:4330) Filchner-Ronne (MRGID:17335) Outflow System	33	4	1.6	0.103
LMA glass Schulz Bank (MRGID:5039)	LMA glass Wairapa (MRGID:34834) Slope	33	5	1.4	0.163
LMA glass Schulz Bank (MRGID:5039)	LMA glass Tromsøflaket (MRGID:5034) C	33	8	-1.2	0.238
LMA glass Schulz Bank (MRGID:5039)	LMA glass Sognefjorden (MRGID:3376)	33	2	0.0	0.977
LMA glass Schulz Bank (MRGID:5039)	LMA glass Vesteris Seamount (MRGID:5046)	33	23	-0.1	0.922
LMA glass Schulz Bank (MRGID:5039)	LMA glass Sularevet (WDPaid:55557185)	33	1	0.0	0.978
LMA glass Sognefjorden (MRGID:3376)	LMA glass Wairapa (MRGID:34834) Slope	2	5	0.8	0.409
LMA glass Sognefjorden (MRGID:3376)	LMA glass Tromsøflaket (MRGID:5034) C	2	8	-0.6	0.574
LMA glass Sognefjorden (MRGID:3376)	LMA glass Sularevet (WDPaid:55557185)	2	1	0.0	0.968
LMA glass Sognefjorden (MRGID:3376)	LMA glass Vesteris Seamount (MRGID:5046)	2	23	0.0	0.994
LMA glass Sognefjorden (MRGID:3376)	LMA glass Weddell Sea (MRGID:4330) Filchner-Ronne (MRGID:17335) Outflow System	2	4	1.0	0.308
LMA glass Sularevet (WDPaid:55557185)	LMA glass Vesteris Seamount (MRGID:5046)	1	23	-0.1	0.957
LMA glass Sularevet (WDPaid:55557185)	LMA glass Tromsøflaket (MRGID:5034) C	1	8	-0.5	0.642
LMA glass Sularevet (WDPaid:55557185)	LMA glass Weddell Sea (MRGID:4330) Filchner-Ronne (MRGID:17335) Outflow System	1	4	0.7	0.456
LMA glass Sularevet (WDPaid:55557185)	LMA glass Wairapa (MRGID:34834) Slope	1	5	0.6	0.558
LMA glass Tromsøflaket (MRGID:5034) C	LMA glass Weddell Sea (MRGID:4330) Filchner-Ronne (MRGID:17335) Outflow System	8	4	2.2	0.030
LMA glass Tromsøflaket (MRGID:5034) C	LMA glass Vesteris Seamount (MRGID:5046)	8	23	1.1	0.286
LMA glass Tromsøflaket (MRGID:5034) C	LMA glass Wairapa (MRGID:34834) Slope	8	5	2.0	0.047
LMA glass Vesteris Seamount (MRGID:5046)	LMA glass Wairapa (MRGID:34834) Slope	23	5	1.4	0.158
LMA glass Vesteris Seamount (MRGID:5046)	LMA glass Weddell Sea (MRGID:4330) Filchner-Ronne (MRGID:17335) Outflow System	23	4	1.6	0.101
LMA glass Wairarapa (MRGID:34834) Slope	LMA glass Weddell Sea (MRGID:4330) Filchner-Ronne (MRGID:17335) Outflow System	5	4	0.3	0.774
seawater Bømlø-Sotra archipelago (MRGID:5285)	seawater Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	2	22	0.0	0.962
seawater Bømlø-Sotra archipelago (MRGID:5285)	seawater Sularevet (WDPaid:55557185)	2	4	-0.1	0.938
seawater Bømlø-Sotra archipelago (MRGID:5285)	seawater Christchurch (MRGID:32447) Canyon/Slope	2	3	-0.3	0.754
seawater Bømlø-Sotra archipelago (MRGID:5285)	seawater off Troms (MRGID:2319) Malangsrunden	2	3	0.1	0.948
seawater Bømlø-Sotra archipelago (MRGID:5285)	seawater Fram Strait (MRGID:26579) LTER Observatory HAUSGARTEN site53	2	4	0.5	0.611
seawater Bømlø-Sotra archipelago (MRGID:5285)	seawater Colville (MRGID:5846) volcano	2	3	1.2	0.233
seawater Bømlø-Sotra archipelago (MRGID:5285)	seawater Weddell Sea (MRGID:4330) Prince Gustav Channel (MRGID:9417)	2	4	1.6	0.112
seawater Bømlø-Sotra archipelago (MRGID:5285)	seawater Finnmark (MRGID:8962) Stjensund	2	9	0.5	0.644
seawater Bømlø-Sotra archipelago (MRGID:5285)	seawater off Pico (MRGID:2460)	2	6	-0.1	0.913
seawater Bømlø-Sotra archipelago (MRGID:5285)	seawater Devonport Seamount Chain (MRGID:26630) seamount114	2	3	2.1	0.035
seawater Bømlø-Sotra archipelago (MRGID:5285)	seawater Tromsøflaket (MRGID:5034) B	2	9	0.2	0.876
seawater Bømlø-Sotra archipelago (MRGID:5285)	seawater off Vesterålen (MRGID:18663) Bleiksdjupet	2	9	0.9	0.380
seawater Bømlø-Sotra archipelago (MRGID:5285)	seawater Gakkel Ridge (GAZ:00002436) Central Seamount	2	1	1.1	0.281
seawater Bømlø-Sotra archipelago (MRGID:5285)	seawater Sognefjorden (MRGID:3376)	2	6	-0.1	0.898
