Host-microbe symbiosis in the marine sponge *Halichondria panicea*

Stephen Knobloch

Faculty of Life and Environmental Sciences
University of Iceland
2019

Host-microbe symbiosis in the marine sponge *Halichondria panicea*

Stephen Knobloch

Dissertation submitted in partial fulfilment of a
Philosophiae Doctor degree in Biology

PhD Committee
Viggó Þór Marteinsson
Ragnar Jóhannsson
Eva Benediktsdóttir

Opponents
Detmer Sipkema
Ólafur Sigmar Andrésson

Faculty of Life and Environmental Sciences
School of Engineering and Natural Sciences
University of Iceland
Reykjavik, May 2019

Host-microbe symbiosis in the marine sponge *Halichondria panicea*
Bacterial symbiosis in *H. panicea*
Dissertation submitted in partial fulfilment of a *Philosophiae Doctor* degree in Biology

Faculty of Life and Environmental Sciences
School of Engineering and Natural Sciences
University of Iceland
Askja, Sturlugata 7
101, Reykjavik
Iceland

Telephone: +354 525 4000

Bibliographic information:
Stephen Knobloch, 2019, Host-microbe symbiosis in the marine sponge *Halichondria panicea*, PhD dissertation, Faculty of Life and Environmental Sciences, University of Iceland, 204 pp.

ISBN: 978-9935-9438-8-0

Author ORCID: 0000-0002-0930-8592

Printing: Háskólaprent
Reykjavik, Iceland, May 2019

Abstract

Sponges (phylum Porifera) are considered one of the oldest extant lineages of the animal kingdom. Their close association with microorganism makes them suitable for studying early animal-microbe symbiosis and expanding our understanding of host-microbe interactions through conserved mechanisms. Moreover, many sponges and sponge-associated microorganisms produce bioactive compounds, making them highly relevant for the biotechnology and pharmaceutical industry. So far, however, it has not been possible to grow biotechnologically relevant sponge species or isolate their obligate microbial symbionts for these purposes.

The marine sponge *Halichondria panicea* inhabits coastal areas worldwide, which makes the sponge easily accessible, and has therefore been a widely studied sponge species in terms of its biology and ecophysiology. However, the stability of its associated microbial community and the functional role of its permanent members has so far remained elusive. Despite previous studies showing bioactive properties of *H. panicea*, it is yet unknown if these are produced by the sponge itself or by microorganisms found in its body.

In this study, the microbial diversity of *H. panicea* from Icelandic waters was analysed through marker gene sequencing, genomics and bacterial cultivation. We showed that *H. panicea* from Icelandic water hosts a dominant bacterial symbiont across geographical locations and seasons whereas other bacteria are geographically and temporally transient. Functional genomic analysis of the dominant symbiont, for which we propose a candidate status with the name "*Candidatus* Halichondribacter symbioticus", showed a lack of regulatory genes in line with its obligate symbiotic lifestyle. In addition, it is enriched in sponge symbiont specific gene families related to host-symbiont interactions and defence. Despite the presence of a putative bacteriocin gene cluster, "*Ca*. H. symbioticus" does not appear to be the source of other bioactive secondary metabolites in this sponge.

A comparison of different methods for *ex situ* cultivation of *H. panicea* showed that, while different cultivation set-ups allowed for successful maintenance of the sponge itself, only certain conditions also maintained "*Ca*. H. symbioticus" at a high relative abundance in the sponge body. Based on these results a novel method for sponge-microbe co-cultivation was demonstrated which increases the enriched bacterial diversity compared to standard plating methods. Although "*Ca*. H. symbioticus" remains un-isolated, other selected isolated bacteria from *H. panicea* were subjected to antimicrobial activity assays, genome sequencing and strain characterisation leading to the description of two novel species of marine bacteria and one novel putative host-associated bacterium.

Due to the presence of a dominant bacterial symbiont and the amenability of *H. panicea* towards laboratory cultivation, the *H. panicea* – "*Ca*. H. symbioticus" system represents a promising model for studying early animal-microbe symbiosis. The results from this study lay the ground work for future studies exploring the interactions between *H. panicea* and "*Ca*. H. symbioticus".

Útdráttur

Svampar (*phylum Porifera*) er talin ein af elstu núlifandi fylkingum dýraríkisins. Náin tengsl þeirra við örverur gera þá áhugaverða og heppilega til rannsókna á fyrstu gerðum samlífis dýra og örvera og til að auka almennt skilning okkar á þessu varðveitta samspili og virkni þess. Að auki framleiða margir svampar og samlífisörverur þeirra lífvirk efnasambönd sem gera þá áhugaverða fyrir líftækni- og lyfjaiðnaðinn. Ekki hefur enn tekist að rækta upp svampa fyrir líftækni eða lyfjaiðnaðinn eða einangra skilyrtar samlíförverur þeirra í þeim tilgangi.

Sjávarsvampurinn *Halichondria panicea* (Brauðsvampur) finnst á strandsvæðum víða um heim. Hann er því aðgengilegur og hefur líffræði hans og vistlífeðlisfræði verið rannsökuð töluvert. Hins vegar er lítið vitað um sambýlisörveruflóru hans og hlutverk hennar í samlífinu. Þrátt fyrir að fyrri rannsóknir hafi sýnt fram á lífvirkni efna sem einangruð voru úr *H. panicea*, er enn ekki vitað hvort þau eru framleidd af svampinum sjálfum eða samlífisörverum hans.

Í þessari rannsókn var fjölbreytileiki örvera *H. panicea* svamps, sem tekin var úr íslensku sjávarumhverfi, skoðaður með merkigenaraðgreiningu, auk raðgreiningu genamengja (metagenom) og erfðamengja ræktaðra baktería. Við sýnum fram á að *H. Panicea*, úr íslensku umhverfi, hýsir eina ríkjandi bakteríutegund. Tegundin sem fékk heitið "*Candidatus* Halicondribacter symbioticus", er einnig til staðar í öðrum *H. panicea* svömpunum sem rannsakaðir hafa verið frá mismunandi stöðum og óháð árstíðum. Hins vegar eru aðrar sambýlisörverur óskilyrtar og meira bundnar við stað og tíma. Greining á genamengi ríkjandi skilyrtu bakteríunnar sýnir fram á að algeng stýrigen vantar í erfðamengi bakteríunnar. Það er í samræmi við hið skilyrta samlífsform en skortur á stýrigenum algengur í ákveðnum genafjölskyldum tengdum samlífisforminu og vörnum þess. Þrátt fyrir að genaklasi fyrir smíði á lífvirka efninu bacteriocin sé til staðar í "*Candidatus* Halichondribacter symbioticus" virðist sem það komi ekki við sögu við framleiðslu lífvirkra efna eða „secondary metabolites".

Bornar voru saman mismunandi aðferðir við ræktun (*ex situ*) á *H. panicea* og sýnt fram á að þó að mismunandi uppsetning búnaðar viðhéldi svampinum á lífi, þá var eingöngu hægt að viðhalda "*Ca. H. symbioticus* " í miklu magni í vefjum svampsins við ákveðnar aðstæður. Á grunni þessara niðurstaðna var þróuð ný aðferð við samræktun á svampi og örverum sem auðgar og eykur fjölbreytileika örveranna samanborið við venjulegar ræktunaraðferðir á agar skálum. Þó "*Ca. H. symbioticus* " sé enn óeinangruð, þá voru aðrar bakteríur einangraðar úr *H. panicea* og voru valdar bakteríur rannsakaðar með tilliti til framleiðslu lífvirkra efna. Ennfremur var erfðamengi þeirra raðgreint og efnaskipti rannsökuð.

Út frá niðurstöðum þessara rannsókna var tveimur nýjum sjávarbakteríutegundum lýst og einni nýrri sambýlistegund. Þar sem sambýlisbakterían er ríkjandi og unnt er að rækta *H. panicea* í rannsóknarumhverfi er "*Ca. H. symbioticus* " samrækt heppilegt módel til rannsókna á samspili samlífisörvara og dýra. Niðurstöðurnar úr þessari rannsókn leggja því grunninn að framtíðarrannsóknum á samspili *H. panicea* og "*Ca. H. symbioticus* "

To Rebecca, Sascha and Nora

For all the love, patience and support.

Table of Contents

List of Figures

List of Tables

List of Papers

This thesis is based on the following papers, one of which is published in a peer-reviewed journal, one is accepted for publication, two have been submitted to international journals and two are presented as manuscripts. They are referred to by their Roman numerals throughout the text. The papers and supplementary material are included at the end of this thesis.

Paper I: **Knobloch, S**., Jóhannsson, R., Marteinsson, V. 2019 Bacterial diversity in the marine sponge *Halichondria panicea* from Icelandic waters and host-specificity of its dominant symbiont "*Candidatus* Halichondribacter symbioticus", *FEMS Microbiology Ecology* 95 (1). fiy220, https://doi.org/10.1093/femsec/fiy220. Published

Paper II: **Knobloch, S**., Jóhannsson, R., Marteinsson, V. Genome analysis of sponge symbiont "*Candidatus* Halichondribacter symbioticus" shows genomic adaptation to a host-dependent lifestyle. Submitted to *Environmental Microbiology*

Paper III: **Knobloch, S**., Jóhannsson, R., Marteinsson, V. Co-cultivation of the marine sponge *Halichondria panicea* and its associated microorganisms. Submitted to *Scientific Reports*

Paper IV: **Knobloch, S**., Daussin, A., Jóhannsson, R., Marteinsson, V. *Pelagibaculum spongiae* gen. nov., sp. nov., isolated from a marine sponge in South-West Iceland. Accepted for publication in *International Journal of Systematic and Evolutionary Microbiology*

Paper V: **Knobloch, S**., Jóhannsson, R., Marteinsson, V. *Tenacibaculum islandicus* sp. nov., isolated from a marine sponge contains secondary metabolite gene clusters. *Manuscript*

Paper VI: **Knobloch, S**., Jóhannsson, R., Marteinsson, V. *Endozoicomonas halichondricola* sp. nov., isolated from the marine sponge *Halichondria panicea*. *Manuscript*

Other papers published during the course of this study:

Borchert, E., **Knobloch, S**., Dwyer, E., Flynn, S., Jackson, S.A., Jóhannsson, R., Marteinsson, V., O'Gara, F., and Dobson, A.D.W. 2017. Biotechnological Potential of Cold Adapted *Pseudoalteromonas* spp. Isolated from 'Deep Sea' Sponges. *Marine Drugs* 15 (6). https://doi.org/10.3390/md15060184.

Steinert, G., Stauffer, C.H., Aas-Valleriani, N., Borchert, E., Bhushan, A., Campbell, A., De Mares, M.C., Costa, M., Gutleben, J., **Knobloch, S**., Lee, R.G., Munroe, S., Naik, D., Peters, E.E., Stokes, E., Wang, W., Einarsdóttir, E., Sipkema., D. 2018. BluePharmTrain: Biology and Biotechnology of Marine Sponges. In *Grand Challenges in Marine Biotechnology* (pp. 505-553). Springer, Cham. https://doi.org/10.1007/978-3-319-69075-9_13.

List of Published Nucleotide Sequences

In the context of this thesis twelve nucleotide sequence datasets were made publicly available across the DDBJ/ENA/GenBank and Sequence Read Archive (SRA) databases. These include:

BioProject accession PRJNA495906 containing 112 BioSamples with 16S rRNA gene amplicon sequences, referred to in **Paper I**.

DDBJ/ENA/GenBank accession numbers MH756603 and MH756604 containing two full mitochondrial genomes of *Halichondria panicea* from South-West and North Iceland, respectively, referred to in **Paper I**.

DDBJ/ENA/GenBank accession numbers MH734183 and MH734529 containing two near full-length 16S rRNA gene sequences of "*Candidatus* Halichondribacter symbioticus" from South-West and North Iceland, respectively, referred to in **Paper I**.

DDBJ/ENA/GenBank accession number RWJS00000000 containing the draft genome sequence of "*Candidatus* Halichondribacter symbioticus" HS1, referred to in **Paper II**.

BioProject accession PRJNA508373 containing two BioSamples with raw metagenomic sequences of *H. panicea*, referred to in **Paper II**.

BioProject accession PRJNA521872 containing 27 BioSamples with 16S rRNA gene amplicon sequences, referred to in **Paper III**.

DDBJ/ENA/GenBank accession number MG877746 containing the full 16S rRNA gene sequence, and whole genome shotgun project QDDL00000000 containing the genome of strain Hp12 *Pelagibaculum spongiae*, referred to in **Paper IV**.

DDBJ/ENA/GenBank accession number MK633875 containing the full 16S rRNA gene sequence, and whole genome shotgun project SIHP00000000 containing the genome of strain Hp32 *Tenacibaculum islandicus*, referred to in **Paper V**.

DDBJ/ENA/GenBank accession number MK633876 containing the full 16S rRNA gene sequence, and whole genome shotgun project SIHQ00000000 containing the genome of strain Hp36 *Endozoicomonas halichondricola*, referred to in **Paper VI**.

Abbreviations

ANI – Average Nucleotide Identity

COG – Cluster of Orthologous Groups

dDDH – digital DNA-DNA hybridisation

DGGE - Denaturing Gradient Gel Electrophoresis

DMSO – Dimethyl sulfoxide

DSMZ – German Collection of Microorganisms and Cell Cultures

FISH – Fluorescence *In Situ* Hybridization

HMA – High microbial abundance

KEGG – Kyoto Encyclopedia of Genes and Genomes

KO – KEGG Orthology

LMA – Low microbial abundance

NGS – Next Generation Sequencing

PBS – Phosphate-buffered saline

R-M – Restriction modification

SCMG – Single copy marker gene

SRA – Sequence Read Archive

Acknowledgements

I would like to thank my supervisors, Viggó Þór Marteinsson for all his guidance and support along the way and Ragnar Jóhannsson for offering me this opportunity and for all his valuable advice. My thanks goes also to Eva Benediktsdóttir for her help as part of my doctorial committee.

This work was supported by the EC Marie Curie Actions ITN Project No. 607786, BluePharmTrain. I am grateful to all the PIs, Postdocs and fellow PhD students in the project and for all the interesting summer schools, workshops and meetings that were organised as part of the training network. It was a pleasure to get to know you all, exchange ideas and learn more about your research.

My thanks also goes to my colleagues at Matís, Alex, Sigurlaug, René, Pauline B, Antoine, Justine, Greg, Mia, Clara, Ólafur, Steinunn, Águsta and Sólveig for your help and advice and for making our lab a fun place to work. A special thanks goes to Pauline Vannier for introducing me to the marvellous world of microbial cultivation and to Elísabet Eik Guðmundsdóttir for answering my many questions concerning all things DNA. Many thanks to Aurélien for his tireless effort to cultivate the uncultivated.

I am grateful to Ásgeir Guðnason for letting me occupy a part of his aquaculture farm for my own little sponge "farm" and always assisting me with his wealth of knowledge. I also thank Þorkell Heiðarsson for allowing me to use the public aquaria at the Reykjavik animal park for one of my experiments. Many thanks to Eydís Einarsdóttir for providing sponge samples, as well as to Þór Ásgeirsson and Jörundur Svavarsson for valuable information on Icelandic sponges. Thank you to Anna María Einarsdóttir and Haukur Einarsson for assisting us with SCUBA diving and sponge collection.

Last but not least, I thank my parents for their never wavering support and for always believing in me. Few things in life are as powerful as that. Along the way I have become a parent myself, to Sascha in 2014 and Nora in 2017. Thank you for your patience when Papa needed time to work on his thesis. Finally, I thank Rebecca, my source of inspiration and without whom this would have been just half the fun.

1 Introduction

1.1 Animal-microbe symbiosis

1.1.1 Definition of symbiosis

Unicellular and multicellular organisms can interact with each other in different ways. This includes mutualistic behaviour where two or more partners benefit from the interaction, commensalism where only one partner benefits and the other goes unharmed, and parasitism in which one partner benefits and thereby harms the other (Leung and Poulin, 2008). The term symbiosis was first coined by de Bary in 1879 summarising the mutualistic, commensal or parasitic interactions between organisms. Since then it has been used with varying meanings, for instance, only including mutualistic relationships (Trager, 2012) or relationships where neither species is harmed (i.e. mutualism, commensalism, and neutralism) (Pianka, 2011). However, nowadays symbiosis is usually used in its broader, original meaning from de Bary (Martin and Schwab, 2013). Symbiotic organisms can further be classified as endosymbionts and ectosymbionts, whereas the former lives inside of another and the later lives on or outside of others (Peacock, 2011).

With an increasing understanding of the strong interdependencies that exist between complex multicellular organisms and microorganisms (Gilbert *et al.*, 2012), new terminologies have come into use to described host-microbe associations. When describing microbial communities associated with a host, an often used term is "microbiota", describing the assemblage of microorganisms present in a defined environment, or "microbiome", encompassing all microorganisms, their genomes and the surrounding environmental conditions (Marchesi and Ravel, 2015). In addition, the terms "holobiont" and "hologenome" have been introduced to describe the ecosystem that is the individual host organisms and its many microbial communities (Bourne *et al.*, 2009). However, this term has been criticised due to its restrictive assumption of a high partner fidelity (Douglas and Werren, 2016). Despite new terminology it remains difficult to define the role of all microorganisms in large host-associated microbiota and evaluate their possible symbiotic relationship towards their host (Taylor *et al.*, 2007). New molecular methods including (meta)genomics/transcriptomics and metabolomics are a way to elucidate these potential interactions (Weinstock, 2012).

1.1.2 Animal-microbe coevolution and interactions

Whereas bacterial life first emerged some 3.5 to 4 billion years ago, animals only evolved from their protist ancestors between 700 and 800 million years ago (Knoll, 2015). Therefore, the current-day interactions of protists with microorganisms was probably already present when the first animals appeared (McFall-Ngai *et al.*, 2013). Major interactions of animals and microorganisms include nutrition, recognition, cell adhesion and signalling. It is thought that these interactions guided the behavioural properties, predation and colony-formation, in

our last common animal ancestor, the choanoflagellates, and were thereby key to the origin of animals (Nichols *et al.*, 2009; Alegado *et al.*, 2012). During diversification of early animals, for instance through the emergence of the gut cavity, organs and the adaptive immune system (Figure 1), animal-microbial interactions continued to shape animal evolution. Through the gut cavity the role of microbes, as nutrient source through predations, was extend to being producers of digestible molecules (Ley *et al.*, 2008). Organs, such as the respiratory system, have likely been influenced by the interactions with microbes (Herbst *et al.*, 2011) and the adaptive immune system is one of the most intricate response mechanisms to microorganisms in vertebrate animals (Eberl, 2010; Hooper and MacPherson, 2010). Inversely, microbial diversity has been shaped by the presence of animals niche systems, shown through highly specific host-associated microbial compositions and taxa (Hongoh, 2010; Kamke *et al.*, 2014; Kwong *et al.*, 2014).

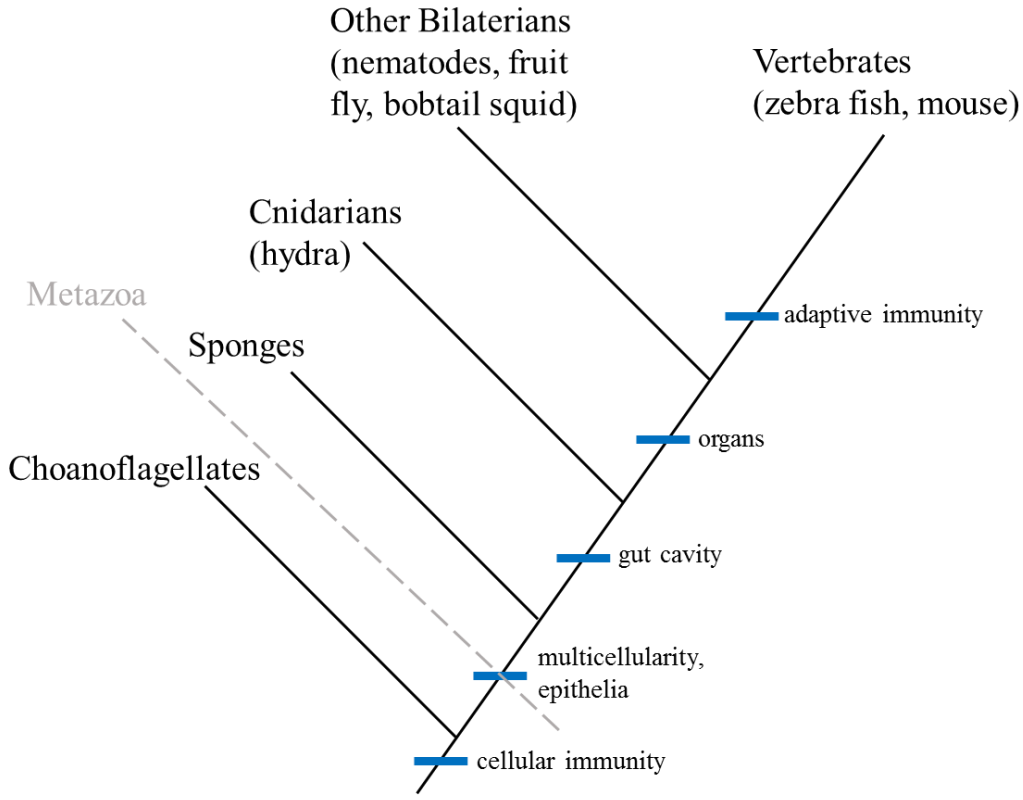


Figure 1: Simplified phylogenetic tree of choanoflagellates and selected animals (metazoa). Evolutionary characteristics relevant to interaction with bacteria are annotated at the branches. Common model organisms are listed in parenthesis. Diagram adapted from McFall-Ngai et al. *(2013).*

Analysis of prokaryotic and eukaryotic genomes shows that most organisms share approximately one third of their genes, reflecting the shared ancestry between all domains of life (Domazet-Lošo and Tautz, 2008). Furthermore, animals genes often have close homologues in bacteria, some of which are integral in signalling between animals and bacteria (Hughes and Sperandio, 2008). As microorganisms have a vastly more diverse metabolic properties, it is not surprising that animal hosts co-opt microorganisms to benefit from their metabolic repertoire (Lapierre and Gogarten, 2009). This greatly increases their metabolic potential and expands both their ecological versatility and adaptability. For example, some invertebrate harbour intracellular bacterial symbionts that are capable of producing amino acids otherwise not produced by themselves (Hansen and Moran, 2011).

Other animals rely on photosynthetic or chemosynthetic symbionts to produce necessary nutrients (Dubilier *et al.*, 2008; Venn *et al.*, 2008). Invertebrates can be particularly rich in symbionts whose function can extend to defending the host from predation (Lopanik, 2014), to protecting the host from intracellular pathogens (Haine, 2008).

Communication between bacteria and animals can occur through cell-to-cell signalling in form of quorum sensing and small molecule exchange (Hughes and Sperandio, 2008; Karlsson *et al.*, 2012). This emerging field known as interdomain or interkingdom signalling (González and Venturi, 2013) is still in its infancy, but with an increased focus on animal-microbe interactions, particularly in the human ecosystem, important new findings have come to light (Shanahan, 2002; Kaper and Sperandio, 2005): Human gut-associated bacteria are, for instance, enriched in N-acyl amide synthesis genes which are capable of interacting with G-protein-coupled receptors that regulate the gastrointestinal tract physiology (Cohen *et al.*, 2017). Conversely, human airway epithelia can inactivate quorum sensing molecules from *Pseudomonas aeruginosa*, thereby interfering with virulence of this opportunistic pathogen (Chun *et al.*, 2004). Animal-microbe communication is not restricted to development and pathogenicity, but have also been shown to modulate behaviour (Collins *et al.*, 2012; Johnson and Foster, 2018), reproduction (Engelstädter and Hurst, 2009), and the immune system (Hooper *et al.*, 2012).

Today it is generally agreed that the understanding of biological processes in animals greatly benefit from viewing animals as host-microbe ecosystems (Camp *et al.*, 2009; McFall-Ngai *et al.*, 2013; Bordenstein, 2016). This viewpoint has generated new insights into human biology and health (Costello *et al.*, 2012), ranging from postnatal colonisation (Dominguez-Bello *et al.*, 2010) to the correlation of the human microbiota with diseases such as diabetes, obesity and cancer (Bäckhed *et al.*, 2004; Louis *et al.*, 2014). As the role of the human microbiota in health and disease continues to be elucidated, new therapies will become available to treat various human diseases (Gilbert *et al.*, 2018). In a larger context of global ecosystem health, a better understanding of host-microbial interactions, weather between microorganisms and animals, plants or fungi, will be needed to mitigate the effects of climate change, restore ecosystems, or support sustainable farming (McFall-Ngai *et al.*, 2013).

1.1.3 Model systems

Model systems are important proxies for studying fundamental relationships in animal-microbe symbiosis (Kostic *et al.*, 2013) and coevolution (Figure 1). Although most studies on these symbiotic interactions focus on humans and vertebrate model systems, such as zebrafish and mice, similar complex interactions can be found in all animal species (McFall-Ngai *et al.*, 2013). The strength of the mouse model system lies in it having many similarities to the human digestive tract. This has been instrumental in studying the role of the gut microbiota in regard to many aspects of mammalian physiology (Kostic *et al.*, 2013). Studies, for instance, on obesity (Cani *et al.*, 2008) or innate and adaptive immune function (Garrett *et al.*, 2010) can be used for evaluating host-microbe interactions applicable to human biology. In addition, mice and humans share most of their genes at the host genetic level, as well as similar gut microbial communities at the phylum and family levels (Spor *et al.*, 2011). Similar to the mouse model, zebrafish (*Danio rerio*) are a powerful model system for exploring host-microbe interactions with an outlook on human physiology (Kostic *et al.*, 2013). Features that make the zebrafish a suitable model are the optical transparency of embryos and larvae, allowing for visualisation of developing cells and microbial

colonisation, as well as a high degree of homology in the adaptive immune system between zebrafish and mammals (Cheesman *et al.*, 2011; Chu and Mazmanian, 2013).

Apart from vertebrate model systems, there are multiple invertebrate models that have been employed to study animal-microbe symbiosis. The green hydra (*Hydra viridissima*) *Chlorella* system, for instance, represents a symbiosis between a complex eumetazoan and a microalgae which benefit from each other through a bidirectional flow of metabolites (Douglas and Smith, 1984; Kovacevic, 2012). Nematodes from the families Steinernematidae and Heterorhabditidae form mutualistic symbiosis with the Gram-negative bacteria *Xenorhabdus* and *Photorhabdus* that enable a pathogenic attack against a separate insect host to provide nutrition for the nematode (Forst and Clarke, 2002). The fruit fly *Drosophila melanogaster* represents a powerful model to explore innate immunity and microbial pathogenesis, in addition to having a high experimental tractability and being a well understood model for basic cellular and developmental biology (Kostic *et al.*, 2013).

One of the best understood host-bacterial systems is the symbiosis between the Hawaiian bobtail squid (*Euprymna scolopes*) and *Vibrio fischeri* (McFall-Ngai and Ruby, 1991). Juvenile bobtail squid must acquire *V. fischeri* from their environment which then populate a ventrally located cavity called a light organ. Bobtail squid feed at night and a cast shadow seen from below could reveal their presence to predators. Counter-illumination from light generated by *V. fischeri* camouflages the squid and is therefore essential to its survival (Ruby and Lee, 1998). This naturally occurring one-on-one relationship has allowed detailed study of the different mechanisms involved in initiation of the symbiosis, the molecular steps necessary for its maintenance and its developmental consequences (McFall-Ngai, 2014).

Another single-organism symbiosis which has provided a larger understanding of animal-microbe symbiotic systems is the relationship between the pea aphid *Acyrthosiphon pisum* and the bacterium *Buchnera aphidicola* (Douglas and Prosser, 1992). Extensive genetic analysis of host and bacterial symbiont coupled with laboratory experimentation have shown that the *A. pisum* /*Buchnera* symbiosis operates at a genomic level to produce essential amino acids needed by the respective partner. In addition, the *Buchnera* genome is reduced and lacks regulatory genes implying that the host plays the dominant role in regulating the symbiosis (Shigenobu and Wilson, 2011).

Invertebrate metamorphosis offers a valuable models for exploring the fundamental interactions of bacterial signalling in animal development (McFall-Ngai *et al.*, 2013). In addition, invertebrates occupying a basal position in the phylogeny of the animal kingdom can provide important insights into the origin of multicellularity. Sponges, for instance, could hold valuable clues as to the role of symbionts in intercellular cohesion (McFall-Ngai *et al.*, 2013). Their phylogenetic and morphological diversity are also useful in studying evolutionary convergence of host-associated microbial communities (Fan *et al.*, 2012) and in determining driving forces in animal-microbe coevolution (Erpenbeck *et al.*, 2002). Sponge-microbe models are therefore fertile ground for studying early animal-microbe symbiosis, revealing these interactions through conserved mechanisms (Pita et al., 2016). Similar to the squid/*Vibrio* and the *A. pisum*/*Buchnera* single-organism systems, marine sponges with single dominant symbionts exist and could be well suited for studying functional interactions of early animal-microbe symbiosis (Croué *et al.*, 2013).

1.1.4 Metataxonomics and genomics as tools to elucidate animal-microbe relationships

It has been known since the beginning of the 20th century that the diversity of microorganisms found on earth far outweighs the diversity detectable through standard cultivation methods (Amann, 1911; Razumov, 1932), a phenomenon termed the "great plate count anomaly" (Staley and Konopka, 1985). In the late 1970s the microbiologists Carl Woese and George Fox pioneered the use of the 16S rRNA gene as a phylogenetic marker, revealing the deep evolutionary history shared between all living organisms and opening up a new path to study the true diversity of microorganisms (Woese and Fox, 1977; Woese, 1987). The use of the 16S rRNA gene rapidly gained popularity as a molecular marker not only for cultivated bacteria, but also as a tool for the direct identification of prokaryotic communities in environmental samples (Lane *et al.*, 1985; Amann *et al.*, 1995). Suitable marker genes such as the 18S rRNA gene and ribosomal internal transcribed spacers further enabled the exploration of microbial diversity of the eukaryotic domain (Kowalchuk, 1999; Díez *et al.*, 2001). Modern sequencing technologies, often termed Next-Generation Sequencing (NGS) or High-Throughput Sequencing, have since enabled the parallel sequencing of million to billion of DNA nucleotides per instrument (Metzker, 2010), thereby vastly increasing sample throughput and deep coverage of microbial communities. This, in turn, has enabled large-scale interdisciplinary research projects such as the Human Microbiome Project (Turnbaugh *et al.*, 2007) and the Earth Microbiome Project (Gilbert *et al.*, 2014), which are instrumental in elucidating the microbial diversity, also referred to as metataxonomics, of animal-associated and general ecosystem-associated microbiota.

Whereas metataxonomics do not directly infer functional characteristics of microbial communities, metagenomics allows for a genomic analysis of microorganism in environmental samples (Hugenholtz and Tyson, 2008). First employed through clone library sequencing (Handelsman *et al.*, 1998; Handelsman, 2004), NGS sequencing coupled with modern bioinformatic tools now enable whole or partial genome recovery from environmental samples (Albertsen *et al.*, 2013), allowing functional characterisation of microbial communities at the genome level. In combination with other tools such as metatranscriptomics, metabolomics and metaproteomics, collectively referred to as "multi-omics", it is also possible to elucidate gene expression, metabolite profiles and protein complement of complex microbial assemblages (Gutleben *et al.*, 2018). These tools have greatly increased our understanding of the functional potential of microbial symbionts (Warnecke *et al.*, 2007; Qin *et al.*, 2010; Fan *et al.*, 2012; Yoon *et al.*, 2015). However, as in the case of sponge-microbe symbiosis, hypothesis-driven, experimental research is highly necessary to move from understanding the functional potential of symbionts to the confirmation of putative symbiont physiology and shared metabolic pathways between the host and their associated microorganisms (Webster and Thomas, 2016).

1.2 Marine sponges and their associated microorganisms

1.2.1 The phylum Porifera

Sponges (phylum Porifera) are a taxonomic group of aquatic animals classified into the four distinct classes Demospongiae (common sponges), Hexactinellida (glass sponges), Calcarea (calcareous sponges) and Homoscleromorpha (encrusting sponges) (Gazave *et al.*, 2012). Over 8500 sponge species have been described so far, the majority of which belong to the class Demospongiae (Van Soest *et al.*, 2012) and are found in the marine environment (Chambers, 2003). Sponges take up a unique position within the metazoans due to their basal position in its phylogenetic tree. Today, it is generally agreed that sponges are the oldest extant lineage of the animal kingdom and thereby the sister group to all other animals (Feuda *et al.*, 2017). This is supported by fossil evidence which dates the emergence of sponges to over 640 million years ago (Yin *et al.*, 2015).

Sponges play important functional roles in ecosystems, including their impact on substrate through reef creation or bioerosion, their role in bentho-pelagic coupling through nutrient cycling, and their association with other organisms, facilitating primary production and provision of microhabitats (Bell, 2008). More recently it has been discovered how sponges create diversity-enhancing sponge grounds in the deep sea (Beazley *et al.*, 2013) and play a vital role in retaining nutrients in coral reefs (de Goeij *et al.*, 2013).

Sponges go through two different life stages: a motile larval phase followed by a sessile adult stage (Ayling, 1980). While some sponge species are carnivorous (Vacelet and Boury-Esnault, 1995), the vast majority of adult sponges are filter feeders, effectively pumping water through their body at rates of up 12 times their own body volume per minute (Reiswig, 1974). Sponges do not possess true tissue but instead have specialised cells that perform specific tasks within the sponge body (Simpson, 1984). A complex network of water canals forms the aquiferous system which contains choanocyte chambers lined with flagellated choanocyte cells (Bergquist, 1978) (Figure 2). Movement of the flagella creates a water current that effectively pumps water through the canals and provides the cells with oxygen and nutrients while removing waste products. Choanocytes and porocytes capture feed organisms and other particles which are transported into the mesohyl to be degraded by archaeocytes, a totipotent cell type that can differentiate into other specialised cells (Bergquist, 1978). Waste particles are then transported back into the water canals for removal. Other specialised cells such as sclerocytes secrete spicules for structural support (Simpson, 2012). The outer layer of the sponge (pinacoderm) is covered with pores (ostia) through which water enters the sponge. Leaving the sponge body, the water is expelled through one or multiple exhalent openings (oscula) (Reiswig, 1971) (Figure 2).

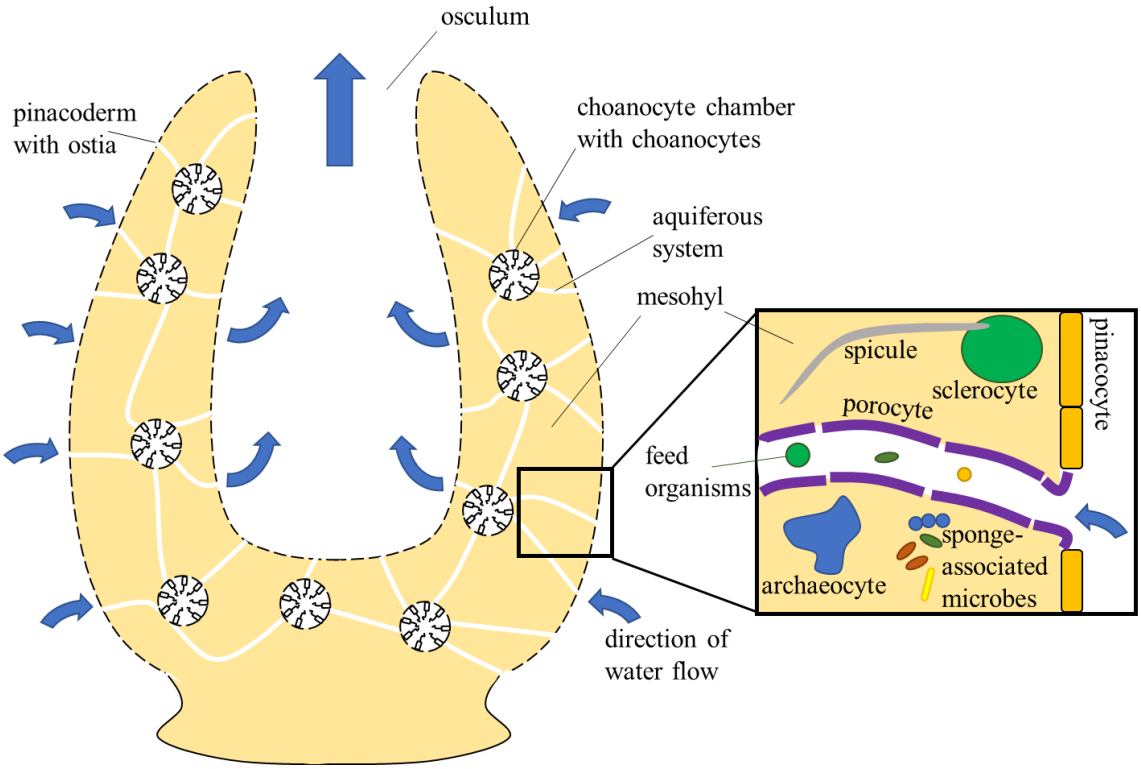


Figure 2: General body structure and main cell types of a "leuconoid" sponge. Water flows through the sponge aquiferous system by means of flagella movement by the choanocytes. Feed organisms and dissolved nutrients are filtered before the water is expelled through the osculum.

1.2.2 Sponge-associated microorganisms

Apart from specialised sponge cells, many sponges contain sponge-associated microorganisms (Wilkinson, 1978; Hentschel *et al.*, 2002). This association was first suggested over 40 years ago through electron microscopy studies of marine sponges (Vacelet and Donadey, 1977). Later, with the advance of modern sequencing technology, our understanding of these microbial communities has expanded significantly (Webster and Thomas, 2016). Sponges can host a remarkably large diversity of microorganisms (Webster *et al.*, 2010; Schmitt, Hentschel, *et al.*, 2012), some of which, such as members of the candidate phylum "Poribacteria", are almost exclusively found in sponges (Lafi *et al.*, 2009; Kamke *et al.*, 2013). Microscopic observations have shown that several billion bacteria can be living in one gram of sponge wet weight (Gloeckner *et al.*, 2014), accounting for up to 40 % of the mesohyl volume (Wilkinson, 1978). Despite this abundance, there appears to be a division between sponge species with a high abundance and diversity of microorganisms and those with a much lower abundant and diverse microbial community. This phenomenon has been termed the high microbial abundance (HMA), low microbial abundance (LMA) sponge dichotomy (Hentschel *et al.*, 2006; Gloeckner *et al.*, 2014).

A key characteristic of sponge-associated microbial communities is the enrichment of specific bacterial taxa, also known as sponge-specific clusters, which are low or absent in the surrounding waters (Hentschel *et al.*, 2002; Simister *et al.*, 2012; Taylor *et al.*, 2012). Although there is little commonality in bacterial species composition or structure across the

sponge phylum, convergent evolution appears to have led to similar sponge-associated community organization and interactions (Easson and Thacker, 2014; Thomas *et al.*, 2016). These sponge-associated communities are characterized by specialist, generalist and opportunist microbes, whereas the core sponge communities usually contain the former two types (Thomas *et al.*, 2016). Due to the large microbial diversity found in sponges, identifying obligate sponge symbionts, as opposed to transient or opportunistic sponge-associated bacteria, is still a matter of debate (Taylor *et al.*, 2007)

Although sponge pathogens exist (Webster, 2007), many sponge-associated microorganisms live in commensal relationships with their host. These include potentially mutualistic behaviour such as defence (Flatt *et al.*, 2005), shared metabolism (Erwin and Thacker, 2007; Freeman and Thacker, 2011; Tian *et al.*, 2016) or the removal of metabolic waste products or contaminants (Schläppy *et al.*, 2010; Siegl *et al.*, 2011; Keren *et al.*, 2017). In addition, an increase in metagenomic and single-cell genomic studies have revealed a range of shared functional genomic features as well as metabolic specialisations between sponge symbionts (Webster and Thomas, 2016). A shared feature between obligate sponge symbionts include the enrichment of mobile genetic elements, the presence of eukaryotic-like proteins, and restriction-modification systems related to bacterial defence (Fan *et al.*, 2012; Nguyen *et al.*, 2014; Díez-Vives *et al.*, 2017; Slaby *et al.*, 2017). Although experimental research remains low, these features are thought to be involved in host-microbe interactions, evading the host immune system and transfer of mobile genetic elements between sponge symbionts. Specialised genetic features of sponge symbionts include secondary metabolite synthesis, functional adaptations to a host-dependent lifestyle (Siegl and Hentschel, 2010; Liu *et al.*, 2011; Gao *et al.*, 2014; Burgsdorf *et al.*, 2015) and specialised metabolic functions such as CO_2-fixation, sulphur reduction and oxidation, urea utilisation and ammonia oxidation (Hallam *et al.*, 2006; Hoffmann *et al.*, 2006; Siegl *et al.*, 2011; Fan *et al.*, 2012; Gauthier *et al.*, 2016). Despite an increased knowledge of the functional features of sponge symbionts, many questions remain, such as defining the molecular determinants of sponge-microbe and microbe-microbe interactions and evaluating the extent to which sponge-associated microbes enhance the environmental adaptation of the sponge (Webster and Thomas, 2016). Moreover, it is not yet clear what role microbial symbionts play in sponge-symbiont coevolution (Webster and Taylor, 2012).

1.2.3 The biotechnological potential of sponges and the "supply problem"

Sponges have been used by humans for thousands of years, playing a significant role in ancient Egyptian, Greek and Roman civilisations (Brümmer and Nickel, 2003; Pronzato, 2003; Duckworth, 2009). Apart from their useful material properties making them ideal as bathing or cleaning tools, sponges have been valued for medicinal purposes since the Greek Antiquity. In several of the works by Greek physician Hippocrates, sponges are described as a remedy for various illnesses and remarkably physicians at that time had already identified the anti-inflammatory properties of sponges when applied to wound healing (Voultsiadou, 2010). Despite these early use cases, it was not until the advance of modern biotechnological techniques in the 20th century that the real pharmaceutical potential of sponges was discovered (Sipkema, Franssen, *et al.*, 2005).

Today sponges are considered one of the richest sources of novel bioactive marine natural products (Mehbub *et al.*, 2014). Due to their sedentary lifestyle and lack of hard protective

layer, sponges use a chemical defence system to ward off predators or compete for surface area (Wulff, 2006). Many of the compounds discovered in sponges exhibit a wide range of bioactive properties, including anti-cancer, anti-microbial and anti-inflammatory activity (Hirata and Uemura, 1986; Hayashi *et al.*, 1992; Sata *et al.*, 1999). Despite the large discovery of novel bioactive compounds only few sponge-derived pharmaceutical, such as Ara-C for the treatment of leukaemia and lymphoma or Ara-A an antiviral drug, have so far reached the market (Sipkema, Franssen, *et al.*, 2005). One of the reasons for this discrepancy is that large amounts of the target compound must be guaranteed before it can enter the stage of preclinical trials. Many sponge derived compounds are however only present in low concentrations (Mendola, 2003) and removing extensive amounts of biomass would have a severe negative ecological impact. This problem has been termed the "supply problem" and has spawned new ideas to source larger amounts of sponge-derived bioactive compounds including large scale sponge farming or the heterologous expression of secondary metabolite gene clusters (Osinga *et al.*, 1999; Sipkema, Osinga, *et al.*, 2005).

With a growing interest in sponge derived natural products, there has also been an increased interest in sponge microbiology (Sacristán-Soriano *et al.*, 2016), not least due to findings that microbes are involved in the production of relevant bioactive metabolites within the sponge body (Piel, 2009). Isolation and cultivation of bioactive compound producing microorganisms has therefore been suggested as a method for overcoming the "supply problem". However, despite this approach isolation of relevant sponge-associated bacteria remains low (Steinert *et al.*, 2018). A multi-omics approach, whereby genomic and metabolic information of yet uncultivated microbes are analysed, shows promising potential to overcome the limitation of traditional microbial isolation strategies (Gutleben *et al.*, 2018).

1.3 The marine sponge *Halichondria panicea*

1.3.1 *H. panicea* biology and ecology

The sponge *Halichondria panicea*, also known under the common name "breadcrumb sponge" due to its crumb-of-bread texture, was first described by zoologist Peter Simon Pallas in 1766. Due to its many different morphological appearances *H. panicea* has been misidentified as a new species and accordingly given 56 different taxonomic names over the years (http://blogs.nature.com, accessed 25.2.2019). Although it has been stated that *H. panicea* is a cosmopolitan species (Barthel and Wolfrath, 1989) its distribution is not yet well defined. A voucher specimen (BELUM: Mc109. Portrush, Co Antrim) from Northern Ireland is considered a valid reference and its distribution along the Atlantic coast of Europe is well established (Ackers *et al.*, 2007). Other reports of *H. panicea* from the East Pacific (Lee *et al.*, 2009) or New Zealand (Bergquist, 1961) are contested, not least due to differences in molecular markers pointing towards different lineages of a *H. panicea*-like species-complex with similar morphological characteristics (Erpenbeck *et al.*, 2004). Despite attempts in assessing the genetic relatedness of *H. panicea* populations (Knowlton *et al.*, 2003) this species complex remains to be resolved.

The shape of *H. panicea* is very polymorphic, ranging from thin sheets or rock surfaces to massive forms, sometimes growing profusely on the stipe of *Laminaria hyperborea* (Ackers *et al.*, 2007) (Figure 3). Both current regimes and wave force have an impact on growth

forms and the body plan of *H. panicea* (Palumbi, 1984, 1986; Barthel, 1991). Its colour can range from green to cream-yellow depending on its exposure to light and the presence of endobionts such as *Microspora ficulinae* (Christensen, 1985). *H. panicea* adapts to a wide range of niches but is usually found on rocks or other hard surfaces along the shoreline, from the intertidal area down to the circalittoral (Ackers *et al.*, 2007).

Figure 3: Clockwise from top-left: Cross-section of H. panicea *attached to the thallus of* Laminaria *sp.; Beach cast and partially dried* H. panicea *after a storm in South-West Iceland. Pores connecting the aquiferous system are visible due to the degradation of the pinacoderm;* H. panicea *associated macrofauna. Red asterisk marks a nematode, red cross marks a crustacean with sponges growing on its carapace.*

As a regular member of coastal ecosystems in the North-East Atlantic, *H. panicea* has been the focus of many ecological and biological studies over the past decades and remains one of the most studied sponge species today (Figure 4). The ecophysiology of *H. panicea* was first investigated in depth by Barthel (1986, 1988) through field experiments and laboratory observations, evaluating substrate specificity, growth rates, predatory activity, energy demands and the integration of the sponge in coastal environmental processes. Even earlier, ecological studies found *H. panicea* to be associated with a large and varied fauna, exhibiting commensal-protective mutualism with the scallop *Chlamys varia* (Forester, 1979) and appearing to exhibit chemical attraction for the skeleton shrimp *Caprella linearis* (Peattie and Hoare, 1981).

H. panicea showed high regenerative capacity when grazed by nudibranchs (Knowlton and Highsmith, 2005) and has the capability to shed its outer cell layer to regularly remove fouling organisms in a process termed tissue sloughing (Barthel and Wolfrath, 1989). Despite this mechanism, reduced water flow rates can induce excessive bacterial growth on

the sponge which can lead to mortality (Hummel *et al.*, 1988). This could be connected to the lack of waste particle removal through low water currents (Wolfrath and Barthel, 1989).