seawater Bømlø-Sotra archipelago (MRGID:5285)	seawater Emerald Basin (MRGID:34188)	2	5	0.0	0.960
seawater Bømlø-Sotra archipelago (MRGID:5285)	seawater Kiwi Seamount (MRGID:7051)	2	6	1.0	0.302
seawater Bømlø-Sotra archipelago (MRGID:5285)	seawater Tromsøflaket (MRGID:5034) C	2	12	0.4	0.675

seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	1.8	0.077
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	1.6	0.107
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	2	1.5	0.129
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	0.6	0.568
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	11	0.3	0.774
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	4	0.3	0.734
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	4	0.4	0.714
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	6	2.2	0.025
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	140	0.7	0.465
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	2.4	0.015
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	0.9	0.350
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	6	2.1	0.034
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	1	0.3	0.728
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	0.0	0.982
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	14	0.5	0.591
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	2.2	0.028
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	4	1.8	0.070
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	1	0.1	0.954
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	6	1.5	0.133
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	1.8	0.068
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	2.4	0.014
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	9	1.0	0.317
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	11	0.7	0.489
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	3.7	0.000
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	9	1.8	0.069
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	1	0.4	0.653
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	1	1.3	0.185
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	2.7	0.007
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	1.4	0.149
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	0.4	0.683
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	0.9	0.366
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	4	2.6	0.010
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	4	0.9	0.382
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	24	-0.1	0.940
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	9	0.4	0.693
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	3.2	0.001
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	6	-0.2	0.880
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	0.0	0.982
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	6	1.9	0.057
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	12	1.0	0.322
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	6	4.1	0.000
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	4	0.6	0.545
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	6	2.7	0.006
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	0.2	0.878
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	6	3.8	0.000
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	14	1.3	0.197
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	140	2.4	0.016
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	2	2.1	0.036
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	3.3	0.001