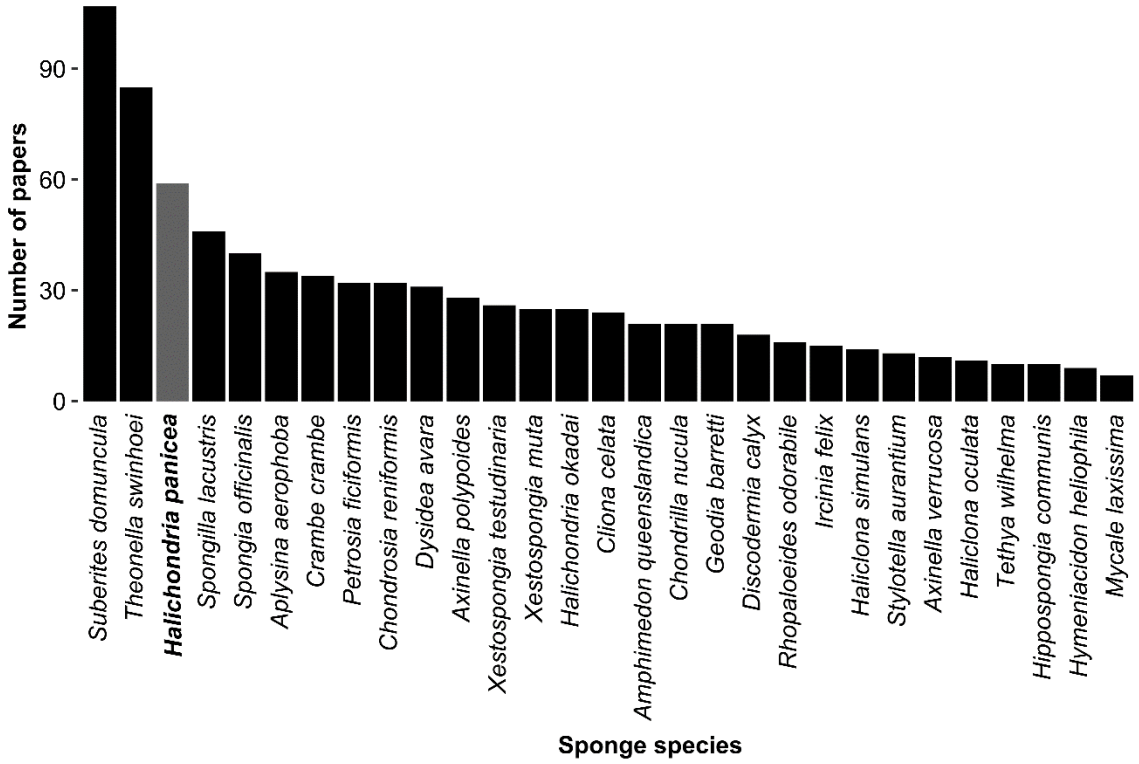


Figure 4: Non-exhaustive list of selected sponge species and number of published papers with full species name in title. Search was performed on Web of Science (https://apps.webofknowledge.com/, accessed 24.2.2019) against Web of Science Core Collection (since 1970).

The reproductive cycle of *H. panicea* has been studied in relationship to reproductive strategies and environmental conditions such as light cues, temperatures and salinity (Amano, 1986; Witte *et al.*, 1994; Gerasimova and Ereskovsky, 2007). *H. panicea* larvae survival and metamorphosis are influenced both by the concentrations of conspecific excretory-secretory products as well as by excretory products from other invertebrate (Khalaman *et al.*, 2014). Although not yet found to have an ecological impact, the characteristic smell exhibited by *H. panicea* was analysed by Christophersen *et al.* (1989) who identified three sulphur compounds responsible for its strong odour.

The filtration and respiration rates of *H. panicea* has been studied extensively under experimental conditions and in the wild, and compared to other marine macroinvertebrate, gaining knowledge on sponge water pumping capacity and retention efficiency (Riisgård *et al.*, 1993; Riisgård and Larsen, 1995; Thomassen and Riisgård, 1995), as well as on utilised carbon sources (Riisgård *et al.*, 2016). *H. panicea* was used to measure uptake of silica, an important mineral in sponges used in the formation of spicules, (Frøhlich and Barthel, 1997; Reincke and Barthel, 1997), as well as seasonality of spicule production (Schönberg and Barthel, 1997). *H. panicea* has further been studied for its suitability as a biomonitoring organism due to its heavy metal accumulation rates (Hansen *et al.*, 1995).

In recent years *H. panicea* has been used as a model species to investigate general sponge specific properties and processes, such as the flagellar apparatus of sponge larvae and adult cell types (Sokolova *et al.*, 2019), the sponge pump and exhalent jet speeds (Goldstein *et al.*, 2019), the relationship between osculum dynamics, filtration activity and respiration (Kumala *et al.*, 2017; Kumala and Canfield, 2018), the condition index in relation to seasonal feed availability (Lüskow *et al.*, 2019), and the formation of primmorphs (Sipkema *et al.*, 2003). In addition, oxygen requirements of *H. panicea* have been studied to make conclusions about the role of oceanic and atmospheric oxygen content in triggering the origin of animal life (Mills *et al.*, 2014).

1.3.2 *H. panicea*-associated microorganisms

Early microscopic observations of *H. panicea* focused on sponge-associated microalgae and detected the presence of an endobiont classified as *Microspora ficulinae* (Wille, 1910; Feldmann, 1933; Christensen, 1985). Later, electron microscope observations classified *H. panicea* as a LMA sponge due to the comparably low abundance of microorganisms found in its mesohyl (Moitinho-Silva, Steinert, *et al.*, 2017). Cultivation studies focused both on the general cultivability of bacteria in *H. panicea* (Imhoff and Stöhr, 2003; Wichels *et al.*, 2006; Lee *et al.*, 2007) as well as on the specific cultivability of members in the phylum *Actinobacteria* (Schneemann, Nagel, *et al.*, 2010). Isolated bacteria belonged to the taxa *Actinobacteria*, *Alpha-*, *Beta-* and *Gammaproteobacteria*, *Deinococcus*, *Cytophagia*, *Flavobacteria, Bacteroidetes* and to a group of low G+C Gram-positive bacteria, however clear sponge-species association between *H. panicea* and the bacterial strains could not be shown.

The first study on the microbial diversity of *H. panicea* using culture-independent molecular methods was conducted by Althoff *et al.* (1998) who collected sponge samples from the North Sea, Baltic Sea and Adriatic Sea and subjected them to 16S rRNA gene based clone library analysis. The results showed that all samples contained a dominant bacterial taxon belonging to the class *Alphaproteobacteria,* which were classified as *Rhodobacter* spp. by the authors due to high sequence similarity with cultivated members of the genus. Based on these findings and electron microscope images of bacteria in the sponge mesohyl, the authors suggested that *H. panicea* lives in a symbiotic relationship with these dominant bacteria. Later, Wichels *et al.* (2006) investigated the bacterial diversity and variability of *H. panicea* from different stations around the island of Heligoland (North Sea) throughout the year and from different "tissue" fractions using denaturing gradient gel electrophoresis (DGGE) and clone libraries. Their findings pointed towards a high variability of the bacterial diversity in relation to time, location and "tissue fraction", whereas most "tissue" samples contained a specific alphaproteobacterium affiliated with the *Roseobacter* group which had high sequence similarity to one of the dominant taxa found by Althoff *et al.* (1998). The authors concluded that the bacterial signature of *H. panicea* could be driven by biotic and abiotic factors but agree that the sponge lives in a close association with the previously detected, dominant alphaproteobacterium. Studying *H. panicea* from the North-East Pacific using culture-independent DNA finger-printing techniques Lee *et al.* (2009) found that its bacterial community was highly variable among sites and that, unlike in previous studies, it did not contain a consistent dominant member in the bacterial community. This led the authors to conclude that there are considerable variations in the bacterial communities associated with *H. panicea* from different locations.

Apart from members of the *Alphaproteobacteria*, a study by Kravchuk *et al.* (2018) focused on epibionts of the genus *Pseudoalteromonas* in *H. panicea* from the White Sea. Whereas the abundance of *Pseudoalteromonas* spp. was dependent on which colour morph was inspected, the authors showed that *H. panicea* with a high abundance of *Pseudoalteromonas* also showed an increased expression of the heat shock protein Hsp70 and a reduction of the proteasomal catalytic beta 5 subunit. The authors conclude that *Pseudoalteromonas* epibionts could affect the ubiquitin-proteasome system of cold-water marine sponges leading to an increased adaptive plasticity.

Using NGS technology Naim *et al.* (2014) revisited the bacterial diversity of *H. panicea* from the Netherlands, showing that three sponge samples contained a dominant OTU affiliated with the *Alphaproteobacteria* at a relative abundance of approximately 80.5 %. The bacterial community dataset of seven sponge taxa analyse through NGS by Steinert *et al.* (2017) confirmed the presence of a dominant alphaproteobacterium in *H. panicea* from the Swedish North Sea coast with high sequence similarity to the dominant taxon in all previous studies.

Although, the microbial community of *H. panicea* has been studied using deep-sequencing NGS technology, thereby greatly increasing the depth of the microbial community structure detected in the sponge, the *H. panicea*-associated microbial community has not been studied across space and time using this approach. Therefore, the identification of a core *H. panicea*-associated microbial community is still elusive. Moreover, the functional role of the core sponge-associated members inside the host ecosystem remain to be elucidated.

1.3.3 Bioactive properties

H. panicea has been associated with a variety of bioactive compounds and well as previously uncharacterized compound discoveries. Cimino *et al.* (1973) discovered five new substances in *H. panicea* which they termed paniceins and elucidated their chemical and spectral properties. Although later also found in other sponge species, panicein A, B_3 and C exhibited cytotoxic activity against cancer cell lines (Casapullo *et al.*, 1993). In addition, four new haliclonadiamine analogues and a new bicyclic guanidine alkaloid were isolated from *H. panicea* collected in Japan, the former compounds showing antimicrobial and antitumor activity (Abdjul *et al.*, 2016). Organic and aqueous extracts from *H. panicea* collected in the Arabian Sea showed various bioactive properties including increased ATP-ase activity, hemolytic activity, increased acetylcholine esterase activity, immunostimulant effects and angiogenic activity (Purushottama *et al.*, 2009). Aqueous extracts from *H. panicea* collected in North-West Spain affected cell appearance, caused significant rates of cell detachment and decrease of mitochondrial membrane potential of a neuroblastoma cell line (Ferreira *et al.*, 2011), leading the authors to conclude that the tested extracts contained compounds with potential pharmacological or biotechnological applications. Similarly, organic extracts from *H. panicea* collected from the North Sea, Baltic Sea and Adriatic Sea displayed cytotoxic properties against leukaemia cells (Althoff *et al.*, 1998). In a study by Toth and Lindeborg (2008) water-soluble compounds from *H. panicea* deterred attachment of barnacles, although it was not determined if this was due to chemical defence from the sponge to prevent fouling or if the barnacles evolved an ability to recognise water-soluble compounds from sponges to avoid attachment to an unsuitable substrate.

In addition to bioactive extracts and compounds derived from the sponge, two bacterial strains isolated from *H. panicea*, showing highest identity to *Antarcticum vesiculatum* and *Psychroserpens burtonensis*, displayed neuroactivity by agonizing N-methyl-D-aspartate receptors in cortical neurons from rats as cell system (Perovic *et al.*, 1998). A bacterium *Microbacterium* sp., isolated from *H. panicea* produced diphosphatidylglycerol and glycoglycerolipids, with the main compound showing modulative properties related to water surface tension and displaying antitumor activity (Wicke *et al.*, 2000; Lang *et al.*, 2004). In a study by Schneemann *et al.* (2010) 46 actinobacterial strains were isolated from *H. panicea* collected in the Baltic Sea. Most strains showed antimicrobial activity and contained biosynthesis genes encoding polyketide synthases and nonribosomal peptide synthesis. One *Streptomyces* strain was selected for further characterisation and was shown to produce the novel compounds mayamycin which exhibited potent cytotoxic activity against eight human cancer cell lines and also showed strong antimicrobial activity against several human pathogens (Schneemann, Kajahn, *et al.*, 2010). Apart from bacteria with antimicrobial and cytotoxic activity, *H. panicea* showed to be a good source of biosurfactant-producing bacteria, with isolated strains showing good response to biosurfactant screening assays in a study by Rizzo *et al.* (2018).

Despite the discovery of many bioactive properties from *H. panicea* it is not yet clear if the core members of the *H. panicea*-associated microbial community are the producers of some of these compounds or if they are instead produced entirely by the sponge or other opportunistic and transient bacteria. In that sense, the original hypothesis from Althoff *et al.* (1998), that *H. panicea* harbours toxic bacteria, requires reinvestigation.

2 Objectives

The aim of this study was to analyse the microbial diversity of the marine sponge *Halichondria panicea* from Icelandic waters using NGS technology and to determine the spatio-temporal stable members of its associated microbiota. Additionally, the aim was to analyse the functional characteristics of its dominant symbiont through genome binning and comparative genomics, and to evaluate its putative role in *H. panicea*. The objective was also to evaluate suitable methods for sponge cultivation and analyse putative changes in the sponge-associated microbial community. Finally, the aim was to analyse the cultivable bacterial diversity and evaluate a method to cultivate the dominant symbiont and other sponge-associated bacteria with potential bioactive properties outside of their host environment.

In **Paper I**, *H. panicea* samples collected from four difference geographical locations in South-West and North Iceland during 2013 to 2016 were subjected to 16S rRNA gene amplicon sequencing to investigate which members of the bacterial community belonged to the core sponge microbiota across seasons and geographical locations. Near full length 16S rRNA gene sequences of the dominant bacterium "*Candidatus* Halichondribacter symbioticus" found within all samples were compared to identify single nucleotide variants and consequently the phylogenetic relationship between the symbiont from different locations. 16S rRNA Fluorescence *In Situ* Hybridisation (FISH) was applied to detect the location of the "*Ca*. H. symbioticus" and other bacteria inside the sponge body.

The aim of **Paper II** was to analyse the functional genomic features of "*Ca*. H. symbioticus" through metagenomic sequencing, binning and genome analysis. In addition, gene expression was studied based on a publicly available RNA-seq dataset of *H. panicea*. Sponge symbiotic features were detected through comparative genome analysis and reconstructed metabolic pathways were used to evaluate the putative role of "*Ca*. H. symbioticus" inside its host.

In **Paper III**, we evaluated suitable conditions for *ex situ* cultivation of *H. panicea* with a focus on maintaining a high relative abundance of "*Ca*. H. symbioticus", similar to its wild counterpart. The aim was to develop a cultivation method suitable for a laboratory with access to seawater, thus enabling experimental research on the *H. panicea* - "*Ca*. H. symbioticus" symbiotic system. In addition, we tested a method to cultivate sponge-associated bacteria outside of the sponge body using a sponge-bacteria co-cultivation chamber. Selected bacteria isolated through standard plating methods were subjected to antimicrobial activity assays and genome mining for secondary metabolite gene clusters, to detect putative bioactivity with biotechnological value.

Three bacteria isolated from *H. panicea* with low 16S rRNA gene sequence similarity to previously characterised species were subjected to extensive strain characterisation in order to evaluate their putative role inside the sponge and determine their taxonomic placement. **Paper IV, V and VI** contain the characterisation for the strains designated Hp12, Hp32 and Hp36.

3 Methods

3.1 Sponge collection, identification and genetics

3.1.1 Sample collection

For sponge cultivation experiments, bacterial isolation and microbial diversity analysis samples of *H. panicea* were collected from four different locations across three years (see Table 1). *H. panicea* grew in subtidal habitats, often attached to the thallus of the macroalgae *Laminaria* sp., or in intertidal areas attached to stones, and was collected from both types of sites (Figure 5). *H. panicea* attached to macroalgae in high wave-action locations were found to be large and globous, whereas in intertidal areas sponge individuals were flat and encrusting. In addition, some sponges were green in colour on the sponge surface, whereas others were yellow or both green and yellow.

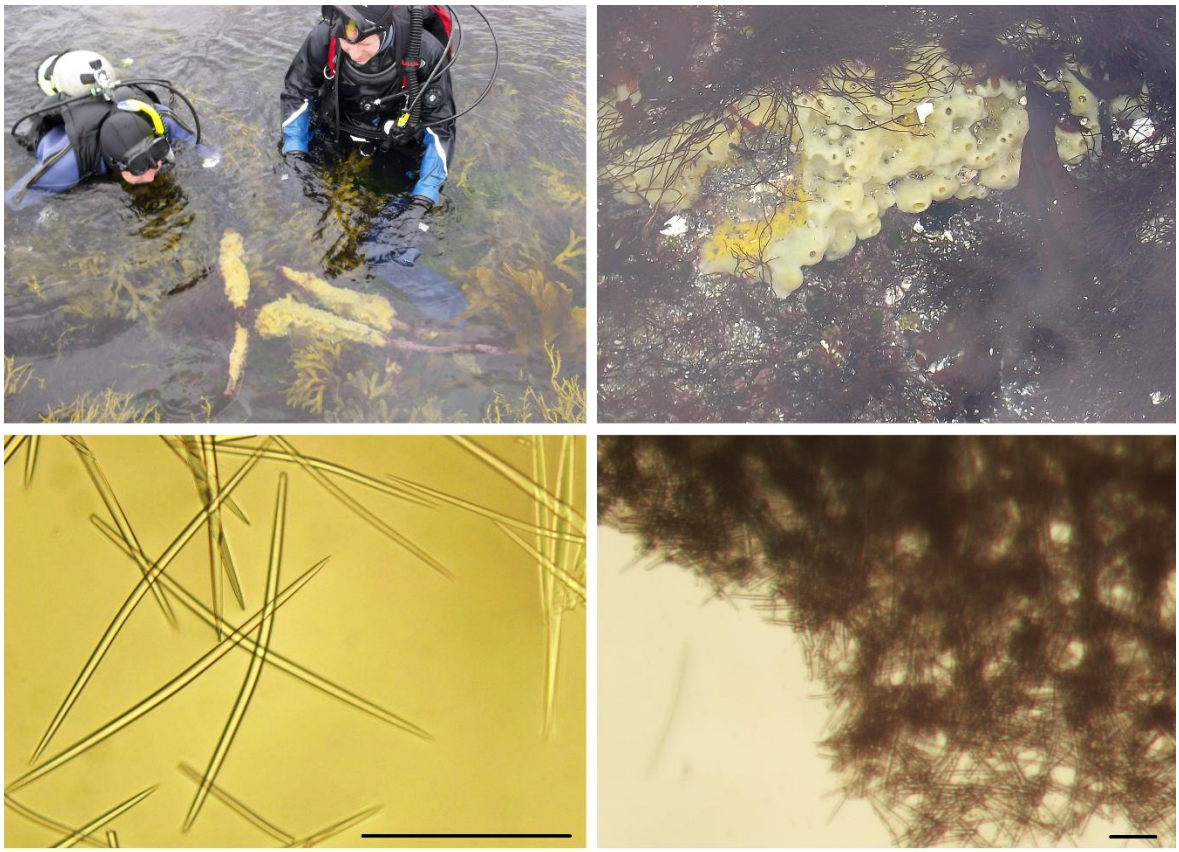


Figure 5: Clockwise from top-left: H. panicea *collected from the subtidal area by scuba diving;* H. panicea *attached to a stone in the intertidal area (cut surface can be seen on the left); Spicule preparation of* H. panicea *sample showing skeletal structure with confused reticulation of megascleres; Spicule preparation from* H. panicea *sample showing individual megascleres; Bar is 0.1 mm.*

Sponges for cultivation experiments were transported to the cultivation facility in 20 litre buckets filled with seawater. For bacterial isolation and microbial diversity analysis, samples were cut from the sponge using a sterile scalpel, rinsed with sterile artificial seawater and transported on ice to the laboratory. Samples for microbial diversity analysis were stored at -80 to -20°C until further processed for DNA extraction.

Table 1: Overview of H. panicea *sample collection for all experiments conducted in this study*

Site	Location	Coordinates	Sampling time	Sample frequency (no. of samples)	Seawater sample (no. of samples)	Experiment
Eyrarbakki	SW-Iceland	63°51' N 21°09' W	Sep. 2015 - Sep. 2016	Monthly (36)	yes (12)	Microbial diversity analysis (**Paper I**)
Grótta	SW-Iceland	64°09' N 22°00' W	Mar. 2015 - Sep. 2016	Seasonally (21)	yes (7)	Microbial diversity analysis (**Paper I**)
Hafnír	SW-Iceland	63°55' N 22°41' W	Apr. 2014 & Nov. 2014	Twice (6)	partially (1)	Microbial diversity analysis (**Paper I**)
Skútan	N-Iceland	65°44' N 18°06' W	Nov. 2013*	Once (6)	no	Microbial diversity analysis (**Paper I**)
Grótta	SW-Iceland	64°09' N 22°00' W	Aug. 2016	Once (2)	no	Metagenomics/genomics (**Paper II**)
Eyrarbakki	SW-Iceland	63°51' N 21°09' W	Mar. 2015	Once (8)	yes (2)	*Ex situ* cultivation (**Paper I & III**)
Eyrarbakki	SW-Iceland	63°51' N 21°09' W	Apr. 2014	Once (1)	no	*Ex situ* cultivation (**Paper III**)
Grótta	SW-Iceland	64°09' N 22°00' W	Apr. 2016	Once (1)	no	*Ex situ* cultivation (**Paper III**)
Grótta	SW-Iceland	64°09' N 22°00' W	Jun. 2016	Once (1)	yes (3)	Co-cultivation (**Paper III & IV**)

* samples provided by Eydís Einarsdóttir from Háskoli Islands

3.1.2 Sponge identification

Sponges were identified via spicule structure, morphological traits and marker gene analysis. For spicule preparation sponge samples were digested in nitric acid for one hour under constant shaking, rinsed three times in distilled water, once in 96 % ethanol and observed under an Olympus BX51 light microscope. Spicule shape and size, as well as morphological characteristics were compared against the species description in Ackers *et al.* (2007). For marker gene analysis the complete mitochondrial DNA was extracted from metagenomes (see section 3.4) and compared against the NCBI nucleotide database (NCBI Resource Coordinators, 2017) using the BLAST tool (Camacho *et al.*, 2009). Highest sequence similarity was detected to *H. panicea* voucher 'BELUM<GBR>:Mc4070' (Genbank: KC869423) and *H. panicea* voucher 'SB16S' (Genbank: KY492588) with 100% and 99% respectively.

3.1.3 Phylogenetic inference

For phylogenetic analysis of *H. panicea* collected in Icelandic waters the partial COI gene covering the conserved Folmer fragment of selected individuals was aligned with publicly available sequences of *H. panicea* and related sponge species. A phylogenetic tree was built in Mega (Kumar *et al.*, 2016) using the Maximum Likelihood method based on the Tamura-Nei model (Tamura and Nei, 1993) with 100 bootstraps (Felsenstein, 1985). Whereas *H. panicea* from South-West Iceland fell within a monophyletic clade with two other *H. panicea* specimens (Figure 6), including the voucher specimen KC869423, the samples from North Iceland as well as a sample from Portugal (KY492588) branched outside of this clade. This difference found in the COI gene could indicate that *H. panicea* from North Iceland represents a new species or sub-species closely related to *H. panicea*. However, the *H. panicea* species complex is not yet well resolved (Erpenbeck *et al.*, 2004), which is further complicated by the high morphological variability and subsequent misidentifications of *H. panicea* and *H. panicea*-like sponges. Spicule preparation and morphological characterisation did not show distinct differences between South-West and North Icelandic specimens, which is why we will refer to the North Icelandic specimens as *H. panicea* throughout this study. The full mitochondrial genomes (including the COI gene) of *H. panicea* from both locations, made publicly available through this study, will facilitate future taxonomical and phylogenetic efforts to further resolve this species-complex.

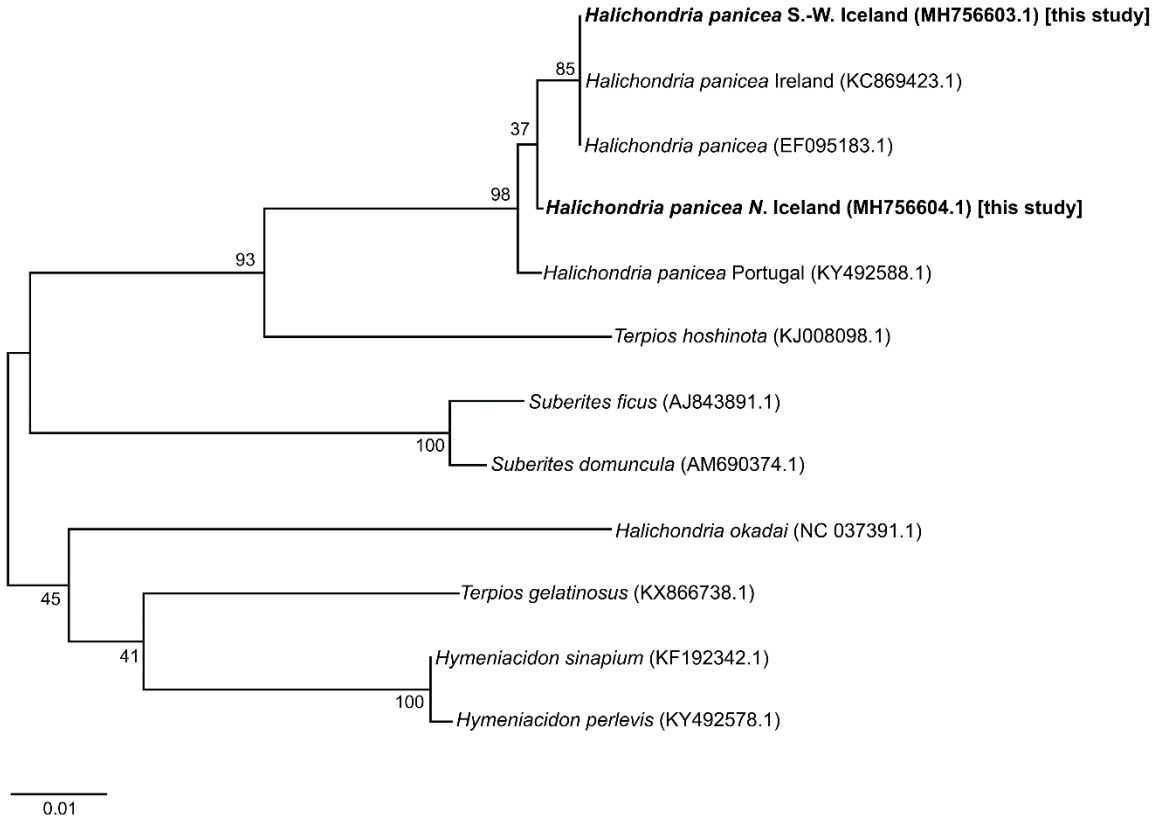


Figure 6: Phylogenetic tree of H. panicea *and related sponge species comparing the Folmer region of the COI gene. Bootstraps in percentage are shown at the branching points.*

3.2 Sponge cultivation

In order to find suitable conditions for sponge cultivation and analyse the microbial diversity during *ex situ* cultivation, five different aquarium set-ups were compared. These included a recirculation system, three semi-recirculation systems and a flow-through system. Moreover, suitable growth substrates and different environmental conditions were compared. Sponge cultivation experiments were conducted in the commercial aquaculture farm Sæbýli in Eyrarbakki, Iceland, and later at the public aquarium in the Reykjavik animal park, Húsdýrargarðurinn, due to reports on positive growth of sponges in or close to public aquaria (Osinga *et al.*, 1999; Hausmann *et al.*, 2006). Many of the procedures for cultivation of *H. panicea* are based on early research of *ex situ* sponge cultivation by Barthel and Theede (1986), in addition to more recent research articles (Thomassen and Riisgård, 1995; Osinga *et al.*, 1999). One of the methods analysed in relationship to the sponge-associated community composition is described in **Paper I** and all methods are compared to each other in **Paper III**.

The recirculation system consisted of a 60 litre tank fitted with plastic foam biofilter mats and upwelling aerators for aeration and water movement. Microalgae cultures consisting of *Dunaliella* sp. and *Phaeodactylum tricornutum* were added to the tank twice a week and debris was removed regularly. The same tank set-up was also used as a semi-recirculating system, exchanging the water continuously at a rate of approximately 600 l per day. The third cultivation set-up consisted of a 360 l tank set-up with a bacterial moving bed biofilter, aeration, and artificial lighting set to 12 hours dark and 12 hours light. The water in this semi-recirculating system was continuously replaced at an exchange rate of two times the tank volume per day with sand-filtered coastal seawater. Additional feed was not added to this system. The fourth system consisted of the aforementioned set-up with an additional 1000 l tank. Live seaweed (*Laminaria* sp.) with intact holdfasts were placed into this tank and artificial light set to 12 hours dark and 12 hours light provided light intensity of 15 klux at the water surface for maintenance of the seaweed. Water was continuously exchanged at approximately four time the tank volume per day. The idea for this set-up arose from observing *H. panicea* growing to comparably large sizes when attached to seaweed in high wave-action coastal area around Iceland. Whereas all of the above-mentioned systems were situated at a commercial aquaculture farm receiving water from a coastal seawater pipeline, the fifth system was set-up at the Reykjavik animal park. This flow-through system consisted of a 15 l tank attached to the outlet of a public display aquarium which was continuously fed with seawater from a coastal borehole. Additional aeration supplied water movement and oxygenation. Water exchange rates were set to approximately one time the tank volume every two minutes.

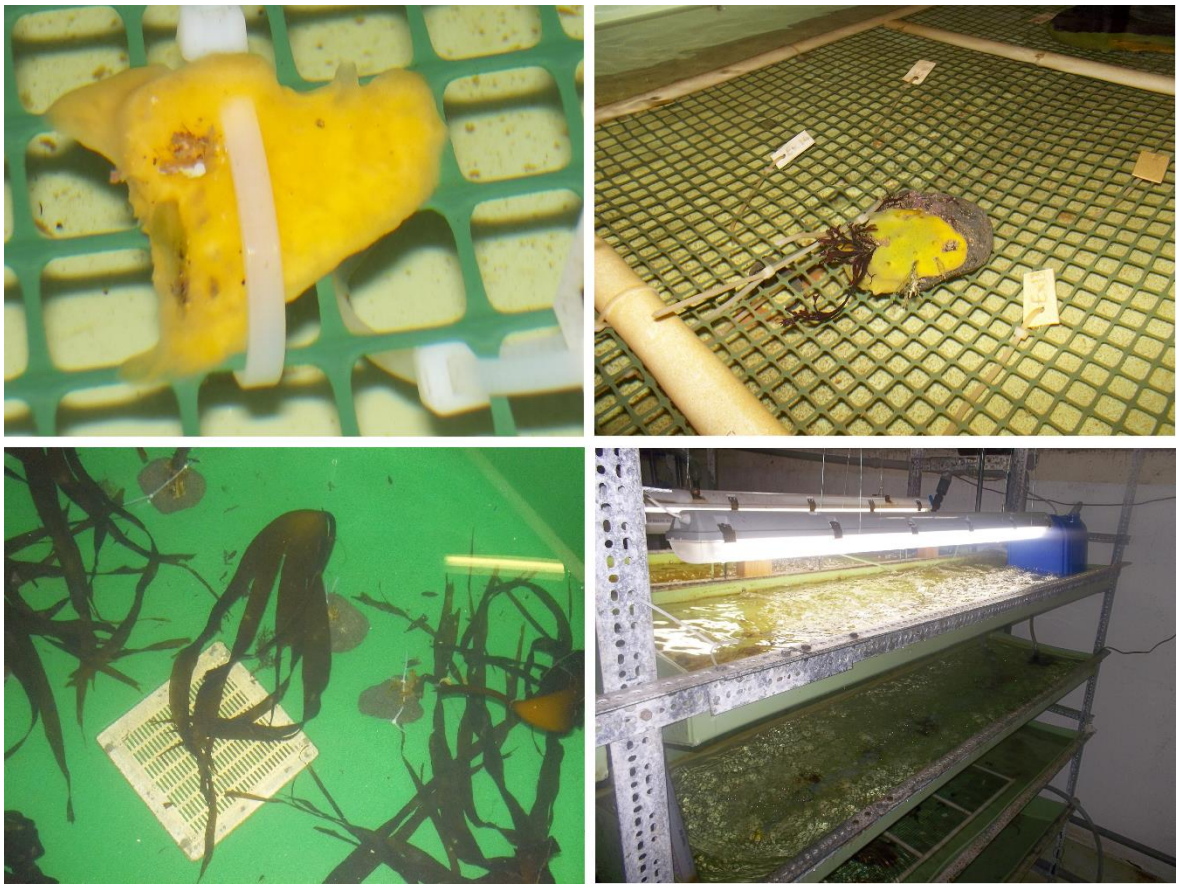


Figure 7: Clockwise from top-left: Sponge explant attached to plastic mesh. Single osculum can be seen on top right side of the explant; Sponge individual on stone collected in an intertidal area; Semi-recirculating tank set-up with lighting; Semi-recirculating tank set-up with sponge explants attached to Laminaria *sp.*

Sponges for the cultivation experiments were either kept on their natural substrate and introduced undamaged into the cultivation system or cut into explants. When cutting sponge explants, it was made sure that at least half of the ectoderm (the outer layer of body cells) was kept intact. Sponge explants were attached to different substrates, including nylon string and the thallus of *Laminaria* sp. or placed on the bottom of multi-well culture plates. The cut surface of the sponge explants usually healed and attached to the substrate within one to two weeks depending on the cultivation method. Sponges introduced into the aquaria on their natural substrate underwent a restructuring of their aquiferous systems, seen through a retraction of oscula and their appearance at other areas on the sponge surface. Pumping was observed by introducing food-grade dye above an osculum and observation of the exhalent jet. Growth as extension of the surface area, survival and attachment to the substrate were measured throughout the cultivation experiments. Water temperatures ranged between 8°C and 15°C, temperatures commonly found in coastal waters around Iceland. Build-up of ammonia and nitrite were regularly observed in the recirculating and semi-recirculating systems using commercial aquarium test kits from Seachem.

3.3 Microbial diversity analysis

3.3.1 16S rRNA gene amplicon sequencing

Microbial diversity in *H. panicea* samples was analysed through 16S rRNA gene amplicon sequencing and 16S rRNA Fluorescence In-Situ Hybridisation (FISH). For the former, two DNA extraction methods were tested: A method using grinding under liquid nitrogen and heat extraction according to Webster *et al.* (2010), and a DNA extraction kit (MasterPure DNA purification Kit, Epicentre) according to the manufacturer's instructions. For the latter method sponge samples were first homogenised in laboratory grade Millipore water by passing them through a 90-gauge syringe ten times, or until sponge cells were visibly disintegrated before continuing with the protocols for DNA purification. DNA quality and quantity were determined on a Nanodrop 1000 Spectrophotometer (Thermo Scientific) and by visual observation on a 1% agarose gel.

Partial 16S rRNA gene amplification was tested using high-fidelity Phusion polymerase (New England BioLabs) and two universal primer sets: The primer pair 515F/806R (5'-GTGYCAGCMGCCGCGGTAA-3' / 5'- GGACTACNVGGGTWTCTAAT-3') (Caporaso *et al.*, 2011; Quince *et al.*, 2011; Apprill *et al.*, 2015) used in the Earth Microbiome Project (Gilbert *et al.*, 2014) and the primer pair 341F/785R (5′-CCTACGGGNGGCWGCAG-3′ / 5′-GACTACHVGGGTATCTAATCC-3′) selected in an *in silico* study by Klindworth *et al.* (2013). The former primer pair covers the variable region V4 of the 16S rRNA gene, whereas the latter primer pair covers the V3 and V4 regions. For both primer pairs the PCR reactions were set up as follows: 5 µl 5X Phusion HF Buffer (New England BioLabs), 50 ng template DNA, 200 µM dNTPs, 0.2 µM of forward and 0.2 µM reverse primers containing Illumina overhang adapters, 0.2 µl Phusion High-Fidelity DNA polymerase and nuclease-free water to reach a total reaction volume of 25 µl. Conditions of the Tetrad 2 thermocycler (Bio-Rad) were set to: Initial Denaturation for 30 s at 98 °C followed by 30 cycles of denaturation for 10 s at 98 °C, annealing for 30 s at 52 °C and extension for 30 s at 72 °C. The final extension was set to 72 °C for 5 min.

A comparison of the number of ribosomal sequence variants (RSVs) detected by both DNA extraction methods (n = 3 samples per DNA extraction method) showed that the Epicentre kit allowed for an increased recovery of RSVs (Figure 8). Between both primer pairs (n = 1 sample per primer pair) the 341F/785R combination detected slightly more RSVs. We therefore chose to use the Epicentre kit DNA extraction method and the 341F/785R primer pair for further analysis, not least due to the fact that studies using the V4 region of the 16S rRNA gene can generally be compared to each other, while using a longer amplicon sequence allows for increased taxonomic resolution of the microbial communities.

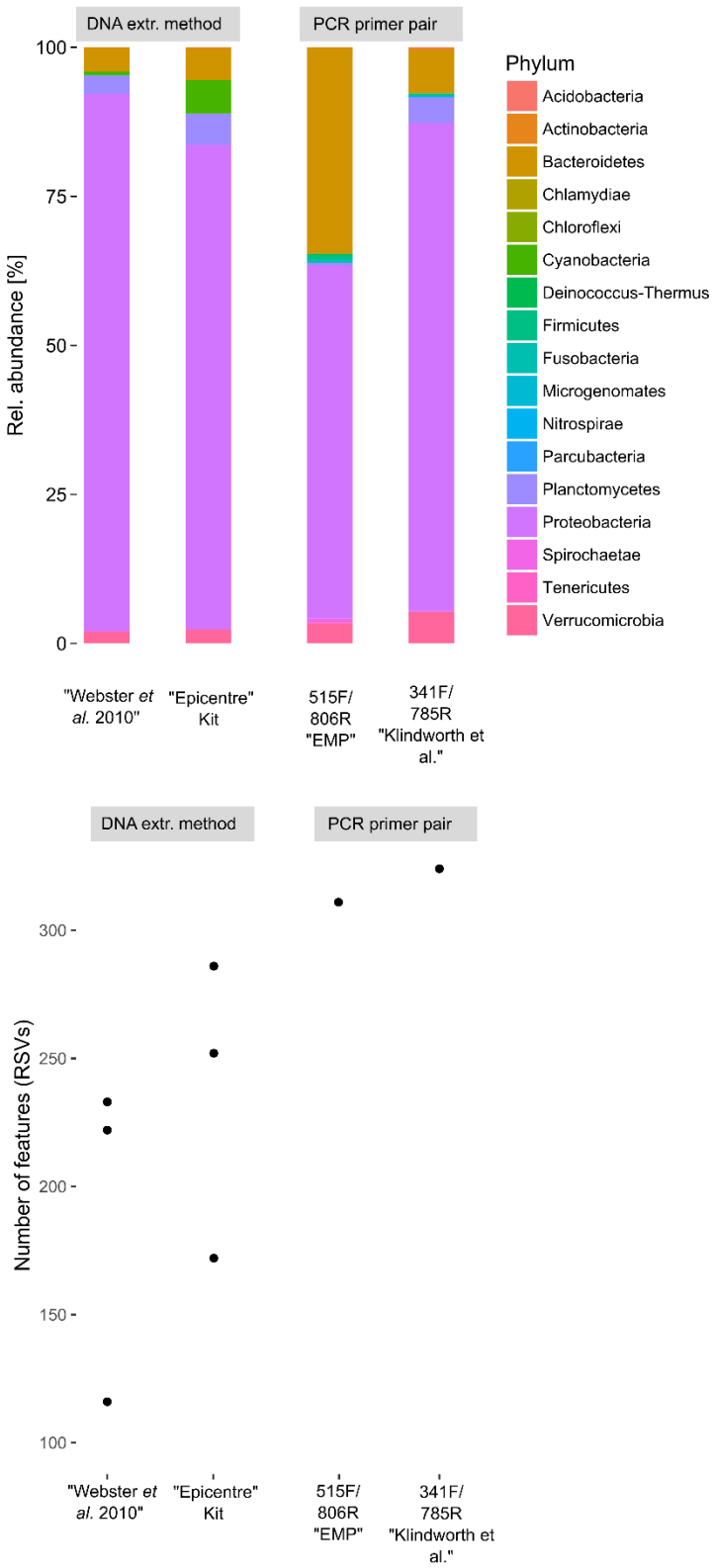


Figure 8: Comparison between two DNA extraction methods ("Webster et al. 2010" and "Epicentre Kit") and two primer pairs (515F/806R and 341F/785R) on samples of the sponge H. panicea. *Relative bacterial abundance at Phylum-level taxonomic assignment (top) and number of RSVs (bottom). Number of samples: 6 (DNA extraction method) and 2 (primer pair test).*

Seawater samples were collected in sterile 1 l containers, filtered on 0.2 µm Sterivex filters (Merck Millipore) and stored at -80 °C. For seawater DNA extraction, the filter containers were opened and filters containing residue were ground in lysis buffer from the Epicentre DNA purification kit prior to following the DNA extraction and PCR amplification as mentioned above.

All PCR amplicons were multiplexed using Nextera XT barcodes (Illumina) and were sequenced on a MiSeq desktop sequencer (Illumina) with v3 chemistry and 2 x 300 cycles over multiple runs. Raw sequences were demultiplexed and processed using to two bioinformatic pipelines according to the desired output. For detection of RSVs the pipeline DADA2 (Callahan *et al.*, 2016) was used. For detection of operational taxonomic units (OTUs) at 97 % sequence identity we used the pipeline UPARSE (Edgar, 2013). Both methods have their advantages and disadvantages. Detection of RSVs is the preferred method for modern microbial diversity analysis due to their higher fidelity and closer resemblance to actual biological sequences (Callahan *et al.*, 2017). However, clustering sequences into OTUs can be advantageous when many sequence variants (i.e. sub-species or strains) of the same species are present or when comparing presence and absence of species across different dataset, thus allowing for a simpler visualisation and more conservative estimation of the microbial diversity. Taxonomic assignment of RSVs was conducted with the RDP classifier (Wang *et al.*, 2007) implemented in the DADA2 package against the SILVA SSU version 128 dataset (Quast *et al.*, 2013). For OTUs, taxonomic assignment was performed with the SINTAX algorithm (Edgar, 2016) against the SILVA v123 LTP database (Yilmaz *et al.*, 2014). Statistical analysis and visual presentation of the data was conducted with the statistical software R (R Core Team, 2017) implemented in RStudio (RStudio Team, 2016) using packages phyloseq (McMurdie and Holmes, 2013), ggplot2 (Wickham, 2009) and other packages detailed in **Paper I**, **II** and **III**.

To detect the presence of bacteria associated with Icelandic *H. panicea* in other sponge species or in offshore seawater, the *H. panicea* microbial diversity data was compared against the Sponge Microbiome Project dataset (Moitinho-Silva, Nielsen, *et al.*, 2017), containing the to-date largest collection of sponge associated bacteria and against a database of seawater samples collected in conjunction with the annual cruise that the Icelandic research vessel "Bjarni Sæmundsson" carried out in May 2011, 12 and 13 (unpublished data). All partial 16S rRNA sequences were compiled into a blast database using the *makeblastdb* command of the BLAST+ suite (Camacho *et al.*, 2009) and used as a reference to query sequences in the *H. panicea* dataset using the *blastn* command.

3.3.2 16S rRNA gene sequence variants of "*Ca*. H. symbioticus"

16S rRNA gene sequence variants of the dominant symbiont "*Ca*. H. symbioticus" between geographical locations was measured by amplifying the near full length 16S rRNA gene of selected samples using primer pair F9/R1510 (5′-GAGTTTGATCCTGGCTCAG-3′ / 5′-GGTTACCTTGTTACGACTT-3′) (Baker *et al.*, 2003; Reynisson *et al.*, 2009). The same PCR reaction and thermocycler conditions were used as mentioned above, apart from that the extension time was increased to 1 min due to the increased size of the amplicons. Amplicons were purified, fragmented, indexed and re-amplified using the Nextera XT Kit (Illumina), and sequenced on an Illumina MiSeq system with v3 chemistry and 2 x 300 cycles. Raw reads were quality filtered, pair end reads merged and assembled in the bioinformatic software Genious (Biomatters). Near full length 16S rRNA gene sequences of

"*Ca*. H. symbioticus" were extracted, based on sequence similarity to its smaller 16S rRNA gene fragment amplicon and were aligned with MUSCLE (Edgar, 2004) against reference sequences of previously published "*Ca*. H. symbioticus" sequences retrieved from GenBank. Alignments between 16S rRNA gene sequences of "Ca. H. symbioticus" from all four geographical locations used in this study and reference sequences were visualised in Boxshade (Hofmann and Baron, 1996).

3.3.3 16S rRNA Fluorescence *In Situ* Hybridisation

In order to visualise the presence of cells of the dominant symbiont "*Ca*. H. symbioticus" in the sponge body of *H. panicea* we designed FISH probes based on its full length 16S rRNA gene sequence (Table 2). Probes were designed in the software package ARB (Ludwig *et al.*, 2004) and tested for specificity against the NCBI nr/nt sequence database (NCBI Resource Coordinators, 2017). In addition, the universal bacterial probes EUB338 I, II and III (Amann *et al.*, 1990; Daims *et al.*, 1999) were used to detect bacterial cells in the sponge body. The procedure for 16S rRNA FISH was conducted according to the description in Pernthaler *et al.* (2001) on formaldehyde fixed cryo-cut sponge sections. In short, sponge samples were fixed in 4 % paraformaldehyde in 1 x Phosphate-buffered saline (PBS) for 12 h at 4 °C and subsequently washed in PBS and stored in 70% ethanol for up to one year. Samples were rehydrated, saturated in 30% sucrose solution in 1 x PBS, covered in tissue freezing medium (Leica Biosystems) and frozen on crushed dry ice. Sponge sections of approximately 5 µm thickness were cut on a Leica CM1800 cryotome set to -35 °C using MX35 Ultra microtome blades (Thermo scientific). Sponge sections were dehydrated in successive baths of ethanol up to 100 % and subsequently air dried. Dried sponge sections on microscope slides were hybridized with Cyanide 3 labelled probes Rhd473 and Rhd1136 and Alexa488 labelled universal bacterial probes EUB338 I – III (each at concentrations of 5 ng DNA μl^{-1}) in hybridization buffer (900 mM NaCl, 20 mM Tris-HCl, 60 % formamide and 0.01 % SDS in deionized water) for 2 h at 46 °C. Sponge sections were subsequently washed in wash buffer (14 mM NaCl, 20 mM Tris-HCl, 5 mM EDTA and 0.01 % SDS in deionised water) for 15 min at 48 °C, rinsed in deionised water, air dried and covered with Fluoroshield antifade containing DAPI (Sigma). FISH Images were taken on a model BX51 epifluorescence microscope (Olympus) and processed with imaging software daime (Daims *et al.*, 2006).

Table 2: Specific 16S rRNA FISH probes designed and used in this study

Probe	**Sequence (5' - 3')**	**Specificity**
Rhd473	AGCCGGGACTTCTTCTACTG	*"Ca.* Halichondribacter symbioticus"
Rhd1136	CTCTTTAGTGTCCCCAACTG	*"Ca.* Halichondribacter symbioticus"

3.4 Symbiont genome analysis

3.4.1 Metagenomic sequencing, "*Candidatus* Halichondribacter symbioticus" genome binning and annotation

To analyse the functional properties of the dominant bacterial symbiont in *H. panicea*, "*Ca*. H. symbioticus", the bacterial fraction was separated from sponge cells by grinding sponge samples in magnesium- and calcium-free artificial seawater and multiple centrifugation steps. Different metagenomic library preparation methods were tested, including the TruSeq DNA Nano and Nextera Mate Pair library preparation kits from Illumina. However, these did not yield high quality libraries, possibly due to low DNA quality of the samples which could not be improved despite using different DNA extraction and purification methods (e.g. phenol-chloroform extraction, Mobio PowerSoil kit). Finally, we constructed a metagenomic library using the Nextera XT (Illumina) kit for low DNA input quantities according to the manufacturer's instructions. Metagenomic libraries were subsequently sequenced on a MiSeq desktop sequencer with paired-end modality, v3 chemistry and 2 x 300 cycles.

Raw sequences were quality filtered using the software Trimmomatic (Bolger *et al.*, 2014) and assembled using the metagenomic algorithm (Nurk *et al.*, 2017) of the SPAdes software (Bankevich *et al.*, 2012). The draft genome of "*Ca*. H. symbioticus" was extracted from the metagenome using the genome binning *Multi-metagenome* workflow (Albertsen *et al.*, 2013) with manual curation of binned contigs. Genome completeness was estimated through the number of single copy marker genes (SCMGs) detected with HMMER (Eddy, 2011) against hidden Markov models of 107 bacterial proteins and the expected number of SCMGs for the *Alphaproteobacteria* class (Albertsen *et al.*, 2013). ORFs predicted through MetaProdigal (Hyatt *et al.*, 2012) were annotated against the COG (Tatusov *et al.*, 2000; Galperin *et al.*, 2015) and KEGG (Kanehisa and Goto, 2000) databases using RPS-BLAST+ and BlastKOALA (Kanehisa *et al.*, 2016) respectively. In order to determine the relative abundance of "*Ca*. H. symbioticus" in the metagenomes, partial 16S rRNA gene sequence were extracted using Meta_RNA(H3) (Huang *et al.*, 2009) and sequences larger than 100 bp taxonomically classified against the SILVA v123 LTP database using the SINTAX algorithm.