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	1	0.1	0.918
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	4	0.6	0.517
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	5	0.2	0.877
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	4	2.9	0.003
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	6	-0.1	0.906
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	4	0.0	0.982
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	9	1.9	0.057
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	4	0.9	0.357
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	9	1.1	0.285
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	1.9	0.061
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	1	1.4	0.177
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	1.5	0.137
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	6	4.1	0.000
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	11	0.8	0.445
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	0.1	0.953
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	3.7	0.000
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	2.5	0.012
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	1	0.5	0.636
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	3.2	0.001
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	12	1.1	0.286
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	4	0.7	0.515
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	-0.4	0.709
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	4	2.7	0.008
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	2.7	0.006
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	4	0.0	0.985
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	0.2	0.849
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	9	0.5	0.647
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	1	0.1	0.900
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	6	-0.1	0.945
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	6	2.8	0.005
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	4	3.0	0.003
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	6	-0.1	0.917
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	14	1.4	0.168
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	0.9	0.344
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	6	3.9	0.000
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	140	2.6	0.009
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	3	3.4	0.001
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	6	2.0	0.048
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	5	0.2	0.841
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	2	2.1	0.032
seawater Bømlø-Sotra archipelago (MRGID:5285)	2	4	0.7	0.488
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	4	1.0	0.342
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	3	0.3	0.745
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	9	1.0	0.332
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	9	0.6	0.541
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	6	3.0	0.003
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	12	0.9	0.347
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	3	1.7	0.092
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	11	0.8	0.436
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	6	0.3	0.797
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	3	2.3	0.020
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	3	1.0	0.322
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	3	0.4	0.672
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	4	2.2	0.030