3.4.2 Phylogenomic inference and comparative genome analysis

Phylogenomic inference of "*Ca*. H. symbioticus" was determined based on 31 conserved proteins predicted through Amphora2 (Wu and Scott, 2012) in "*Ca*. H. symbioticus" and related reference genomes from NCBI which were selected based on their 16S rRNA gene similarity. All 31 proteins from the genomes were individually aligned using the MUSCLE algorithm (Edgar, 2004) implemented in ARB (Ludwig *et al.*, 2004). Concatenated alignments were used to construct a maximum-likelihood phylogenetic tree using the PhyML algorithm (Guindon and Gascuel, 2003) as described previously (Wu and Eisen, 2008).

Comparitive genome analysis against closely related genomes and four selected sponge symbiont genomes from the *Alphaproteobacteria* class was performed in statistical software R using COG and KEGG Orthologue (KO) tables. The KEGG mapper (https://www.genome.jp/kegg/mapper.html) was used to reconstruct functional KEGG

pathways and detect complete or incomplete pathways in "*Ca*. H symbioticus" and the reference genomes. Data visualisation and statistical analysis was conducted in R implemented in RStudio and other packages and custom scripts detailed in **Paper II**.

3.4.3 Ancestral gene content prediction and gene expression

In order to evaluate gene loss and gain events of the "*Ca*. H symbioticus" genome in relation to its closest free-living relatives, we performed ancestral gene content prediction using the software Count (Csuos, 2010). Gene loss and gain events were predicted with asymmetric Wagner parsimony (Swofford and Maddison, 1987; Csűrös, 2008) with a gain penalty of 1.3 and Dollo parsimony (Farris, 1977) on COG distribution of "*Ca*. H symbioticus" and reference genomes. Both methods yielded comparable results (R>0.79, Pearson correlation).

Gene expression of "*Ca*. H. symbioticus" was calculated based on publicly available RNA-seq data from *H. panicea* under BioProject PRJNA394213, accession SRX4378335. These samples had been collected, processed and sequenced as described in Waldron *et al.* (2018). The relative abundance of "*Ca*. H. symbioticus" in the dataset was measured as for the previously mentioned *H. panicea* metagenomes. Raw sequences were quality filtered and trimmed using Trimmomatic and then mapped against ORFs of the "*Ca*. H. symbioticus" HS1 draft genome using Bowtie2 (Langmead and Salzberg, 2012). Read coverage was calculated using the depth function in SAMtools (Li *et al.*, 2009) and summarised by assigned COG and COG categories respectively.

3.5 Bacterial isolation, co-cultivation and strain characterization

3.5.1 Sponge-bacteria co-cultivation and bacterial enrichment

For the sponge-bacteria co-cultivation experiment we used a Bio-Dot SF microfiltration apparatus (BioHit) to separate sponge explants from the bacterial inoculum while allowing both to grow in close proximity to each other. The idea behind this method was to create an inoculum of sponge-associated bacteria with one or more bacterial cells through serial dilution which would get distributed onto solid media and be separated through a membrane with 0.2 µm pore size from the sponge. This proximity to the sponge would allow bacteria, that necessitate an environment similar to that found inside the sponge body or require certain metabolites from the sponge, to grow while normally not being able to proliferate in complete absence of their host. This could facilitate both the cultivation of host-dependent symbionts as well as increase the diversity of isolated sponge-associated bacteria. Similar methods, termed *in situ* diffusion growth chambers have been used previously, most notably for the isolation of soil microbes (Nichols *et al.*, 2010), but also for the isolation of sponge-associated bacteria (Steinert *et al.*, 2014) and have shown promising results for the isolation of previously uncultivated bacteria. In comparison to the work by Steinert *et al.* (2014), we propose a method which can be applied outside of the sponge's natural environment in a laboratory setting, thereby facilitating the procedure of co-cultivation. The use of a multi-chamber device also allows for a higher throughput of the co-cultivation procedure, and the application of extinction dilution in several wells at once.

For the co-cultivation experiment a 10^5 serial dilution of a bacterial suspension was made by rinsing 10 g of a wild *H. panicea* specimen with sterile artificial seawater (450 mM NaCl, 10 mM KCl, 9 mM $CaCl_2$, 14 mM $MgCl_2$, 8 mM $MgSO_4$) to remove bacteria attached to the outer sponge surface and grinding the sponge in sterile calcium-magnesium free seawater (450 mM NaCl, 10 mM KCl). The sponge suspension was centrifuged at 500 x g for 2 minutes to remove the majority of sponge cells, which was visually observed under a light microscope. The bottom part of the microfiltration apparatus was then filled with Marine Agar (Difco) under sterile conditions through the effluent tube and until the medium reached the surface of the well grid. The medium was left to harden and then 3 µl of the sponge bacterial suspension was pipetted onto each well (Figure 9). The bottom part of the apparatus was then covered by 0.2 µm nylon membrane filters (Whatman) and seven layers of 25 µm filter paper (Miracloth, Merk-Millipore) to prevent the membrane from clogging. Then ca. 0.125 cm^3 sized *H. panicea* explants were cut and distributed into the 48 wells on the top part of the apparatus which was placed on top of the bottom part and tightened using the screws in each corner. This effectively created a sealed chamber which prevented the exchange of bacterial cells in and out of the device whereas metabolites would be able to pass through the 0.2 µm membrane between the bacterial inoculum and the sponge explants. The complete device was then deployed into a 15 l tank receiving seawater from a large public aquarium as described in section 3.2 (Figure 9). After 10 weeks, the co-cultivation chamber was removed from the aquarium and opened. Visually detectable colony growth in the wells and attached to the underside of the membrane was collected and stored in a microcentrifuge tube at -80 °C for sequencing and microbial diversity analysis as described in section 3.3. In addition, a sample from the original sponge used for the experiment as well as its explants after 3, 5 and 10 weeks were collected, rinsed and stored at -80 °C for

microbial diversity analysis to evaluate if the original bacterial diversity in the sponge, particularly the dominant symbiont in *H. panicea* where still present in the sponge body during cultivation, a prerequisite for co-cultivation of this symbiont.

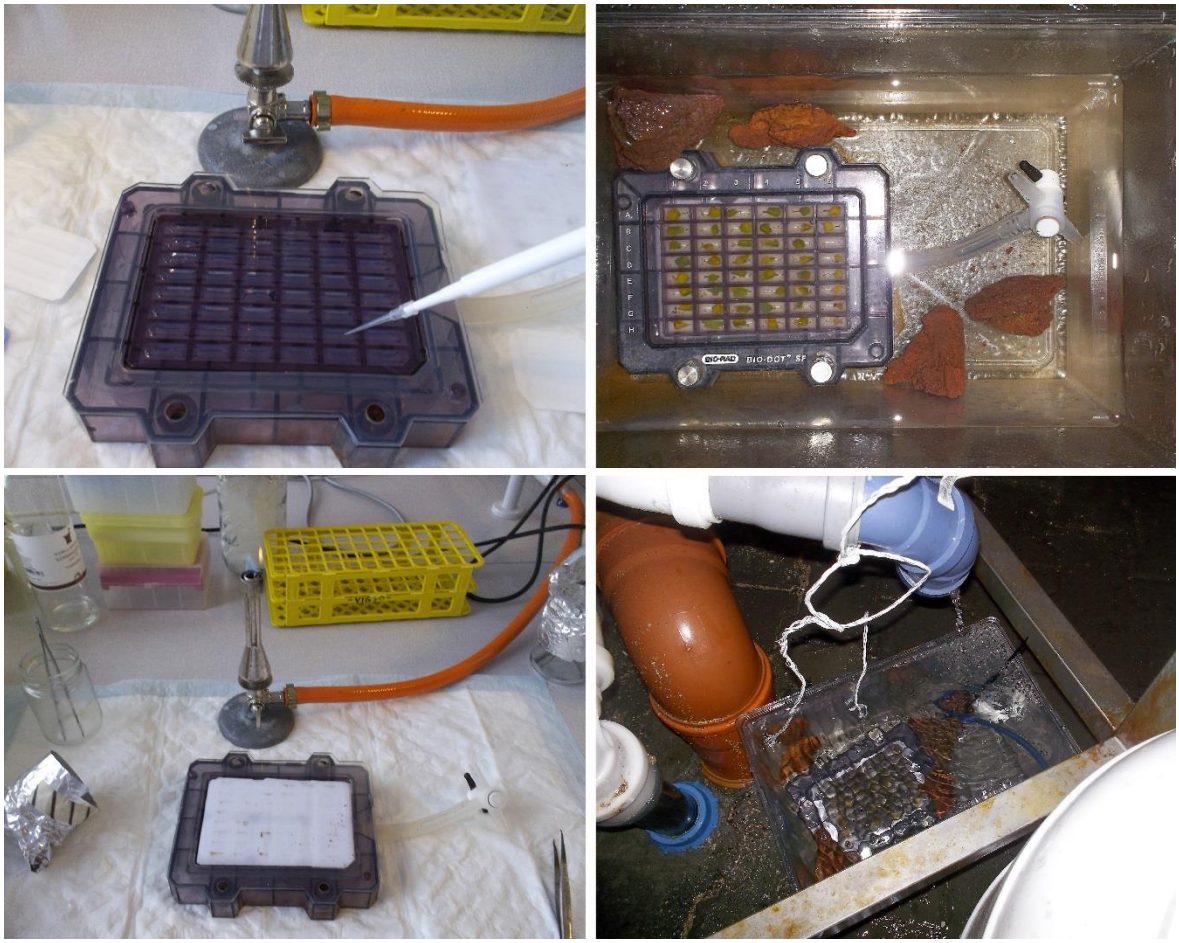


Figure 9: Clockwise from top-left: Inoculation of the wells of the sponge-bacterial co-cultivation device with a bacterial suspension; Deployment of the device in a 15 l tank, with H. panicea *explants placed into each compartment of the device, separated from the bacterial inoculum through a membrane and filter paper; Co-cultivation tank receiving water from the outlet of a public display aquarium; Recovery of the co-cultivation device after 10 weeks and removal of the membrane under sterile conditions.*

3.5.2 Bacterial isolation through standard plating method

The same sponge bacterial suspension used for the co-cultivation experiment was also used to inoculate standard solid media plates to evaluate the general cultivability of bacteria found in *H. panicea*. The plating media used was Marine Agar and Starch Yeast-Extract Peptone Sea Water Agar (10 g potato starch, 4 g Bacto Yeast Extract, 2 g Bacto Peptone, 15 g Bacto Agar in 1 l 0.2 µm filtered seawater), two general media used for cultivation and enumeration of marine bacteria. In addition, both media were prepared with an aqueous sponge extract made from a *H. panicea* suspension filtered through a 0.2 µm filter and added to the media at 0.5 % (v/v) after autoclaving the media. Adding sponge extract to media for the cultivation of sponge-associated media as previously been conducted to improve the number of isolated bacteria (Kennedy *et al.*, 2009; Abdelmohsen *et al.*, 2010), although this was not always correlated to an increase of the isolated bacterial diversity (Webster *et al.*, 2001).

Inoculated plates were incubated at 10 and 22 °C and colonies were picked and re-isolated on Marine Agar plates at least three times to assure that the strains were isolated. In order to evaluate the diversity of isolates strains without sequencing a marker gene of each isolate, we used a MALDI-TOF biotyper to rapidly characterize the relative relatedness of all strains through peptide mass spectra (Sogawa *et al.*, 2011). This was conducted on a Microflex MALDI-TOF mass spectrometer (Bruker) after formic acid extraction of each strain. Only one strain from distinctly distant clades on a main spectra profile dendrogram were selected for sequencing of the partial 16S rRNA gene through Sanger sequencing. In short, a fragment of the 16S rRNA gene was amplified using the primer pair F27/R806 (5'-AGAGTTTGATCMTGGCTCAG-3' / 5'- GGACTACVSGGGTATCTAAT-3) and sequenced on a 3730 DNA Analyser (Applied Biosystems, Hitachi). In order to compare the sequences to those from the microbial diversity analysis of the co-cultivation enrichment, sequences were trimmed to position 341 and 785 (relative to the *E. coli* 16S rRNA gene) using the bioinformatics software Geneious (Biomatters) and classified taxonomically using the SINTAX algorithm against the SILVA v123 LTP database.

Sequences from the co-cultivation enrichment and isolates retrieved from standard plating were compared to each other based on class level assignment and a phylogenetic tree with all sequences was constructed in ARB using the PhyML algorithm. To determine if sequences from the enrichment and the isolates were related to previously isolated sequences from sponges, all sequences were compared to entries in the NCBI nr/nt database using BLAST.

3.5.3 Secondary metabolite clusters and antimicrobial activity assay of selected strains

Three strains isolated through the standard plating method on Marine Agar were selected for antimicrobial activity tests. These strains showed low 16S rRNA sequence similarity to previously cultivated and described bacteria possibly indicating the presence of novel species with yet uncharacterised traits. Strain Hp12 had a 16S rRNA gene sequence similarity of 91 % to *Marinobacter litoralis*, strain Hp32 had a 97 % sequence similarity to *Tenacibaculum ovolyticum*, and strain Hp36 had a 95 % sequence similarity to *Endozoicomonas numazuensis*. All three strains were subjected to whole genome sequencing to mine the genome for potential secondary metabolite clusters. In short, DNA from a colony of each strains was extracted using the Epicentre DNA purification kit. Purified DNA was fragmented, tagged and amplified using the Nextera XT kit from Illumina. All three libraries were barcoded and sequenced on the same run using a MiSeq desktop sequencer with v3 chemistry and 2 x 300 cycles. Raw reads were trimmed using the software Trimmomatic, merged using FLASH (Magoc and Salzberg, 2011) and both merged and unmerged reads were assembled in SPAdes. All draft genomes were searched for secondary metabolite gene clusters using the bacterial version of the AntiSMASH web server (Blin *et al.*, 2017).

Extracts of strains Hp12, 32 and 36 for antimicrobial activity assays were prepared by growing the strains in Marine Broth (Difco) at 22 °C until the late exponential phase and extracting both the aqueous and organic phase of the cultures. In short, bacterial cultures were sonicated for 3 minutes in a Sonifier 250 (Branson Ultrasonics) with the duty cycle on constant and output control on 4. Cultures were kept in an ice bath during sonification. Water-saturated butanol with 10 % (v/v) methanol was added in equal volumes to the culture and shook vigorously, followed by centrifugation for 10 min at 10.000 x g and 4 °C. the

upper organic phase was separated and the bottom phase was subjected to a second extraction with water-saturated butanol containing 10% methanol. The upper phase from the second extraction was pooled with the organic phase from the first extraction. The bottom aqueous phase was separated from the residual pellet, froze at -80°C and freeze-dried on a LyoLab 3000 (Heto-Holten). The organic extract was filtered on a 0.45 µm PTFE filter (Merck Millipore) and concentrated on a rotary evaporator at 40°C. The dry organic extract was dissolved in Dimethyl sulfoxide (DMSO) to obtain a final concentration of 25 mg ml^{-1}.

The antimicrobial capacity of the extracts was tested against the test strains *Escherichia coli* (ATCC 8739), *Staphylococcus aureus* (ATCC 9027) and *Pseudomonas aeruginosa* (ATCC 8739) on a Bioscreen C (Oy Growth Curves) at 37°C. The test strains were inoculated in duplicates into 300 µl of Brain-Heart-Infusion media containing each extract at concentrations of 0, 0.16, 0.33, 0,66 and 0.98 % (v/v). Growth curves were compared to positive controls containing only growth media, as well as growth media with DMSO at 0.98 % (v/v).

3.5.4 Characterisation of strains Hp12, Hp32 and Hp36

Strains Hp12, Hp32, and Hp36 were characterized according to the guidelines outlined in Tindall *et al.* (2010) using a polyphasic approach. In order to specify if the strains were sponge-specific, their 16S rRNA genes were compared to the microbial community of all *H. panicea* samples collected in this study and also against the SRA database using IMNGS (Lagkouvardos *et al.*, 2016). All strains were routinely grown on MA or MB at 20°C. Cell morphology was observed by light microscopy (BX51, Olympus) and for Hp12 also by field emission scanning electron microscopy (Supra 25, Zeiss). The Gram reaction was performed using a test kit (BD BBL Gram Stain Kit) according to the manufacturer's instructions. The growth range was determined on MA at temperatures from 4 to 50°C. Optimal growth temperature, pH and NaCl concentrations were determined in MB using a Bioscreen C. Growth under anaerobic conditions was tested in a GasPak (BD) on MA and MA supplemented with nitrate, nitrite, fumerate, sulphite, thiosulfate and sulfate. Oxidase and catalase activity tests were performed on DrySlide reagent slides and with catalase reagent droppers (BD) respectively. In addition, hydrolysis of casein, Tween 20 and starch, as well as susceptibility to 13 antibiotics was investigated. API ZYM and API 20NE systems (bioMérieux) were used to determine enzymatic activities and other biochemical and physiological traits. Analysis of cellular fatty acids, respiratory lipoquinones and polar lipids for strain Hp12 was carried out from freeze dried cells by the Identification Service of the German Collection of Microorganisms and Cell Cultures (DSMZ), Braunschweig, Germany.

The DNA G+C content of all strains was determined through the draft genome sequence in the software suite Geneious. Digital DNA-DNA hybridization (dDDH) was performed by the Genome-to-Genome Distance Calculator on http://ggdc.dsmz.de (Meier-Kolthoff *et al.*, 2013) and Average Nucleotide Identity (ANI) was calculated on EZBioCloud (Yoon, S. Ha, *et al.*, 2017). The 16S rRNA gene sequence was extracted from the genome using RNAmmer (Lagesen *et al.*, 2007). The full length 16S rRNA gene sequence was aligned to the global SSU SILVA alignment with SINA (Pruesse *et al.*, 2012) and merged with the LTPs132 database (Yilmaz *et al.*, 2014) using the ARB software. Related sequences were selected and neighbour joining (ARB), maximum parsimony (Felsenstein, 2004) and maximum-likelihood (Guindon and Gascuel, 2003) trees constructed. Tree topology was tested with 30, 40 and 50 % positional conservatory filters and 100 bootstraps.

4 Results and Discussions

4.1 Bacterial diversity in *Halichondria panicea* (Paper I)

4.1.1 Spatio-temporal stability of the sponge-associated bacterial community

In **Paper I** we analysed the bacterial diversity of 72 *H. panicea* samples from four different geographical locations and across various times of the year, showing that only a single bacterial taxon was present in all sponge samples whereas other taxa were either only represented locally or were temporally transient (Figure 10). The single stable taxon was also the dominant taxon in all samples, constituting approximately 63 ± 20 % of the relative abundance. This was in accordance with previous studies on the bacterial composition of *H. panicea* which showed the presence of the same dominant taxon in *H. panicea* samples from the North Sea and Baltic Sea (Althoff *et al.*, 1998; Wichels *et al.*, 2006; Naim *et al.*, 2014; Steinert *et al.*, 2017). However, the nature of the remaining bacterial diversity detected in *H. panicea* was previously unknown. Using a high-throughput sequencing approach and high sequencing depth of approximately 20,000 reads per sample we could show that even at a very low relative abundance *H. panicea* from the study area did not consistently contain any other symbionts over time and location.

It was however not possible to conclude that, apart from the dominant symbiont, all other bacterial taxa were transient seawater bacteria which were either passing through the aquiferous system by means of the sponge pump or incorporated as feed organisms. On average and after excluding the dominant bacterial taxon, over half of the reads detected in the sponge body were not detected in the surrounding seawater, meaning that the sponge harboured additional bacteria associated specifically with the sponge body. Only two sponge sample sets showed higher similarity to their respective seawater sample than to the other sponge samples based on calculation of Bray-Curtis dissimilarities (Figure 10), an interesting observation which might be connected to tissue sloughing previously described in *H. panicea* (Barthel and Wolfrath, 1989). This finding could indicate that most of the bacterial diversity found on the sponge is confined to the outer layer of the sponge which is subjected to the process of tissue sloughing.

Due to the stable relationship between the dominant bacterial symbiont and the sponge *H. panicea*, not only in Icelandic waters but also in the North Sea and Baltic Sea, we decided to propose a candidate status with the name "*Candidatus* Halichondribacter symbioticus" for the symbiont to facilitate future comparative work on this sponge-bacterium symbiosis. Phylogenetic inference of "*Ca*. H. symbioticus" using the near full length 16S rRNA gene sequences from all sampling sites and reference sequences showed that "*Ca*. H. symbioticus" from South-West Iceland was closer related to "*Ca*. H. symbioticus" found in *H. panicea* from the North Sea and Baltic Sea, which fell into a monophyletic clade, than to "*Ca*. H.

symbioticus" from North Iceland which branched outside of this clade. This distinct difference in ribosomal sequence variants between the same sponge symbiont could allow for a better resolution of the coevolutionary processes of this sponge-symbiont relationship, especially in relation to the COI gene evolution of the sponge.

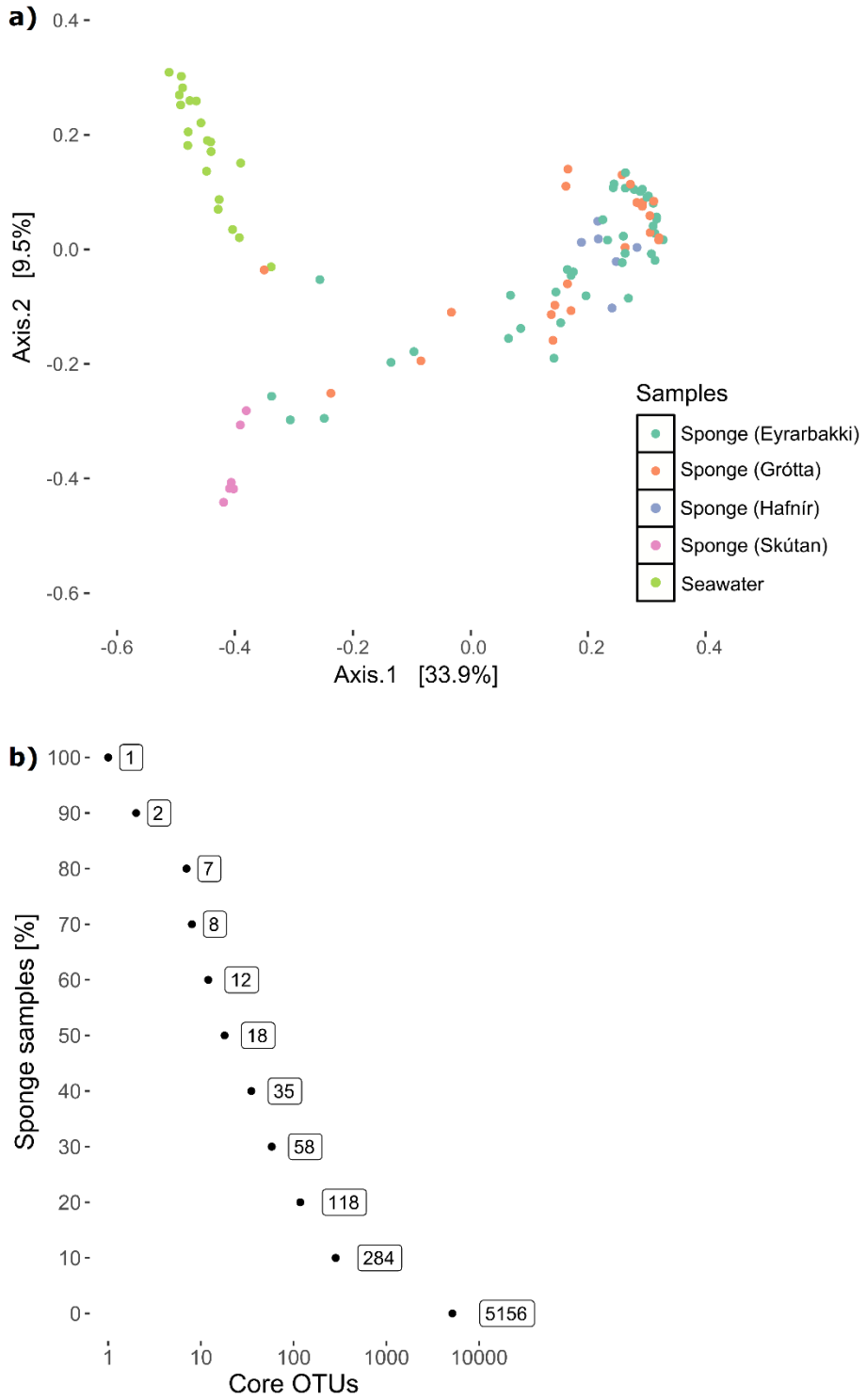


*Figure 10: (a) Principle coordinate analysis of Bray-Curtis dissimilarity indices of all wild sponge and seawater samples collected in Paper I; (b) Number of OTUs detected in proportion of sponge samples. Only a single OTU, referring to "*Ca*. H. symbioticus", was detected in 100% of the samples.*

4.1.2 Presence of dominant symbiont "*Candidatus* Halichondribacter symbioticus" within the sponge body

Using 16S rRNA FISH with specific and universal bacterial probes we were able to confirm that "*Ca*. H. symbioticus" was highly abundant in the sponge body, but also show that its distribution was uniform throughout the mesohyl without areas of particular high or low abundance (Figure 11). Compared directly against 16S rRNA gene amplicon sequencing, the fluorescence microscopy method revealed a higher relative abundance of "*Ca*. H. symbioticus" in relation to other bacteria present in the sponge body. This discrepancy between the two methods might have been due to a higher detection of bacterial sequences in the aquiferous system and on the sponge surface which were removed during the sample processing for FISH.

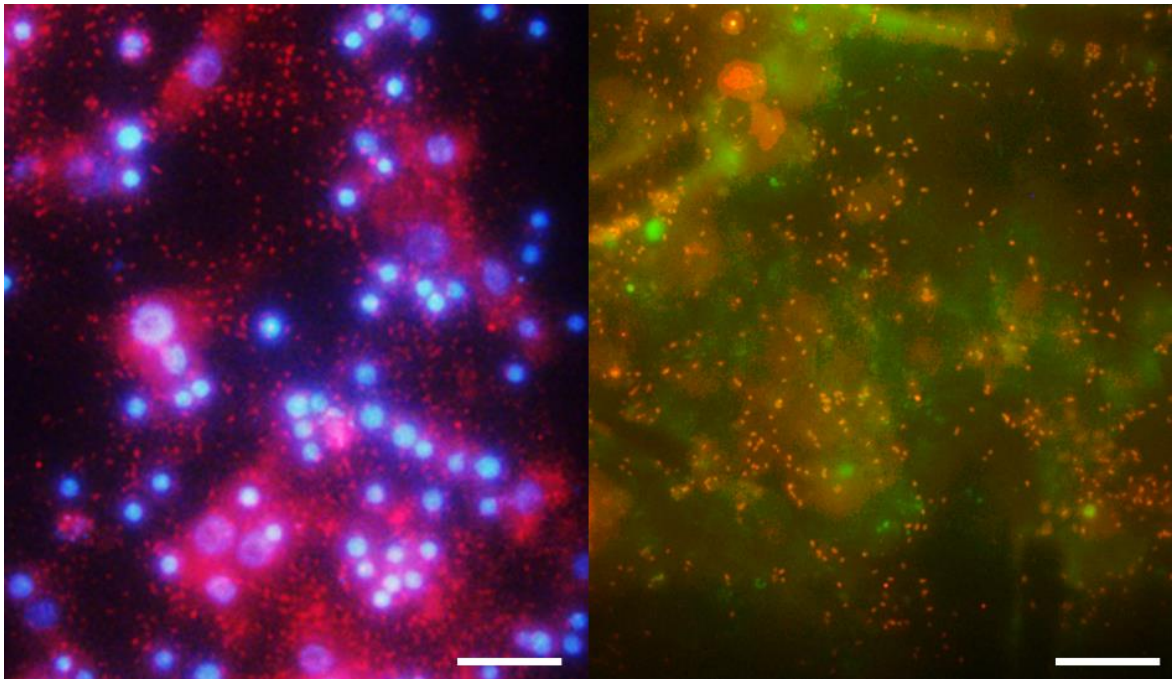


Figure 11: Two 16S rRNA FISH images of H. panicea *mesohyl. Left: Cells of "*Ca*. H. symbioticus" (red) and sponge cell nuclei (blue). Right: Cells of "*Ca*. H. symbioticus" (orange), other bacterial cells (green). Bar = 10 µm*

Another important discovery was that when sponges were kept *ex situ* under less favourable conditions, "*Ca*. H. symbioticus" did not disappear from the sponge body entirely but was restricted to clusters of few cells sparsely distributed throughout the mesohyl. This indicates both that "*Ca*. H. symbioticus" is dependent on its host metabolism through the provision of nutrients and that this sponge-bacterium symbiosis can be maintained even under less favourable conditions. In addition, the latter point could be linked to vertical transmission of the symbiont during the larval stage of the sponge.

4.1.3 Global distribution and sponge species specificity of "*Candidatus* Halichondribacter symbioticus"

A comparison of the 16S rRNA gene sequence of "*Ca*. H. symbioticus" against the so far largest dataset of sponge-associated bacteria within the framework of the Sponge Microbiome Project, showed that *H. panicea* was the only sponge species which contained

the symbiont at a high relative abundance of above 66 %. However, "*Ca*. H. symbioticus" was also found at a much lower abundance in other sponge species from various locations in the Atlantic, Pacific and Indian Ocean. This led to the conclusion that "*Ca*. H. symbioticus" is an opportunistic sponge-associated bacterium which thrives exceptionally well in *H. panicea*. Understanding why this symbiont proliferates so well in *H. panicea* and constitutes the dominant bacteria in the sponge body should be the focus of future studies.

Apart from the global distribution of the sponge symbiont, we found that several sponges marked as *H. panicea* in the dataset did not contain "*Ca*. H. symbioticus" at all. This applied to sponges identified as *H. panicea* from the East Pacific and Mediterranean Sea, but also to locations in the North Sea where "*Ca*. H. symbioticus" had previously been detected in *H. panicea*. Whereas this could be due to misidentification of *H. panicea* due to its wide range of morphological forms, we point out that further studies using molecular sponge markers would be needed to address the phylogeny of the *H. panicea* species complex and in this context the absence or presence of "*Ca*. H. symbioticus".

4.2 Genome analysis of "*Candidatus* Halichondribacter symbioticus" (Paper II)

4.2.1 Genome comparison of "*Candidatus* Halichondribacter symbioticus" draft genome HS1

In **Paper II** we present a draft genome of "*Ca*. H. symbioticus" from *H. panicea* collected in South-West Iceland. The draft genome, designated HS1, was 2.9 Mbp in size with a GC content of 49.4 %. Despite a high number of contigs (N50: 2152) the genome completeness was estimated to be approximately 93 %, with key functional pathways such as the citrate cycle and peptidoglycan biosynthesis present and complete. Phylogenomic analysis showed that HS1 was closely related to two genomes assigned to the genus *Amylibacter* and the genome HTCC2255 of an uncharacterised *Rhodobacterales* bacterium. A comparison with these related reference genomes, showed that 983 out of 1170 COGs found in HS1 were shared between all four genomes, indicating that the symbiont genome still contained a core genome with its free-living relatives. Compared against a larger reference dataset of 14 related free-living bacterial genomes and four selected *Alphaproteobacteria* sponge symbionts previously extracted from the metagenome of the sponge *Aplysina aerophoba*, we could show that HS1 had a similar COG category distribution with HTCC2255 and a sponge symbiont designated bin131. Since bin131 was only distantly related to HS1 this indicates a convergent evolution of functional features between the two sponge symbionts.

Compared against the genomes of the four closely related free-living reference bacteria, HS1 had a reduced number of regulatory genes coding for transcription regulators and response regulators (Table 3). Such a reduction of regulatory genes has previously been detected in a range of animal symbionts and can point towards static transcriptional dynamics which, as in the case of the pea aphid symbiont *Buchnera aphidicola*, could benefit the host (Shigenobu and Wilson, 2011). Further, HS1 was enriched in genes associated with toxin-antitoxin and restriction-modification (R-M) system, features previously found in other sponge symbionts and indicating an enhanced defence mechanism against foreign DNA (Thomas *et al.*, 2010). Additionally, they could act as stabilisers for mobile genetic elements and thus allow exchange of DNA between sponge symbionts sharing similar R-M systems (Fan *et al.*, 2012). Another symbiotic feature detected in HS1 was the high abundance of eukaryotic-like proteins in form of tetratricopeptide repeats and ankyrin repeats. These eukaryotic-like proteins are thought to be involved in host-symbiont interactions and could for instance prevent the symbiont from being targeted by the sponge immune system (Gao *et al.*, 2014; Nguyen *et al.*, 2014).

4.2.2 Putative role of "*Candidatus* Halichondribacter symbioticus" in *H. panicea*

Transcriptomic analysis of "*Ca*. H. symbioticus" showed a high relative expression of stress response genes, possibly linked to stressors affecting the sponge during sample collection and storage and could give valuable indications as to how the symbiont responds to external stress to maintain its presence in its host. Apart from these genes, high relative expression was also observed for amino acid transport systems with highest similarity to branched-chain amino acids transporters, general L-amino acid transporters and glycine betaine-binding periplasmic proteins, as well as carbohydrate transport systems associated with an ABC-type

sugar transport system and TRAP-type mannitol/chloroaromatic compound transport system, possibly related to the exchange of amino acids and carbohydrates between host and symbiont.

Despite possible exchange of carbohydrates between host and symbiont, HS1 showed a reduced potential to utilize certain sugars such as malate, succinate and fumarate and in general lacked other carbohydrate transport systems found in the free-living reference genomes and in symbiont genome bin131 (Table 3). In additional, lack of expression of carbon catabolite repression genes further point to a limited metabolic versatility of HS1. Most amino acid biosynthesis pathways which were detected in the reference genomes were also reconstructed in HS1 except for the tyrosine and phenylalanine biosynthesis pathways which lacked the *tyrB* gene, and serine biosynthesis which was lacking the *serA*, *serC* and *serB* genes. Ammonia assimilation was present in HS1 through the GS-GOGAT pathway, yet the gene *amtB*, associated with ammonia transport under limited ammonia concentrations, was not expressed, suggesting that HS1 is entirely dependent on ambient ammonia concentrations and its unspecific diffusion across the cell membrane.

Genes involved in vitamin B_{12} production were enriched in HS1 compared to the reference genomes pointing towards a potential beneficial role of the symbiont to the host for which vitamin B_{12} is an essential cofactor. A search for secondary metabolite gene clusters through antiSMASH showed that HS1 contained a bacteriocin gene cluster, suggesting involvement in preventing foreign bacteria from entering and proliferating in the sponge body. Although bacteriocins are antibacterial agents, they often only display inhibitory activity against other, closely related bacteria (Jack *et al.*, 1995). Therefor general cytotoxic activity previously found in *H. panicea* extracts (Althoff *et al.*, 1998; Ferreira *et al.*, 2011) is unlikely to be attributed to "*Ca*. H. symbioticus".

The genome analysis of HS1 points toward a mutualistic relationship between host and symbiont, in which the symbiont has a protected habitat and a stable supply of nutrients while the host benefits from the availability of essential nutrients and cofactors, the removal of metabolic waste products, and growth inhibition of foreign or pathogenic bacteria. Many of the traits found in HS1, such as the reduction of regulatory genes, enrichment of toxin-antitoxin systems, R-M systems and eukaryotic-like proteins, appear to be shared between sponge symbiont across host phylogeny and geographical distribution. The discovery of these features in "*Ca*. H. symbioticus" supports the hypothesis of a convergent evolution of sponge symbiont functions.

*Table 3: Number of genes per enriched or depleted COGs in "*Ca*. H. symbioticus" draft genome HS1 and reference genomes of related free-living bacteria and selected sponge symbionts, divided by COG category.*

COG Description	"*Ca*. H. symbioticus" *HS1*	free-living seawater bacteria			sponge symbionts			
		R. HTCC2255	*A. kogurei*	*A. cionae*	bin36	bin65	bin129	bin131
Carbohydrate transport and metabolism								
TRAP-type C4-dicarboxylate transport systems	11	76	27	56	55	60	130	25
ABC-type transport systems	34	28	39	46	224	3	7	79
Coenzyme transport and metabolism								
Cobalamin biosynthesis	10	3	3	2	3	3	3	3
Transcription								
Transcriptional regulators	4	44	53	130	43	39	29	10
Predicted transcriptional regulators	2	9	18	21	7	10	11	3
Response regulator	0	2	4	8	2	0	2	0
Superfamily II DNA/RNA helicases, SNF2 family	1	0	0	0	6	0	1	6
Replication, recombination and repair								
DNA methylase	55	1	1	1	26	4	4	45
Superfamily I DNA and RNA helicases	6	1	1	1	6	1	1	5
Predicted nuclease of the RecB family	2	0	0	0	1	0	0	0
Defence mechanisms								
Type II restriction enzyme, methylase subunits	3	0	0	0	1	0	0	0
Restriction endonuclease	1	0	0	0	1	0	1	2
Type I restriction-modification system	2	0	2	1	7	1	1	8
General function prediction only								
TPR repeats	17	2	2	2	1	5	5	12
Ankyrin repeat	9	0	0	0	2	1	0	4
Predicted helicase	5	0	0	1	3	0	1	10
Function Unknown								
Toxin-antitoxin systems	14	0	2	1	20	8	7	18

4.3 Sponge-bacteria co-cultivation and strain characterisation (Paper III, IV, V and VI)

4.3.1 Bacterial diversity of *H. panicea* during *ex situ* cultivation and during sponge-bacterial co-cultivation

In **Paper III** we evaluate suitable conditions for *ex situ* cultivation of *H. panicea* by testing five different cultivation set-ups. Maintenance of sponge individuals and sponge explants, measured through active pumping, attachment to substrate and survival, was successful in four out of the five set-ups. The only cultivation set-up that did not succeed in maintaining sponge explants was the method that used a fully recirculating system. Although feed in form of live microalgae was added to the recirculating system, it is possible that the sponges did not receive sufficient nutrients and thus did not survive. Keeping sponges under completely controlled conditions has remained difficult (Osinga *et al.*, 1999), however successful maintenance of sponges in large mesocosm systems with limited water exchange have previously been shown (Webster *et al.*, 2011). Maintaining *H. panicea* in small aquaria under fully controlled conditions will require further research into the nutrient requirements and suitable environmental conditions, but if successful could yield benefits for laboratory-scale experiments.

Maintenance of *H. panicea* under *ex situ* conditions in flow-through and semi-recirculating systems has been shown previously (Barthel and Theede, 1986; Thomassen and Riisgård, 1995), however, the microbial diversity of *H. panicea* during cultivation has not yet been studied. In this study we analysed the microbial diversity in *H. panicea* before and after cultivation in the different tested set-ups. Although the sponges survived over the course of several week in all semi-recirculating and flow-through systems, the microbial community of *H. panicea* changed over time. Over half of the bacterial taxa found in the original sponges used for the experiment were lost after cultivation. In addition, the dominant symbiont "*Ca*. H. symbioticus" was reduced to less than 10 % of the relative abundance in all set-ups apart from the method that used the effluent water from a public display aquarium. In this method "*Ca*. H. symbioticus" remained at an average of 85 % of the relative abundance after 18 weeks of cultivation. As the effluent water from the public aquarium likely carried feed particles which were replaced rapidly through the high water exchange rate, this suggests that the abundance of "*Ca*. H. symbioticus" in its host body is linked to the presence of available nutrients, either in the form of waste products from the sponge or dissolved nutrients directly usable by the symbiont. Although this cultivation set-up worked best for maintaining both the sponge and its main symbiont, future experiment should determine the exact reason for this success. If it is not based on the water source, but rather the high water exchange rate or explants size relative to the tank size, this cultivation set-up could readily be set-up in a laboratory with access to unfiltered seawater.

Based on the knowledge from the sponge cultivation experiment we choose the set-up with the high water exchange from the public aquarium to conduct a sponge-bacteria co-cultivation experiment. Survival of the sponge explants in the co-cultivation device was high, with 41 out of 46 explants surviving for as long as 10 weeks after start of the experiment. After three weeks in the co-cultivation set-up the relative abundance of "*Ca*. H. symbioticus" had dropped from approximately 53 % to 22 % in the analysed explant, whereas the explants analysed after five and ten weeks, resumed a high abundance of 83 %

and 84 % respectively. A prerequisite for co-cultivation is that the host is able maintain its associated bacteria. Maintaining sponge-associated, non-phototrophic bacteria under *ex situ* conditions has been difficult in the past (Webster and Taylor, 2012). Here we show a method which allows for the maintenance of both the sponge and its dominant symbiont under *ex situ* conditions.

4.3.2 Comparison between co-cultivation enrichment of bacteria and standard solid media isolation

The bacterial enrichment from the co-cultivation experiment contained 61 different OTUs assigned to nine different taxonomic classes (Figure 12). Using the standard plating method, bacteria isolated were assigned to only six classes. Although studies on the diversity of the cultivable bacteria from *H. panicea* have previously been conducted (Imhoff and Stöhr, 2003; Wichels *et al.*, 2006; Schneemann, Nagel, *et al.*, 2010), this is the first time that bacteria from the classes *Spirochaetia*, *Fusobacteriia*, *Delta-* and *Epsilonproteobacteria*, and *Clostridia* were enriched. In addition, the co-cultivation enrichment contained reads matching the partial 16S rRNA sequence of "*Ca*. H. symbioticus". The relative abundance of "*Ca*. H. symbioticus" was however low at 0.03 % of all reads and it is not excludable that this might represent left over DNA from the original inoculation. Further experiments, for instance without explants in the co-cultivation chamber would need to be conducted to evaluate if the detection of "*Ca*. H. symbioticus" was due to co-cultivation success or an artefact from the experimental set-up.

Three strains isolated through standard plating techniques which showed lowest sequence similarity from previously cultivated and described species were selected for genome sequencing and antimicrobial activity assays. Draft genome sequences of strain Hp12, Hp32 and Hp36 were constructed and analysed for secondary metabolite gene clusters on AntiSMASH. Only Strain Hp32 contained positive matches with highest similarity to Siderophone, Terpene, T3pks, Nrps and Lantipeptide-Ladderane producing gene clusters. Aqueous and organic extracts of all three strains were produced and tested for growth inhibition against the strains *Escherichia coli*, *Staphylococcus aureus* and *Pseudomonas aeruginosa*. None of the extracts reduced the growth rate of the test strains relative to the positive control without the extracts, pointing towards a lack of antimicrobial activity of the isolates under the tested conditions.

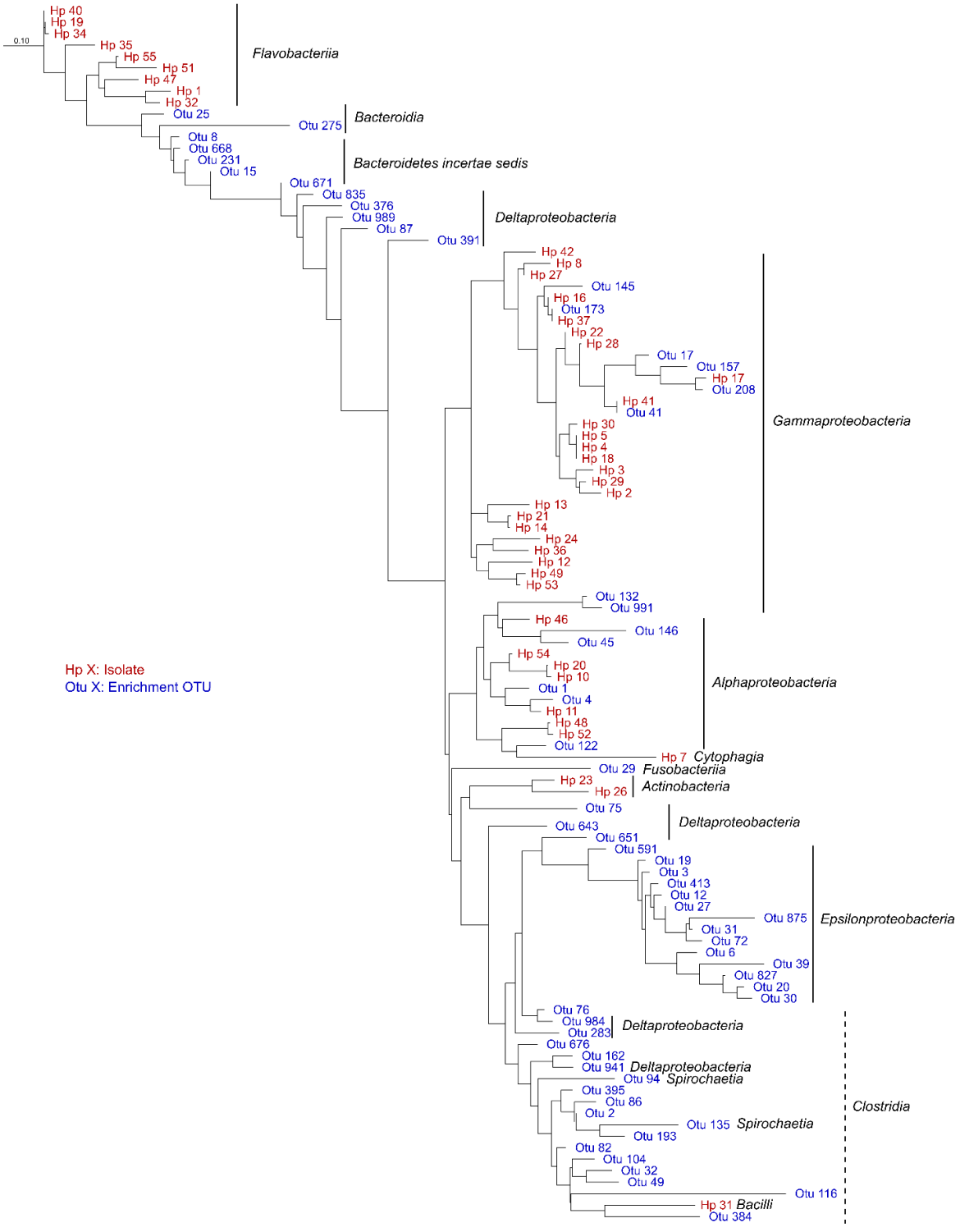


Figure 12: Maximum-likelihood phylogenetic tree of all isolated bacterial strains and enriched OTUs. The tree is based on the V3-V4 region of the 16S rRNA gene. Class level taxonomic assignment is marked to the right of the tree. See Paper III for taxonomic classification of strains and enriched OTUs.

4.3.3 Characterisation of strains Hp12, Hp32 and Hp36

Due to the sheer diversity of microorganisms in sponges, it is often difficult to determine if a bacterial taxon detected by means of molecular methods or traditional cultivation can in

fact be considered sponge-associated or if it is a transient seawater bacterium. For this reason, we decided to analyse three isolated strains, retrieved from *H. panicea* by standard plating method on MA, and see if they could be considered as either sponge-specific or as transient seawater bacteria. Comparison of the 16S rRNA gene sequences to the microbial diversity of *H. panicea* samples from all locations revealed that strains Hp12 and Hp32 were present only in one sponge sample and absent in all other samples (Table 4). Even in the sample used for inoculation the relative abundance of both strains was lower than 0.02 %. Strain Hp36, on the other hand, was detected in approximately 57 % of all sponge samples collected in this study at an average relative abundance of 0.35 %. Although strain Hp36 was also detected in a third of the seawater samples collected in this study, its average relative abundance was lower than in the sponge samples at 0.14 % (Table 4). Searched against the SRA database, the 16S rRNA gene sequence of Hp12, Hp32 and Hp36 were detected in 56, 7 and 1 seawater metagenomic datasets respectively, with 99 % or high sequence similarity.