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	1	0.3	0.758
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	3	2.7	0.007
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	3	3.1	0.002
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	1	0.6	0.538
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	6	1.6	0.110
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	14	1.1	0.277
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	4	0.3	0.754
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	4	0.8	0.447
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	3	2.2	0.031
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	3	2.8	0.005
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	6	0.3	0.780
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	9	1.5	0.145
seawater Bømlø-Sotra archipelago (MRGID:5285)	3	2	2.0	0.048
seawater Raoul Island (MRGID:32903) Denham Bay (MRGID:58148)	2	3	1.8	0.077
seawater off East Cape (MRGID:24347)	2	3	1.6	0.107
seawater Weddell Sea (MRGID:4330) Filchner-Ronne (MRGID:17335) Outflow System	2	2	1.5	0.129
seawater Otago (MRGID:26502) Slope	2	3	0.6	0.568
seawater Tromsøflaket (MRGID:5034) A	2	11	0.3	0.774
seawater Fram Strait (MRGID:26579) LTER Observatory HAUSGARTEN siteN4	2	4	0.3	0.734
seawater Fram Strait (MRGID:26579) LTER Observatory HAUSGARTEN siteHG4	2	4	0.4	0.714
seawater Kermadec Ridge (MRGID:7213)	2	6	2.2	0.025
seawater Schulz Bank (MRGID:5039)	2	140	0.7	0.465
seawater Kermadec Trench (MRGID:7571)	2	3	2.4	0.015
seawater Wairarapa (MRGID:34834) Slope	2	3	0.9	0.350
seawater off Hawke's Bay (MRGID:32521) seamount986	2	6	2.1	0.034
seawater Gakkell Ridge (GAZ:00002436) Karasik Seamount	2	1	0.3	0.728
seawater Tromsøflaket (MRGID:5034)	2	3	0.0	0.982
seawater Sambro Bank (MRGID:34408)	2	14	0.5	0.591
seawater off East Cape (MRGID:24347) seamount1247	2	3	2.2	0.028
seawater Fram Strait (MRGID:26579) LTER Observatory HAUSGARTEN siteSV2	2	4	1.8	0.070
seawater Gakkell Ridge (GAZ:00002436) Northern Seamount	2	1	0.1	0.954
seawater Vesteris Seamount (MRGID:5046)	2	6	1.5	0.133
seawater Colville (MRGID:5846) volcano	2	3	1.8	0.068
seawater off East Cape (MRGID:24347)	2	3	2.4	0.014
seawater Finnmark (MRGID:8962) Stjernsund	2	9	1.0	0.317
seawater Tromsøflaket (MRGID:5034) A	2	11	0.7	0.489
seawater Kermadec Trench (MRGID:7571)	2	3	3.7	0.000
seawater off Vesterålen (MRGID:18663) Bleiksdjupet	2	9	1.8	0.069
seawater Gakkell Ridge (GAZ:00002436) Karasik Seamount	2	1	0.4	0.653
seawater Gakkell Ridge (GAZ:00002436) Central Seamount	2	1	1.3	0.185
seawater Raoul Island (MRGID:32903) Denham Bay (MRGID:58148)	2	3	2.7	0.007
seawater Wairarapa (MRGID:34834) Slope	2	3	1.4	0.149
seawater Christchurch (MRGID:32447) Canyon/Slope	2	3	0.4	0.683
seawater Otago (MRGID:26502) Slope	2	3	0.9	0.366
seawater Weddell Sea (MRGID:4330) Prince Gustav Channel (MRGID:9417)	2	4	2.6	0.010
seawater Fram Strait (MRGID:26579) LTER Observatory HAUSGARTEN siteS3	2	4	0.9	0.382
seawater Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	2	24	-0.1	0.940
seawater Tromsøflaket (MRGID:5034) B	2	9	0.4	0.693
seawater Devonport Seamount Chain (MRGID:26630) seamount114	2	3	3.2	0.001
seawater Sognefjorden (MRGID:3376)	2	6	-0.2	0.880
seawater Tromsøflaket (MRGID:5034)	2	3	0.0	0.982
seawater Kiwi Seamount (MRGID:7051)	2	6	1.9	0.057
seawater Tromsøflaket (MRGID:5034) C	2	12	1.0	0.322
seawater Kermadec Ridge (MRGID:7213)	2	6	4.1	0.000
seawater Fram Strait (MRGID:26579) LTER Observatory HAUSGARTEN siteN4	2	4	0.6	0.545
seawater Vesteris Seamount (MRGID:5046)	2	6	2.7	0.006
seawater off Troms (MRGID:23319) Malangsgrunnen	2	3	0.2	0.878