Table 4: Strain description of strains Hp12, Hp32 and Hp36 isolated from H. panicea *using standard plating methods. Detection (= 99% sequence similarity) in sponge and seawater 16S rRNA gene sequence dataset from this study, SRA database and sponge-EMP dataset.*

Strain:	**Hp12**	**Hp32**	**Hp36**
GenBank acc. 16S rRNA gene	MG877746	MK633875	MK633876
GenBank acc. whole genome project	QDDL00000000	SIHP00000000	SIHQ00000000
Proposed name	*Pelagibaculum spongiae* gen. nov., sp. nov.	*Tenacibaculum islandicus* sp. nov.	*Endozoicomonas halichondricola* sp. nov.
Habitat	Seawater-associated	Seawater-associated	Host-associated
Detected in *H. panicea* samples (this study)	ca. 1%	ca. 1%	ca. 57%
Average relative abundance in *H. panicea* (this study)	0.01%	0.01%	0.35%
Detected in seawater samples (this study)	ca. 5%	ca. 33%	ca. 33%
Average relative abundance in seawater (this study)	0.11%	0.17%	0.14%
Number of sponge samples in sponge-EMP dataset* with strain detection	36	18	0
Number of seawater / sponge metagenomes in SRA† with strain detection	56 / 0	7 / 0	1 / 0

* Moitinho-Silva, Nielsen, *et al.*, 2017
† https://www.ncbi.nlm.nih.gov/sra **(**Leinonen *et al.*, 2011**),** accessed Dec. 2017 through IMNGS (Lagkouvardos *et al.*, 2016)

The relative abundance of strains Hp12 and Hp32 was on average 0.005 % and 0.014 % in these bacterial communities respectively. This suggests that strains Hp12 and Hp32 are not sponge-specific bacteria, but rather low abundant seawater bacteria which were

coincidentally found inside the sponge body through pumping activity or being attached to the surface of the sponge. Strain Hp36, on the other hand, was detected in over half of the *H. panicea* samples collected in this study and could thus be considered a putative sponge-associated bacterium. However, a lack of detection in sponge datasets stored on SRA (Leinonen *et al.*, 2011), as part of the sponge microbiome project (Moitinho-Silva, Nielsen, *et al.*, 2017), or in previously described sponge-specific sequence clusters (Simister *et al.*, 2012), suggests that strain Hp36 is not a general sponge-specific bacterium. 16S rRNA sequence comparison to closely related characterised strains on EzBioCloud (Yoon, S.-M. Ha, *et al.*, 2017) showed that strain Hp36 belonged to the genus *Endozoicomonas*, members of which are symbionts to a range of marine animals (Neave *et al.*, 2016). We therefore conclude that strain Hp36, similar to other members of its genus, is a marine animal-associated bacterium, which, due to its partial presence in sponge samples collected in this study, exhibits an opportunistic lifestyle in *H. panicea* from Icelandic waters. Future studies need to determine if Hp36 also participates in host-symbiont mutualistic behaviour, such as host-associated protein and carbohydrate transport and cycling, as suggested for other *Endozoicomonas* spp. (Neave *et al.*, 2016).

In addition to the characterisation of their host-association, the novel position of strains Hp12 and Hp32 within the class *Gammaproteobacteria* and strain Hp32 in the class *Flavobacteria* warranted further investigation into their phylogenetic placement and characterisation which are presented in **Paper IV**, **V** and **VI**.

The 16S rRNA gene sequence similarity of strain Hp12, dDDH and ANI values of its genome, and phylogenetic placement relative to its cultivated and officially described relatives strongly suggested that strain Hp12 belongs to a new species in a new genus. Its closest relatives in the phylogenetic tree were members of the genus *Alcanivorax* in the family *Oceanospirillaceae*. Other phenotypical differences such as cell size, maximum salinity tolerance, temperature growth range, catalase activity, DNA G+C content and fatty acid concentrations further differentiated strain Hp12 from members of the genus *Alcanivorax* and of other related genera in the order *Oceanospirillales*. We proposed the name *Pelagibaculum spongiae* for which Hp12 is the type strain and deposited it in two culture collections under the strain designations DSM 104963^{T} and CECT 9367^{T}. Official characterisation of novel strains and their taxonomic placement is an important step in resolving microbial taxonomy (Yarza *et al.*, 2008) and as a future reference for other new strain discoveries. Especially in the class *Gammaproteobacteria* the phylogeny has been difficult to resolve (Williams *et al.*, 2010). In this context we provide a reference strain for future phylogenetic work in this large bacterial class.

Polyphasic analysis showed that strain Hp32 belonged to the genus *Tenacibaculum* of the class *Flavobacteria*, but several phenotypical traits and a 16S rRNA gene sequence similarity below 97.2 % to its closest cultivated relatives differentiated Hp36 from other type strains with validly published names. We therefore propose a new species with the name *Tenacibaculum islandicus*. Although Hp32 possesses secondary metabolite biosynthetic gene clusters, antimicrobial activity assays did not inhibit the growth of three bacterial test strains. Further examination of other biotechnologically relevant bioactive properties such as anti-tumour, anti-fungal or anti-parasitic activity could be conducted in future studies.

Strain Hp36 showed highest 16S rRNA gene sequence similarity to members of the genus *Endozoicomonas* in the class *Gammaproteobacteria* and phylogenetic analysis confirmed the placement of strain Hp36 in the genus *Endozoicomonas*. Low ANI values and dDDH

estimates as well as differential phenotypical traits between Hp36 and other related type strains clearly showed that strain Hp36 presents a novel species for which we propose the name *Endozoicomonas halichondricola*. All validly published members of the genus *Endozoicomonas* have been isolated from benthic marine animals and the two closest related *Endozoicomonas* spp. to Hp36 were isolated from marine sponges. Further investigation into the similarities between these putative sponge-associated species could yield insights into shared genomic features and provide valuable information into the host-symbiotic interactions of *Endozoicomonas* spp. associated with sponges.

5 Conclusions and Future Perspectives

This thesis aimed at broadening our understanding of the symbiotic relationship between *Halichondria panicea*, one of the most studied sponge species to date, and its bacterial community. By using high-throughput sequencing methods it is possible to detect a large diversity of microorganisms in marine sponges, yet the task of differentiating between sponge-specific, opportunistic and transient microbes is far from resolved for most sponge species (Taylor *et al.*, 2007; Schmitt, Tsai, *et al.*, 2012). By applying a high-throughput sequencing approach to *H. panicea* from different geographical locations and across different points in time we could show that most bacteria detected in *H. panicea* from Icelandic waters do not belong to a stable sponge species specific core-group of bacteria. In fact, this core-group consisted of a single bacterial taxon which had previously been detected in *H. panicea* from various locations in the North Atlantic. Although we show that other bacterial taxa detected in *H. panicea* do not belong to this core-group, it was apparent that they were at least temporarily specific to the sponge ecosystem, shown through their dissimilarity to the surrounding seawater bacterial community. Future studies must investigate to which degree these bacteria are opportunists that inhabit, for instance, the surface of sponges and possibly other invertebrate, or if they play a significant role inside their host. The finding that, in two cases, the bacterial community of *H. panicea* was closely related to that of its surrounding seawater, indicates that the sponge is capable of, or subject to, a periodic removal of its bacterial community which could play an important role in determining the nature of its temporary inhabitants.

Close animals-microbe relationships can be characterised by coevolutionary processes in which both or multiple participants exert selection pressure on each other (Easson and Thacker, 2014). The discovery of geographically distinct subpopulations of "*Ca*. H. symbioticus" and reports of sponges identified as *H. panicea* that do not contain this symbiont, point towards a sponge species complex with possible divergent lineages based on the presence or absence of "*Ca*. H. symbioticus". Further studies would need to address this hypothesis through analysis of sponge and symbiont marker genes from all locations that report the presence of *H. panicea* or *H. panicea*-like species. Next, changes at the sponge genome level related to the presence of "*Ca*. H. symbioticus", for instance in form of shared metabolic functions, would need to be evaluated to determine coevolutionary processes. In a similar manner, genomic comparison of "*Ca*. H. symbioticus" found in *H. panicea* and "*Ca*. H. symbioticus" found in other sponge species could be analysed for host species specific traits.

We present the first draft genome of "*Ca*. H. symbioticus" providing insights into the functional genomic repertoire of this sponge symbiont and its putative role inside of its host. Shared features between "*Ca*. H. symbioticus" and other previously described sponge symbionts warrant further investigation into potential mechanisms that are involved in establishing these convergent features, for instance horizontal gene transfer, or insertion of mobile genetic elements (Fan *et al.*, 2012). Moreover, the relative simplicity of the *H. panicea* core microbial community, consisting of a dominant bacterial taxon, could facilitate

future hypothesis-driven, experimental research targeting gene expression. Cultivation of *H. panicea* under semi-controlled conditions could enable the manipulation of single environmental conditions, as for instance temperature, the availability of different nutrient sources, or the introduction of pathogenic microorganisms, followed by an analysis of the gene expression response by "*Ca*. H. symbioticus". Availability of the *H. panicea* genome would further provide greater understanding on the interrelation between host and symbiont under different experimental conditions and provide valuable insights into sponge-microbe symbiosis in general.

New methods are needed to exploit the biotechnological potential of sponge-associated microbes and produce their bioactive compounds at a scale to overcome the "supply problem". One method is to enhance the cultivability of sponge-associated microbes (Steinert *et al.*, 2018). In light of "*Ca*. H. symbioticus" genomic features, especially the lack of regulatory genes which could bind metabolic functions and even reproductive mechanisms to its host, it is unlikely that this symbiont could be isolated using standard cultivation methods even with complex media compositions targeting specific metabolic pathways. Here we present a new method of growing microorganisms outside of, but in close proximity to, their host through the use of a co-cultivation chamber. We were able to demonstrate that *ex situ* cultivation conditions were suitable to maintain "*Ca*. H. symbioticus" inside the sponge body, a prerequisite for co-cultivation outside of the host. Whereas we were able to provide this proof of concept, larger scale experiments are needed to explore the full potential of this method.

With an increasing awareness of animal-associated microbiota and their role in animal physiology beyond pathogenicity, not least in the context of the human microbiota, it is appropriate to elucidate the microbial interactions of the most basal members of the animal kingdom and reveal coevolutionary processes in the animal lineage through conserved mechanisms. The results presented in this study form a good basis for future studies focusing not only on host-microbe interactions in the sponge *H. panicea*, but also on other sponge-symbiont relationships.

References

Abdelmohsen, U.R., Pimentel-Elardo, S.M., Hanora, A., Radwan, M., Abou-El-Ela, S.H., Ahmed, S., et al. (2010) Isolation, Phylogenetic Analysis and Anti-infective Activity Screening of Marine Sponge-Associated Actinomycetes. *Mar. Drugs* **8**: 399–412.

Abdjul, D.B., Yamazaki, H., Kanno, S., Takahashi, O., Kirikoshi, R., Ukai, K., and Namikoshi, M. (2016) Haliclonadiamine Derivatives and 6- *epi* -Monanchorin from the Marine Sponge *Halichondria panicea* Collected at Iriomote Island. *J. Nat. Prod.* **79**: 1149–1154.

Ackers, R.G., Moss, D., and Picton, B.E. (2007) Sponges of the British Isles (" Sponge V"): a colour guide and working document, 1992 edition, reset with modifications Marine Conservation Society.

Albertsen, M., Hugenholtz, P., Skarshewski, A., Nielsen, K.L., Tyson, G.W., and Nielsen, P.H. (2013) Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nat Biotech* **31**: 533–538.

Alegado, R.A., Brown, L.W., Cao, S., Dermenjian, R.K., Zuzow, R., Fairclough, S.R., et al. (2012) A bacterial sulfonolipid triggers multicellular development in the closest living relatives of animals. *Elife* **1**: e00013.

Althoff, K., Schütt, C., Steffen, R., Batel, R., and Müller, W.E.G. (1998) Evidence for a symbiosis between bacteria of the genus *Rhodobacter* and the marine sponge *Halichondria panicea*: Harbor also for putatively toxic bacteria? *Mar. Biol.* **130**: 529–536.

Amann, J. (1911) Die direkte Zahlung der Wasserbakterien mittels des Ultramikroskops. *Cent. f. Bakteriol.* **29**: 381–384.

Amann, R.I., Binder, B.J., Olson, R.J., Chisholm, S.W., Devereux, R., and Stahl, D.A. (1990) Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl. Environ. Microbiol.* **56**: 1919–1925.

Amann, R.I., Ludwig, W., and Schleifer, K.-H. (1995) Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* **59**: 143–169.

Amano, S. (1986) Larval release in response to a light signal by the intertidal sponge Halichondria panicea. *Biol. Bull.* **171**: 371–378.

Apprill, A., McNally, S., Parsons, R., and Weber, L. (2015) Minor revision to V4 region SSU rRNA 806R gene primer greatly increases detection of SAR11 bacterioplankton. *Aquat. Microb. Ecol.* **75**: 129–137.

Ayling, A.L. (1980) Patterns of sexuality, asexual reproduction and recruitment in some subtidal marine Demospongiae. *Biol. Bull.* **158**: 271–282.

Bäckhed, F., Ding, H., Wang, T., Hooper, L. V, Koh, G.Y., Nagy, A., et al. (2004) The gut microbiota as an environmental factor that regulates fat storage. *Proc. Natl. Acad. Sci. U. S. A.* **101**: 15718–23.

Baker, G.C., Smith, J.J., and Cowan, D.A. (2003) Review and re-analysis of domain-specific 16S primers. *J. Microbiol. Methods* **55**: 541–555.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., et al. (2012) SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J. Comput. Biol.* **19**: 455–477.

Barthel, D. (1991) Influence of Different Current Regimes on the Growth Form of

Halichondria panicea Pallas. In, *Fossil and Recent Sponges*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 387–394.
Barthel, D. (1988) On the ecophysiology of the sponge Halichondria panicea in Kiel Bight. 2. Biomass, production, energy budget and integration in environmental processes. *Mar. Ecol. Prog. Ser. Oldend.* **43**: 87–93.
Barthel, D. (1986) On the ecophysiology of the sponge Halichondria panicea in Kiel Bight. I. Substrate specificity, growth and reproduction. *Mar. Ecol. Prog. Ser.* **32**: 291–298.
Barthel, D. and Theede, H. (1986) A new method for the culture of marine sponges and its application for experimental studies. *Ophelia* **25**: 75–82.
Barthel, D. and Wolfrath, B. (1989) Tissue sloughing in the sponge Halichondria panicea: a fouling organism prevents being fouled. *Oecologia* **78**: 357–360.
de Bary, A. (1879) Die erscheinung der symbiose Verlag von Karl J. Trübner.
Beazley, L.I., Kenchington, E.L., Murillo, F.J., and Sacau, M. del M. (2013) Deep-sea sponge grounds enhance diversity and abundance of epibenthic megafauna in the Northwest Atlantic. *ICES J. Mar. Sci.* **70**: 1471–1490.
Bell, J.J. (2008) The functional roles of marine sponges. *Estuar. Coast. Shelf Sci.* **79**: 341–353.
Bergquist, P.R. (1961) A collection of Porifera from Northern New Zealand, with descriptions of seventeen new species. *Pacific Sci.* **15**: 33–48.
Bergquist, P.R. (1978) Sponges Univ of California Press.
Blin, K., Wolf, T., Chevrette, M.G., Lu, X., Schwalen, C.J., Kautsar, S.A., et al. (2017) antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* **45**: W36–W41.
Bolger, A.M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.
Bordenstein, S.R. (2016) Lean on We: How Eco-Evolution Permeates Animal–Microbe Associations. *Trends Ecol. Evol.* **31**: 815–816.
Bourne, D.G., Garren, M., Work, T.M., Rosenberg, E., Smith, G.W., and Harvell, C.D. (2009) Microbial disease and the coral holobiont. *Trends Microbiol.* **17**: 554–562.
Brümmer, F. and Nickel, M. (2003) Sustainable Use of Marine Resources: Cultivation of Sponges. Springer, Berlin, Heidelberg, pp. 143–162.
Burgsdorf, I., Slaby, B.M., Handley, K.M., Haber, M., Blom, J., Marshall, C.W., et al. (2015) Lifestyle evolution in cyanobacterial symbionts of sponges. *MBio* **6**: e00391-15.
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.
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 Meth* **13**: 581–583.
Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009) BLAST+: architecture and applications. *BMC Bioinformatics* **10**: 421.
Camp, J.G., Kanther, M., Semova, I., and Rawls, J.F. (2009) Patterns and scales in gastrointestinal microbial ecology. *Gastroenterology* **136**: 1989–2002.
Cani, P.D., Bibiloni, R., Knauf, C., Waget, A., Neyrinck, A.M., Delzenne, N.M., and Burcelin, R. (2008) Changes in gut microbiota control metabolic endotoxemia-induced inflammation in high-fat diet–induced obesity and diabetes in mice. *Diabetes* **57**: 1470–1481.
Caporaso, J.G., Lauber, C.L., Walters, W.A., Berg-Lyons, D., Lozupone, C.A., Turnbaugh, P.J., 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 Suppl 1**: 4516–22.
Casapullo, A., Minale, L., and Zollo, F. (1993) Paniceins and Related Sesquiterpenoids from the Mediterranean Sponge Reniera fulva. *J. Nat. Prod.* **56**: 527–533.
Chambers, S. (2003) Systema porifera. A guide to the classification of sponges,edited by J.N.A. Hooper and R.W.M. van Soest. Kluwer Academic/Plenum Publishers, Dordrecht, 2002, 1756pp. ISBN 0-306-47260-0. *Aquat. Conserv. Mar. Freshw. Ecosyst.* **13**: 461–462.
Cheesman, S.E., Neal, J.T., Mittge, E., Seredick, B.M., and Guillemin, K. (2011) Epithelial cell proliferation in the developing zebrafish intestine is regulated by the Wnt pathway and microbial signaling via Myd88. *Proc. Natl. Acad. Sci.* **108**: 4570–4577.
Christensen, T. (1985) Microspora ficulinae, a green alga living in marine sponges. *Br. Phycol. J.* **20**: 5–7.
Christophersen, C., Anthoni, U., Nielsen, P.H., Jacobsen, N., and Tendal, O.S. (1989) Source of a nauseating stench from the marine sponge, Halichondria panicea, collected at Clever Bank in the North Sea. *Biochem. Syst. Ecol.* **17**: 459–461.
Chu, H. and Mazmanian, S.K. (2013) Innate immune recognition of the microbiota promotes host-microbial symbiosis. *Nat. Immunol.* **14**: 668–675.
Chun, C.K., Ozer, E.A., Welsh, M.J., Zabner, J., and Greenberg, E.P. (2004) Inactivation of a Pseudomonas aeruginosa quorum-sensing signal by human airway epithelia. *Proc. Natl. Acad. Sci. U. S. A.* **101**: 3587–90.
Cimino, G., De Stefano, S., and Minale, L. (1973) Paniceins, unusual aromatic sesquiterpenoids linked to a quinol or quinone system from the marine sponge Halichondria panicea. *Tetrahedron* **29**: 2565–2570.
Cohen, L.J., Esterhazy, D., Kim, S.-H., Lemetre, C., Aguilar, R.R., Gordon, E.A., et al. (2017) Commensal bacteria make GPCR ligands that mimic human signalling molecules. *Nature* **549**: 48–53.
Collins, S.M., Surette, M., and Bercik, P. (2012) The interplay between the intestinal microbiota and the brain. *Nat. Rev. Microbiol.* **10**: 735.
Costello, E.K., Stagaman, K., Dethlefsen, L., Bohannan, B.J.M., and Relman, D.A. (2012) The application of ecological theory toward an understanding of the human microbiome. *Science (80-.).* **336**: 1255–1262.
Croué, J., West, N.J., Escande, M.-L., Intertaglia, L., Lebaron, P., and Suzuki, M.T. (2013) A single betaproteobacterium dominates the microbial community of the crambescidine-containing sponge Crambe crambe. *Sci. Rep.* **3**: 2583.
Csuos, M. (2010) Count: evolutionary analysis of phylogenetic profiles with parsimony and likelihood. *Bioinformatics* **26**: 1910–1912.
Csűrös, M. (2008) Ancestral Reconstruction by Asymmetric Wagner Parsimony over Continuous Characters and Squared Parsimony over Distributions. Springer, Berlin, Heidelberg, pp. 72–86.
Daims, H., Brühl, A., Amann, R., Schleifer, K.-H., and Wagner, M. (1999) The Domain-specific Probe EUB338 is Insufficient for the Detection of all Bacteria: Development and Evaluation of a more Comprehensive Probe Set. *Syst. Appl. Microbiol.* **22**: 434–444.
Daims, H., Lücker, S., and Wagner, M. (2006) daime, a novel image analysis program for microbial ecology and biofilm research. *Environ. Microbiol.* **8**: 200–213.
Díez-Vives, C., Moitinho-Silva, L., Nielsen, S., Reynolds, D., and Thomas, T. (2017) Expression of eukaryotic-like protein in the microbiome of sponges. *Mol. Ecol.* **26**: 1432–1451.
Díez, B., Pedrós-Alió, C., and Massana, R. (2001) Study of genetic diversity of eukaryotic

picoplankton in different oceanic regions by small-subunit rRNA gene cloning and sequencing. *Appl. Environ. Microbiol.* **67**: 2932–41.
Domazet-Lošo, T. and Tautz, D. (2008) An ancient evolutionary origin of genes associated with human genetic diseases. *Mol. Biol. Evol.* **25**: 2699–2707.
Dominguez-Bello, M.G., Costello, E.K., Contreras, M., Magris, M., Hidalgo, G., Fierer, N., and Knight, R. (2010) Delivery mode shapes the acquisition and structure of the initial microbiota across multiple body habitats in newborns. *Proc. Natl. Acad. Sci.* **107**: 11971–11975.
Douglas, A. and Smith, D.C. (1984) The green hydra symbiosis. VIII. Mechanisms in symbiont regulation. *Proc. R. Soc. London. Ser. B. Biol. Sci.* **221**: 291–319.
Douglas, A.E. and Prosser, W.A. (1992) Synthesis of the essential amino acid tryptophan in the pea aphid (Acyrthosiphon pisum) symbiosis. *J. Insect Physiol.* **38**: 565–568.
Douglas, A.E. and Werren, J.H. (2016) Holes in the Hologenome: Why Host-Microbe Symbioses Are Not Holobionts. *MBio* **7**: e02099.
Dubilier, N., Bergin, C., and Lott, C. (2008) Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. *Nat. Rev. Microbiol.* **6**: 725–740.
Duckworth, A. (2009) Farming sponges to supply bioactive metabolites and bath sponges: A review. *Mar. Biotechnol.* **11**: 669–679.
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.
Eberl, G. (2010) A new vision of immunity: homeostasis of the superorganism. *Mucosal Immunol.* **3**: 450.
Eddy, S.R. (2011) Accelerated Profile HMM Searches. *PLoS Comput. Biol.* **7**: e1002195.
Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**: 1792–1797.
Edgar, R.C. (2016) SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. *bioRxiv* 074161.
Edgar, R.C. (2013) UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Meth* **10**: 996–998.
Engelstädter, J. and Hurst, G.D.D. (2009) The ecology and evolution of microbes that manipulate host reproduction. *Annu. Rev. Ecol. Evol. Syst.* **40**: 127–149.
Erpenbeck, D., Breeuwer, J., van der Velde, H., and van Soest, R. (2002) Unravelling host and symbiont phylogenies of halichondrid sponges (Demospongiae, Porifera) using a mitochondrial marker. *Mar. Biol.* **141**: 377–386.
Erpenbeck, D., Knowlton, A.L., Talbot, S.L., Highsmith, R.C., and Van Soest, R.W. (2004) A molecular comparison of Alaskan and Northeast Atlantic Halichondria panicea (Pallas 1766)(Porifera: Demospongiae) populations. *Boll Mus Ist Univ Genova* **61**: 319–325.
Erwin, P.M. and Thacker, R.W. (2007) Incidence and identity of photosynthetic symbionts in Caribbean coral reef sponge assemblages. *J. Mar. Biol. Assoc. UK* **87**: 1683–1692.
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. U. S. A.* **109**: E1878-87.
Farris, J.S. (1977) Phylogenetic Analysis Under Dollo's Law. *Syst. Biol.* **26**: 77–88.
Feldmann, J. (1933) Sur quelques cyanophycées vivant dans le tissu des éponges de banyules. *Arch Zool Exp Gén* **75**: 331–404.
Felsenstein, J. (1985) Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution (N. Y).* **39**: 783–791.
Felsenstein, J. (2004) PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by

the author. *http://www.evolution.gs.washington.edu/phylip.html*.
Ferreira, M., Cabado, A.G., Chapela, M.-J., Fajardo, P., Atanassova, M., Garrido, A., et al. (2011) Cytotoxic activity of extracts of marine sponges from NW Spain on a neuroblastoma cell line. *Environ. Toxicol. Pharmacol.* **32**: 430–437.
Feuda, R., Dohrmann, M., Pett, W., Philippe, H., Rota-Stabelli, O., Lartillot, N., et al. (2017) Improved Modeling of Compositional Heterogeneity Supports Sponges as Sister to All Other Animals. *Curr. Biol.* **27**: 3864–3870.e4.
Flatt, P.M., Gautschi, J.T., Thacker, R.W., Musafija-Girt, M., Crews, P., and Gerwick, W.H. (2005) Identification of the cellular site of polychlorinated peptide biosynthesis in the marine sponge Dysidea (Lamellodysidea) herbacea and symbiotic cyanobacterium Oscillatoria spongeliae by CARD-FISH analysis. *Mar. Biol.* **147**: 761–774.
Forester, A.J. (1979) The association between the sponge Halichondria panicea (Pallas) and scallop Chlamys varia (L.): A commensal-protective mutualism. *J. Exp. Mar. Bio. Ecol.* **36**: 1–10.
Forst, S. and Clarke, D. (2002) Bacteria-nematode symbiosis. *Entomopathog. Nematol.* 57–77.
Freeman, C.J. and Thacker, R.W. (2011) Complex interactions between marine sponges and their symbiotic microbial communities. *Limnol. Oceanogr.* **56**: 1577–1586.
Frøhlich, H. and Barthel, D. (1997) Silica uptake of the marine sponge Halichondria panicea in Kiel Bight. *Mar. Biol.* **128**: 115–125.
Galperin, M.Y., Makarova, K.S., Wolf, Y.I., and Koonin, E. V. (2015) Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res.* **43**: D261–D269.
Gao, Z.-M., Wang, Y., Tian, R.-M., Wong, Y.H., Batang, Z.B., Al-Suwailem, A.M., et al. (2014) Symbiotic adaptation drives genome streamlining of the cyanobacterial sponge symbiont "Candidatus Synechococcus spongiarum" *MBio* **5**: e00079-14.
Garrett, W.S., Gordon, J.I., and Glimcher, L.H. (2010) Homeostasis and inflammation in the intestine. *Cell* **140**: 859–870.
Gauthier, M.-E.A., Watson, J.R., and Degnan, S.M. (2016) Draft Genomes Shed Light on the Dual Bacterial Symbiosis that Dominates the Microbiome of the Coral Reef Sponge Amphimedon queenslandica. *Front. Mar. Sci.* **3**: 196.
Gazave, E., Lapébie, P., Ereskovsky, A. V., Vacelet, J., Renard, E., Cárdenas, P., and Borchiellini, C. (2012) No longer Demospongiae: Homoscleromorpha formal nomination as a fourth class of Porifera. *Hydrobiologia* **687**: 3–10.
Gerasimova, E.I. and Ereskovsky, A. V (2007) Reproduction of two species of Halichondria (Demospongiae: Halichondriidae) in the White Sea. *Porifera Res. biodiversity, Innov. Sustain. Livros. Mus. Nac. Rio Janeiro, Proc. 7th Int. Sponge Symp.*
Gilbert, J.A., Blaser, M.J., Caporaso, J.G., Jansson, J.K., Lynch, S. V, and Knight, R. (2018) Current understanding of the human microbiome. *Nat. Med.* **24**: 392–400.
Gilbert, J.A., Jansson, J.K., and Knight, R. (2014) The Earth Microbiome project: successes and aspirations. *BMC Biol.* **12**: 69.
Gilbert, S.F., Sapp, J., and Tauber, A.I. (2012) A Symbiotic View of Life: We Have Never Been Individuals. *Q. Rev. Biol.* **87**: 325–341.
Gloeckner, V., Wehrl, M., Moitinho-Silva, L., Gernert, C., Hentschel, U., Schupp, P., 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 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**: 108–10.