seawater off Hawke's Bay (MRGID:32521) seamount986	2	6	3.8	0.000
seawater Sambro Bank (MRGID:34408)	2	14	1.3	0.197
seawater Schulz Bank (MRGID:5039)	2	140	2.4	0.016
seawater Weddell Sea (MRGID:4330) Filchner-Ronne (MRGID:17335) Outflow System	2	2	2.1	0.036
seawater off East Cape (MRGID:24347) seamount1247	2	3	3.3	0.001
seawater Gakkell Ridge (GAZ:00002436) Northern Seamount	2	1	0.1	0.918
seawater Fram Strait (MRGID:26579) LTER Observatory HAUSGARTEN siteHG4	2	4	0.6	0.517
seawater Emerald Basin (MRGID:34188)	2	5	0.2	0.877
seawater Fram Strait (MRGID:26579) LTER Observatory HAUSGARTEN siteSV2	2	4	2.9	0.003
seawater off Pico (MRGID:2460)	2	6	-0.1	0.906
seawater Sularevet (WDPAD:555557185)	2	4	0.0	0.982
seawater off Vesterålen (MRGID:18663) Bleiksdjupet	2	9	1.9	0.057
seawater Fram Strait (MRGID:26579) LTER Observatory HAUSGARTEN siteS3	2	4	0.9	0.357
seawater Finnmark (MRGID:8962) Stjernsund	2	9	1.1	0.285
seawater Colville (MRGID:5846) volcano	2	3	1.9	0.061
seawater Gakkell Ridge (GAZ:00002436) Central Seamount	2	1	1.4	0.177
seawater Wairarapa (MRGID:34834) Slope	2	3	1.5	0.137
seawater Kermadec Ridge (MRGID:7213)	2	6	4.1	0.000
seawater Tromsøflaket (MRGID:5034) A	2	11	0.8	0.445
seawater Tromsøflaket (MRGID:5034)	2	3	0.1	0.953
seawater Kermadec Trench (MRGID:7571)	2	3	3.7	0.000
seawater off East Cape (MRGID:24347)	2	3	2.5	0.012
seawater Gakkell Ridge (GAZ:00002436) Karasik Seamount	2	1	0.5	0.636
seawater Devonport Seamount Chain (MRGID:26630) seamount114	2	3	3.2	0.001
seawater Tromsøflaket (MRGID:5034) C	2	12	1.1	0.286
seawater Fram Strait (MRGID:26579) LTER Observatory HAUSGARTEN siteN4	2	4	0.7	0.515
seawater Christchurch (MRGID:32447) Canyon/Slope	2	3	-0.4	0.709
seawater Weddell Sea (MRGID:4330) Prince Gustav Channel (MRGID:9417)	2	4	2.7	0.008
seawater Raoul Island (MRGID:32903) Denham Bay (MRGID:58148)	2	3	2.7	0.006
seawater Sularevet (WDPAD:555557185)	2	4	0.0	0.985
seawater off Troms (MRGID:23319) Malangsgrunnen	2	3	0.2	0.849
seawater Tromsøflaket (MRGID:5034) B	2	9	0.5	0.647
seawater Gakkell Ridge (GAZ:00002436) Northern Seamount	2	1	0.1	0.900
seawater off Pico (MRGID:2460)	2	6	-0.1	0.945
seawater Vesteris Seamount (MRGID:5046)	2	6	2.8	0.005
seawater Fram Strait (MRGID:26579) LTER Observatory HAUSGARTEN siteSV2	2	4	3.0	0.003
seawater Sognefjorden (MRGID:3376)	2	6	-0.1	0.917
seawater Sambro Bank (MRGID:34408)	2	14	1.4	0.168
seawater Otago (MRGID:26502) Slope	2	3	0.9	0.344
seawater off Hawke's Bay (MRGID:32521) seamount986	2	6	3.9	0.000
seawater Schulz Bank (MRGID:5039)	2	140	2.6	0.009
seawater off East Cape (MRGID:24347) seamount1				

Table S17 Results of statistical testing (PERMANOVAs based on weighted UniFrac distances; 999 permutations) conducted to assess differences in the microbial community composition between geographic locations for the different sampletypes.

Group1	Group2	Sample no.	Pseudo F	p-value
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	23 : 11	7.2	0.001
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	23 : 52	16.6	0.001
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA Gakkel Ridge (GAZ:0002436) Saddle between Seamounts	23 : 2	9.4	0.004
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA Colville (MRGID:5846) volcano	23 : 11	21.4	0.001
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA off Hawke's Bay (MRGID:32521) seamount986	23 : 2	3.9	0.004
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA Gakkel Ridge (GAZ:0002436) Central Seamount	23 : 1	4.2	0.049
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA Troms�ffaket (MRGID:5034) B	23 : 12	9.2	0.001
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA Davis Strait (MRGID:4250)	23 : 15	17.4	0.001
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA Schulz Bank (MRGID:5039)	23 : 76	27.8	0.001
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA Raoul Island (MRGID:32903) Denham Bay (MRGID:58148)	23 : 2	8.9	0.004
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA Hordaland (MRGID:18643) Korsfjorden	23 : 2	5.4	0.001
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA Troms�ffaket (MRGID:5034) C	23 : 8	4.1	0.001
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA Sularevet (WDPaid:555557185)	23 : 9	6.6	0.001
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA Finnmark (MRGID:8962) Stjernsund	23 : 17	6.4	0.001
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA Gakkel Ridge (GAZ:0002436) Karasik Seamount	23 : 17	30.8	0.001
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA Troms�ffaket (MRGID:5034) A	23 : 17	8.6	0.001
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA Gakkel Ridge (GAZ:0002436) Northern Seamount	23 : 5	14.8	0.001
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA off Troms (MRGID:23319) Malangsgrunnen	23 : 4	1.7	0.090
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA Vesteris Seamount (MRGID:5046)	23 : 46	30.5	0.001
HMA B�mlo-Sotra archipelago (MRGID:5285)	HMA Kermadec Ridge (MRGID:7213)	23 : 16	16.4	0.001
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA Kermadec Ridge (MRGID:7213)	11 : 16	5.1	0.001
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA off Troms (MRGID:23319) Malangsgrunnen	11 : 4	3.3	0.002
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA Schulz Bank (MRGID:5039)	11 : 76	16.4	0.001
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA Gakkel Ridge (GAZ:0002436) Saddle between Seamounts	11 : 2	6.7	0.017
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA Hordaland (MRGID:18643) Korsfjorden	11 : 10	6.9	0.001
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA Troms�ffaket (MRGID:5034) A	11 : 17	9.2	0.001
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA Gakkel Ridge (GAZ:0002436) Central Seamount	11 : 1	3.1	0.173
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA Raoul Island (MRGID:32903) Denham Bay (MRGID:58148)	11 : 2	7.6	0.012
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	11 : 52	1.8	0.035
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA off Hawke's Bay (MRGID:32521) seamount986	11 : 2	1.7	0.073
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA Finnmark (MRGID:8962) Stjernsund	11 : 17	7.1	0.001
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA Colville (MRGID:5846) volcano	11 : 11	8.9	0.002
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA Davis Strait (MRGID:4250)	11 : 15	15.4	0.001
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA Sularevet (WDPaid:555557185)	11 : 9	6.3	0.001
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA Gakkel Ridge (GAZ:0002436) Karasik Seamount	11 : 17	16.3	0.001
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA Vesteris Seamount (MRGID:5046)	11 : 46	16.8	0.001
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA Troms�ffaket (MRGID:5034) B	11 : 12	7.5	0.001
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA Gakkel Ridge (GAZ:0002436) Northern Seamount	11 : 5	10.1	0.001
HMA Cantabrian (MRGID:3758) Sea Avil�s Canyon (MRGID:36126)	HMA Troms�ffaket (MRGID:5034) C	11 : 8	5.3	0.001
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Gakkel Ridge (GAZ:0002436) Karasik Seamount	52 : 17	37.4	0.001
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Davis Strait (MRGID:4250)	52 : 15	30.3	0.001
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Gakkel Ridge (GAZ:0002436) Central Seamount	52 : 1	4.1	0.019
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Troms�ffaket (MRGID:5034) B	52 : 12	15.8	0.001
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Finnmark (MRGID:8962) Stjernsund	52 : 17	18.7	0.001
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Gakkel Ridge (GAZ:0002436) Northern Seamount	52 : 5	15.4	0.001
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Sularevet (WDPaid:555557185)	52 : 9	12.8	0.001
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Schulz Bank (MRGID:5039)	52 : 76	57.5	0.001
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Kermadec Ridge (MRGID:7213)	52 : 16	12.6	0.001
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Gakkel Ridge (GAZ:0002436) Saddle between Seamounts	52 : 2	9.1	0.001
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Colville (MRGID:5846) volcano	52 : 11	24.5	0.001
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Troms�ffaket (MRGID:5034) C	52 : 8	8.9	0.001
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Hordaland (MRGID:18643) Korsfjorden	52 : 10	12.0	0.001
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Raoul Island (MRGID:32903) Denham Bay (MRGID:58148)	52 : 2	9.4	0.001
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Vesteris Seamount (MRGID:5046)	52 : 46	58.6	0.001
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA off Troms (MRGID:23319) Malangsgrunnen	52 : 4	4.1	0.002
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA Troms�ffaket (MRGID:5034) A	52 : 17	17.1	0.001
HMA Cantabrian (MRGID:3758) Sea Le Danois Bank (MRGID:4719)	HMA off Hawke's Bay (MRGID:32521) seamount986	52 : 2	2.3	0.031
HMA Colville (MRGID:5846) volcano	HMA Finnmark (MRGID:8962) Stjernsund	11 : 17	10.6	0.001
HMA Colville (MRGID:5846) volcano	HMA Gakkel Ridge (GAZ:0002436) Karasik Seamount	11 : 17	16.2	0.001
HMA Colville (MRGID:5846) volcano	HMA Davis Strait (MRGID:4250)	11 : 15	20.9	0.001
HMA Colville (MRGID:5846) volcano	HMA Gakkel Ridge (GAZ:0002436) Saddle between Seamounts	11 : 2	3.4	0.037
HMA Colville (MRGID:5846) volcano	HMA Troms�ffaket (MRGID:5034) C	11 : 8	8.4	0.005
HMA Colville (MRGID:5846) volcano	HMA Kermadec Ridge (MRGID:7213)	11 : 16	13.1	0.001
HMA Colville (MRGID:5846) volcano	HMA off Hawke's Bay (MRGID:32521) seamount986	11 : 2	2.5	0.073
HMA Colville (MRGID:5846) volcano	HMA Hordaland (MRGID:18643) Korsfjorden	11 : 10	9.9	0.001
HMA Colville (MRGID:5846) volcano	HMA Gakkel Ridge (GAZ:0002436) Northern Seamount	11 : 5	6.0	0.004
HMA Colville (MRGID:5846) volcano	HMA Vesteris Seamount (MRGID:5046)	11 : 46	24.6	0.001
HMA Colville (MRGID:5846) volcano	HMA Raoul Island (MRGID:32903) Denham Bay (MRGID:58148)	11 : 2	3.6	0.022
HMA Colville (MRGID:5846) volcano	HMA Troms�ffaket (MRGID:5034) B	11 : 12	10.9	0.001
HMA Colville (MRGID:5846) volcano	HMA off Troms (MRGID:23319) Malangsgrunnen	11 : 4	4.9	0.027
HMA Colville (MRGID:5846) volcano	HMA Schulz Bank (MRGID:5039)	11 : 76	32.1	0.001
HMA Colville (MRGID:5846) volcano	HMA Troms�ffaket (MRGID:5034) A	11 : 17	15.9	0.001
HMA Colville (MRGID:5846) volcano	HMA Sularevet (WDPaid:555557185)	11 : 9	7.2	0.002
HMA Colville (MRGID:5846) volcano	HMA Gakkel Ridge (GAZ:0002436) Central Seamount	11 : 1	1.5	0.173
HMA Colville (MRGID:5846) volcano	HMA off Troms (MRGID:23319) Malangsgrunnen	15 : 4	7.1	0.001
HMA Davis Strait (MRGID:4250)	HMA Hordaland (MRGID:18643) Korsfjorden	15 : 10	18.3	0.001
HMA Davis Strait (MRGID:4250)	HMA off Hawke's Bay (MRGID:32521) seamount986	15 : 2	4.2	0.008
HMA Davis Strait (MRGID:4250)	HMA Gakkel Ridge (GAZ:0002436) Karasik Seamount	15 : 17	34.1	0.001
HMA Davis Strait (MRGID:4250)	HMA Troms�ffaket (MRGID:5034) A	15 : 17	25.4	0.001
HMA Davis Strait (MRGID:4250)	HMA Raoul Island (MRGID:32903) Denham Bay (MRGID:58148)	15 : 2	5.9	0.007
HMA Davis Strait (MRGID:4250)	HMA Gakkel Ridge (GAZ:0002436) Saddle between Seamounts	15 : 2	8.7	0.010
HMA Davis Strait (MRGID:4250)	HMA Sularevet (WDPaid:555557185)	15 : 9	18.5	0.001
HMA Davis Strait (MRGID:4250)	HMA Vesteris Seamount (MRGID:5046)	15 : 46	39.6	0.001
HMA Davis Strait (MRGID:4250)	HMA Troms�ffaket (MRGID:5034) B	15 : 12	23.3	0.001
HMA Davis Strait (MRGID:4250)	HMA Schulz Bank (MRGID:5039)	15 : 76	31.3	0.001
HMA Davis Strait (MRGID:4250)	HMA Troms�ffaket (MRGID:5034) C	15 : 8	14.7	0.001
HMA Davis Strait (MRGID:4250)	HMA Kermadec Ridge (MRGID:7213)	15 : 16	20.8	0.001
HMA Davis Strait (MRGID:4250)	HMA Finnmark (MRGID:8962) Stjernsund	15 : 17	21.4	0.001
HMA Davis Strait (MRGID:4250)	HMA Gakkel Ridge (GAZ:0002436) Northern Seamount	15 : 5	15.6	0.001
HMA Davis Strait (MRGID:4250)	HMA Gakkel Ridge (GAZ:0002436) Central Seamount	15 : 1	4.1	0.063
HMA Finnmark (MRGID:8962) Stjernsund	HMA Troms�ffaket (MRGID:5034) A	17 : 17	3.5	0.002
HMA Finnmark (MRGID:8962) Stjernsund	HMA Raoul Island (MRGID:32903) Denham Bay (MRGID:58148)	17 : 2	6.4	0.008
HMA Finnmark (MRGID:8962) Stjernsund	HMA Troms�ffaket (MRGID:5034) C	17 : 8	1.5	0.129
HMA Finnmark (MRGID:8962) Stjernsund	HMA Troms�ffaket (MRGID:5034) B	17 : 12	3.0	0.012
HMA Finnmark (MRGID:8962) Stjernsund	HMA Hordaland (MRGID:18643) Korsfjorden	17 : 10	2.0	0.050
HMA Finnmark (MRGID:8962) Stjernsund	HMA Kermadec Ridge (MRGID:7213)	17 : 16	12.3	0.001
HMA Finnmark (MRGID:8962) Stjernsund	HMA Gakkel Ridge (GAZ:0002436) Northern Seamount	17 : 5	7.0	0.001
HMA Finnmark (MRGID:8962) Stjernsund	HMA Gakkel Ridge (GAZ:0002436) Central Seamount	17 : 1	2.3	0.060
HMA Finnmark (MRGID:8962) Stjernsund	HMA Vesteris Seamount (MRGID:5046)	17 : 46	14.0	0.001
HMA Finnmark (MRGID:8962) Stjernsund	HMA Schulz Bank (MRGID:5039)	17 : 76	18.4	0.001
HMA Finnmark (MRGID:8962) Stjernsund	HMA off Troms (MRGID:23319) Malangsgrunnen	17 : 4	1.3	0.237
HMA Finnmark (MRGID:8962) Stjernsund	HMA Gakkel Ridge (GAZ:0002436) Saddle between Seamounts	17 : 2	5.2	0.008
HMA Finnmark (MRGID:8962) Stjernsund	HMA Sularevet (WDPaid:555557185)	17 : 9	2.0	0.052
HMA Finnmark (MRGID:8962) Stjernsund	HMA Gakkel Ridge (GAZ:0002436) Karasik Seamount	17 : 17	15.1	0.001
HMA Finnmark (MRGID:8962) Stjernsund	HMA off Hawke's Bay (MRGID:32521) seamount986	17 : 2	3.4	0.034

Declaration

10

Hiermit erkläre ich, Kathrin Busch (geboren am 08.04.1991 in Berlin), dass ich die vorliegende Dissertation bis auf die Beratung durch meine Betreuer, in Inhalt und Form selbstständig und ohne Zuhilfenahme unerlaubter Hilfsmittel angefertigt habe. Ich versichere, dass ich diese Arbeit unter Einhaltung der Regeln guter wissenschaftlicher Praxis der Deutschen Forschungsgesellschaft verfasst habe. Alle benutzten Quellen und Hilfsmittel habe ich vollständig angegeben und die Zusammenarbeit mit anderen Wissenschaftlern habe ich kenntlich gemacht. Weder diese noch eine ähnliche Arbeit wurden an einer anderen Abteilung oder Hochschule im Rahmen eines Prüfungsverfahrens vorgelegt, veröffentlicht oder zur Veröffentlichung vorgelegt. Mir wurde kein akademischer Grad entzogen und dies ist mein erstes und einziges Promotionsverfahren.

Kathrin Busch
Kiel, den 14. November 2020