Goldstein, J., Riisgård, H.U., and Larsen, P.S. (2019) Exhalant jet speed of single-osculum explants of the demosponge Halichondria panicea and basic properties of the sponge-pump. *J. Exp. Mar. Bio. Ecol.* **511**: 82–90.

González, J.F. and Venturi, V. (2013) A novel widespread interkingdom signaling circuit. *Trends Plant Sci.* **18**: 167–174.

Guindon, S. and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**: 696–704.

Gutleben, J., Chaib De Mares, M., van Elsas, J.D., Smidt, H., Overmann, J., and Sipkema, D. (2018) The multi-omics promise in context: from sequence to microbial isolate. *Crit. Rev. Microbiol.* **44**: 212–229.

Haine, E.R. (2008) Symbiont-mediated protection. *Proc. R. Soc. B Biol. Sci.* **275**: 353–361.

Hallam, S.J., Mincer, T.J., Schleper, C., Preston, C.M., Roberts, K., Richardson, P.M., and DeLong, E.F. (2006) Pathways of carbon assimilation and ammonia oxidation suggested by environmental genomic analyses of marine Crenarchaeota. *PLoS Biol.* **4**: e95.

Handelsman, J. (2004) Metagenomics: application of genomics to uncultured microorganisms. *Microbiol. Mol. Biol. Rev.* **68**: 669–85.

Handelsman, J., Rondon, M.R., Brady, S.F., Clardy, J., and Goodman, R.M. (1998) Molecular biological access to the chemistry of unknown soil microbes: a new frontier for natural products. *Chem. Biol.* **5**: R245–R249.

Hansen, A.K. and Moran, N.A. (2011) Aphid genome expression reveals host-symbiont cooperation in the production of amino acids. *Proc. Natl. Acad. Sci. U. S. A.* **108**: 2849–54.

Hansen, I. V., Weeks, J.M., and Depledge, M.H. (1995) Accumulation of copper, zinc, cadmium and chromium by the marine sponge Halichondria panicea Pallas and the implications for biomonitoring. *Mar. Pollut. Bull.* **31**: 133–138.

Hausmann, R., Vitello, M.P., Leitermann, F., and Syldatk, C. (2006) Advances in the production of sponge biomass Aplysina aerophoba-A model sponge for ex situ sponge biomass production. *J. Biotechnol.* **124**: 117–127.

Hayashi, K., Hamada, Y., and Shioiri, T. (1992) Synthesis of nazumamide A, a thrombin-inhibitory linear tetrapeptide, from a marine sponge, Theonella sp. *Tetrahedron Lett.* **33**: 5075–5076.

Hentschel, U., Hopke, J., Horn, M., Anja, 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.

Hentschel, U., Usher, K.M., and Taylor, M.W. (2006) Marine sponges as microbial fermenters. *FEMS Microbiol. Ecol.* **55**: 167–177.

Herbst, T., Sichelstiel, A., Schär, C., Yadava, K., Bürki, K., Cahenzli, J., et al. (2011) Dysregulation of Allergic Airway Inflammation in the Absence of Microbial Colonization. *Am. J. Respir. Crit. Care Med.* **184**: 198–205.

Hirata, Y. and Uemura, D. (1986) Halichondrins-antitumor polyether macrolides from a marine sponge. *Pure Appl. Chem.* **58**: 701–710.

Hoffmann, F., Rapp, H.T., and Reitner, J. (2006) Monitoring microbial community composition by fluorescence in situ hybridization during cultivation of the marine cold-water sponge Geodia barretti. *Mar. Biotechnol.* **8**: 373–379.

Hofmann, K. and Baron, M.D. (1996) Boxshade 3.21. In, *Pretty printing and shading of multiple-alignment files. Kay Hofmann ISREC Bioinformatics Group, Lausanne, Switzerland*.

Hongoh, Y. (2010) Diversity and genomes of uncultured microbial symbionts in the termite

gut. *Biosci. Biotechnol. Biochem.* **74**: 1145–1151.
Hooper, L. V., Littman, D.R., and Macpherson, A.J. (2012) Interactions between the microbiota and the immune system. *Science (80-.).* **336**: 1268–1273.
Hooper, L. V. and MacPherson, A.J. (2010) Immune adaptations that maintain homeostasis with the intestinal microbiota. *Nat. Rev. Immunol.* **10**: 159–169.
http://blogs.nature.com (2008) Ocean census reveals the beast with 56 names : News blog. *http://blogs.nature.com/news/2008/06/ocean_census_reveals_the_beast.html, accessed 25.2.2019*.
Huang, Y., Gilna, P., and Li, W. (2009) Identification of ribosomal RNA genes in metagenomic fragments. *Bioinformatics* **25**: 1338–1340.
Hugenholtz, P. and Tyson, G.W. (2008) Metagenomics. *Nature* **455**: 481–483.
Hughes, D.T. and Sperandio, V. (2008) Inter-kingdom signalling: communication between bacteria and their hosts. *Nat. Rev. Microbiol.* **6**: 111.
Hummel, H., Sepers, A., de Wolf, L., and Melissen, F. (1988) Bacterial growth of the marine sponge Halichondria panicea induced by reduced waterflow rate . *Mar. Ecol. Prog. Ser.* **42**: 195–198.
Hyatt, D., LoCascio, P.F., Hauser, L.J., and Uberbacher, E.C. (2012) Gene and translation initiation site prediction in metagenomic sequences. *Bioinformatics* **28**: 2223–2230.
Imhoff, J.F. and Stöhr, R. (2003) Sponge-Associated Bacteria: General Overview and Special Aspects of Bacteria Associated with Halichondria panicea. In, *Sponges (Porifera)*. Springer, Berlin, Heidelberg, pp. 35–57.
Jack, R.W., Tagg, J.R., and Ray, B. (1995) Bacteriocins of gram-positive bacteria. *Microbiol. Rev.* **59**: 171–200.
Johnson, K.V.-A. and Foster, K.R. (2018) Why does the microbiome affect behaviour? *Nat. Rev. Microbiol.* 1.
Kamke, J., Rinke, C., Schwientek, P., Mavromatis, K., Ivanova, N., Sczyrba, 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.
Kamke, J., Sczyrba, 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.
Kanehisa, M. and Goto, S. (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**: 27–30.
Kanehisa, M., Sato, Y., and Morishima, K. (2016) BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *J. Mol. Biol.* **428**: 726–731.
Kaper, J.B. and Sperandio, V. (2005) Bacterial cell-to-cell signaling in the gastrointestinal tract. *Infect. Immun.* **73**: 3197–209.
Karlsson, T., Turkina, M. V., Yakymenko, O., Magnusson, K.-E., and Vikström, E. (2012) The Pseudomonas aeruginosa N-Acylhomoserine Lactone Quorum Sensing Molecules Target IQGAP1 and Modulate Epithelial Cell Migration. *PLoS Pathog.* **8**: e1002953.
Kennedy, J., Baker, P., Piper, C., Cotter, P.D., Walsh, M., Mooij, M.J., et al. (2009) Isolation and analysis of bacteria with antimicrobial activities from the marine sponge haliclona simulans collected from irish waters. *Mar. Biotechnol.* **11**: 384–396.
Keren, R., Mayzel, B., Lavy, A., Polishchuk, I., Levy, D., Fakra, S.C., et al. (2017) Sponge-associated bacteria mineralize arsenic and barium on intracellular vesicles. *Nat. Commun.* **8**: 14393.
Khalaman, V. V., Korchagina, N.M., and Komendantov, A.Y. (2014) The impact of

waterborne cues from conspecifics and other species on the larvae of Halichondria panacea Pallas, 1766 (Porifera: Demospongiae). *Russ. J. Mar. Biol.* **40**: 36–42.
Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., and Glöckner, F.O. (2013) Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41**: e1–e1.
Knoll, A.H. (2015) Life on a Young Planet: The First Three Billion Years of Evolution on Earth-Updated Edition Princeton University Press.
Knowlton, A.L. and Highsmith, R.C. (2005) Nudibranch-sponge feeding dynamics: Benefits of symbiont-containing sponge to Archidoris montereyensis (Cooper, 1862) and recovery of nudibranch feeding scars by Halichondria panicea (Pallas, 1766). *J. Exp. Mar. Bio. Ecol.* **327**: 36–46.
Knowlton, A.L., Pierson, B.J., Talbot, S.L., and Highsmith, R.C. (2003) Isolation and characterization of microsatellite loci in the intertidal sponge Halichondria panicea. *Mol. Ecol. Notes* **3**: 560–562.
Kostic, A.D., Howitt, M.R., and Garrett, W.S. (2013) Exploring host-microbiota interactions in animal models and humans. *Genes Dev.* **27**: 701–18.
Kovacevic, G. (2012) Value of the Hydra model system for studying symbiosis. *Int. J. Dev. Biol.* **56**: 627–635.
Kowalchuk, G.A. (1999) New perspectives towards analysing fungal communities in terrestrial environments. *Curr. Opin. Biotechnol.* **10**: 247–251.
Kravchuk, O.I., Lavrov, A.I., Finoshin, A.D., Gornostaev, N.G., Georgiev, A.A., Abaturova, S.B., et al. (2018) The Role of Epibionts of Bacteria of the Genus Pseudoalteromonas and Cellular Proteasomes in the Adaptive Plasticity of Marine Cold-Water Sponges. *Dokl. Biochem. Biophys.* **479**: 77–79.
Kumala, L. and Canfield, D.E. (2018) Contraction Dynamics and Respiration of Small Single-Osculum Explants of the Demosponge Halichondria panicea. *Front. Mar. Sci.* **5**: 410.
Kumala, L., Riisgård, H.U., and Canfield, D.E. (2017) Osculum dynamics and filtration activity studied in small single-osculum explants of the demosponge Halichondria panicea. *Mar Ecol Prog Ser* **572**: 117–128.
Kumar, S., Stecher, G., and Tamura, K. (2016) MEGA7 : Molecular Evolutionary Genetics Analysis version 7 . 0 for bigger datasets. *Mol. Biol. Evol.* **33**: 1–11.
Kwong, W.K., Engel, P., Koch, H., and Moran, N.A. (2014) Genomics and host specialization of honey bee and bumble bee gut symbionts. *Proc. Natl. Acad. Sci.* **111**: 11509–11514.
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**: 5695–9.
Lagesen, K., Hallin, P., Rødland, E.A., Stærfeldt, H.-H., Rognes, T., and Ussery, D.W. (2007) RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **35**: 3100–3108.
Lagkouvardos, I., Joseph, D., Kapfhammer, M., Giritli, S., Horn, M., Haller, D., and Clavel, T. (2016) IMNGS: A comprehensive open resource of processed 16S rRNA microbial profiles for ecology and diversity studies. *Sci. Rep.* **6**: 33721.
Lane, D.J., Pace, B., Olsen, G.J., Stahl, D.A., Sogin, M.L., and Pace, N.R. (1985) Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses. *Proc. Natl. Acad. Sci. U. S. A.* **82**: 6955–9.
Lang, S.., Beil, W.., Tokuda, H.., Wicke, C.., and Lurtz, V.. (2004) Improved Production of Bioactive Glucosylmannosyl-Glycerolipid by Sponge-Associated Microbacterium

Species. *Mar. Biotechnol.* **6**: 152–6.
Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Meth* **9**: 357–359.
Lapierre, P. and Gogarten, J.P. (2009) Estimating the size of the bacterial pan-genome. *Trends Genet.* **25**: 107–110.
Lee, O.O., Tsoi, M.M.Y., Li, X., Wong, P.-K., and Qian, P.-Y. (2007) Thalassococcus halodurans gen. nov., sp. nov., a novel halotolerant member of the Roseobacter clade isolated from the marine sponge Halichondria panicea at Friday Harbor, USA. *Int. J. Syst. Evol. Microbiol.* **57**: 1919–1924.
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.
Leinonen, R., Sugawara, H., and Shumway, M. (2011) The Sequence Read Archive. *Nucleic Acids Res.* **39**: D19–D21.
Leung, T.L.F. and Poulin, R. (2008) Parasitism, commensalism, and mutualism: exploring the many shades of symbioses. *Vie Milieu* **58**: 107.
Ley, R.E., Lozupone, C.A., Hamady, M., Knight, R., and Gordon, J.I. (2008) Worlds within worlds: evolution of the vertebrate gut microbiota. *Nat. Rev. Microbiol.* **6**: 776.
Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.
Liu, M.Y., Kjelleberg, S., and Thomas, T. (2011) Functional genomic analysis of an uncultured δ-proteobacterium in the sponge Cymbastela concentrica. *ISME J.* **5**: 427–435.
Lopanik, N.B. (2014) Chemical defensive symbioses in the marine environment. *Funct. Ecol.* **28**: 328–340.
Louis, P., Hold, G.L., and Flint, H.J. (2014) The gut microbiota, bacterial metabolites and colorectal cancer. *Nat. Rev. Microbiol.* **12**: 661–672.
Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, et al. (2004) ARB: a software environment for sequence data. *Nucleic Acids Res.* **32**: 1363–1371.
Lüskow, F., Riisgård, H., Solovyeva, V., and Brewer, J. (2019) Seasonal changes in bacteria and phytoplankton biomass control the condition index of the demosponge Halichondria panicea in temperate Danish waters. *Mar. Ecol. Prog. Ser.* **608**: 119–132.
Magoc, T. and Salzberg, S.L. (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**: 2957–2963.
Marchesi, J.R. and Ravel, J. (2015) The vocabulary of microbiome research: a proposal. *Microbiome* **3**: 31.
Martin, B.D. and Schwab, E. (2013) Current usage of symbiosis and associated terminology. *Int. J. Biol.* **5**: 32.
McFall-Ngai, M. (2014) Divining the essence of symbiosis: insights from the squid-vibrio model. *PLoS Biol.* **12**: e1001783.
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. U. S. A.* **110**: 3229–36.
McFall-Ngai, M.J. and Ruby, E.G. (1991) Symbiont recognition and subsequent morphogenesis as early events in an animal-bacterial mutualism. *Science* **254**: 1491–4.
McMurdie, P.J. and Holmes, S. (2013) phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One* **8**: e61217.
Mehbub, M., Lei, J., Franco, C., and Zhang, W. (2014) Marine sponge derived natural products between 2001 and 2010: Trends and opportunities for discovery of bioactives.

Mar. Drugs **12**: 4539–4577.
Meier-Kolthoff, J.P., Auch, A.F., Klenk, H.-P., and Göker, M. (2013) Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* **14**: 60.
Mendola, D. (2003) Aquaculture of three phyla of marine invertebrates to yield bioactive metabolites: Process developments and economics. *Biomol. Eng.* **20**: 441–458.
Metzker, M.L. (2010) Sequencing technologies — the next generation. *Nat. Rev. Genet.* **11**: 31–46.
Mills, D.B., Ward, L.M., Jones, C., Sweeten, B., Forth, M., Treusch, A.H., and Canfield, D.E. (2014) Oxygen requirements of the earliest animals. *Proc. Natl. Acad. Sci.* **111**: 4168–4172.
Moitinho-Silva, L., Nielsen, S., Amir, A., Gonzalez, A., Ackermann, G.L., Cerrano, C., et al. (2017) The sponge microbiome project. *Gigascience* **6**: 1–7.
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**: 752.
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.
NCBI Resource Coordinators (2017) Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **45**: D12–D17.
Neave, M.J., Apprill, A., Ferrier-Pagès, C., and Voolstra, C.R. (2016) Diversity and function of prevalent symbiotic marine bacteria in the genus Endozoicomonas. *Appl. Microbiol. Biotechnol.* **100**: 8315–8324.
Nguyen, M.T.H.D., Liu, M., and Thomas, T. (2014) Ankyrin-repeat proteins from sponge symbionts modulate amoebal phagocytosis. *Mol. Ecol.* **23**: 1635–1645.
Nichols, D., Cahoon, N., Trakhtenberg, E.M., Pham, L., Mehta, A., Belanger, A., et al. (2010) Use of ichip for high-throughput in situ cultivation of "uncultivable" microbial species. *Appl. Environ. Microbiol.* **76**: 2445–50.
Nichols, S.A., Dayel, M.J., and King, N. (2009) Genomic, phylogenetic, and cell biological insights into metazoan origins. *Anim. Evol. Genes, Genomes, Foss. Trees* 24–32.
Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P.A. (2017) metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* **27**: 824–834.
Osinga, R., Tramper, J., and Wijffels, R.H. (1999) Cultivation of marine sponges. *Mar. Biotechnol.* **1**: 509–532.
Palumbi, S.R. (1986) How Body Plans Limit Acclimation: Responses of a Demosponge to Wave Force. *Ecology* **67**: 208–214.
Palumbi, S.R. (1984) Tactics of acclimation: morphological changes of sponges in an unpredictable environment. *Science (80-.).* **225**: 1478–1480.
Peacock, K.A. (2011) Symbiosis in Ecology and Evolution. *Philos. Ecol.* **11**: 219–250.
Peattie, M.E. and Hoare, R. (1981) The sublittoral ecology of the Menai Strait: II. The sponge Halichondria panicea (Pallas) and its associated fauna. *Estuar. Coast. Shelf Sci.* **13**: 621–635.
Pernthaler, J., Glöckner, F.-O., Schönhuber, W., and Amann, R.B.T.-M. in M. (2001) Fluorescence in situ hybridization (FISH) with rRNA-targeted oligonucleotide probes. In, *Marine Microbiology*. Academic Press, pp. 207–226.
Perovic, S., Wichels, A., Schütt, C., Gerdts, G., Pahler, S., Steffen, R., and Müller, W.E.G. (1998) Neuroactive compounds produced by bacteria from the marine sponge Halichondria panicea: Activation of the neuronal NMDA receptor. *Environ. Toxicol.*

Pharmacol. **6**: 125–133.
Pianka, E.R. (2011) Evolutionary ecology Eric R. Pianka.
Piel, J. (2009) Metabolites from symbiotic bacteria. *Nat. Prod. Rep.* **26**: 338–362.
Pronzato, R. (2003) Mediterranean sponge fauna: a biological, historical and cultural heritage. *Biogeogr. J. Integr. Biogeogr.* **24**:.
Pruesse, E., Peplies, J., and Glöckner, F.O. (2012) SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**: 1823–1829.
Purushottama, G., Venkateshvaran, K., Pani Prasad, K., and Nalini, P. (2009) Bioactivities of extracts from the marine sponge Halichondria panicea. *J. Venom. Anim. Toxins Incl. Trop. Dis.* **15**: 444–459.
Qin, J., Li, R., Raes, J., Arumugam, M., Burgdorf, K.S., Manichanh, C., et al. (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* **464**: 59–65.
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**: D590–D596.
Quince, C., Lanzen, A., Davenport, R.J., and Turnbaugh, P.J. (2011) Removing Noise From Pyrosequenced Amplicons. *BMC Bioinformatics* **12**: 38.
R Core Team (2017) A language and environment for statistical computing. In, *R foundation for Statistical Computing*. Vienna, Austria.
Razumov, A.S. (1932) The direct method of calculation of bacteria in water: comparison with the Koch method. *Mikrobiologija* **1**: 131–146.
Reincke, T. and Barthel, D. (1997) Silica uptake kinetics of Halichondria panicea in Kiel Bight. *Mar. Biol.* **129**: 591–593.
Reiswig, H.M. (1971) 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.
Reynisson, E., Lauzon, H.L., Magnússon, H., Jónsdóttir, R., Olafsdóttir, G., Marteinsson, V., and Hreggvidsson, G.O. (2009) Bacterial composition and succession during storage of North-Atlantic cod (Gadus morhua) at superchilled temperatures. *BMC Microbiol.* **9**: 250.
Riisgård, H.U., Kumala, L., and Charitonidou, K. (2016) Using the *F/R* -ratio for an evaluation of the ability of the demosponge *Halichondria panicea* to nourish solely on phytoplankton versus free-living bacteria in the sea. *Mar. Biol. Res.* **12**: 907–916.
Riisgård, H.U. and Larsen, P.S. (1995) Filter-feeding in marine macro-invertebrates: Pump characteristics, modelling and energy cost. *Biol. Rev.* **70**: 67–106.
Riisgård, H.U., Thomassen, S., Jakobsen, H., Weeks, J.M., and Larsen, P.S. (1993) Suspension feeding in marine sponges Halichondria panicea and Haliclona urceolus: effects of temperature on filtration rate and energy cost of pumping. *Mar. Ecol. Prog. Ser.* **96**: 177–188.
Rizzo, C., Syldatk, C., Hausmann, R., Gerçe, B., Longo, C., Papale, M., et al. (2018) The demosponge *Halichondria (Halichondria) panicea* (Pallas, 1766) as a novel source of biosurfactant-producing bacteria. *J. Basic Microbiol.* **58**: 532–542.
RStudio Team (2016) Rstudio: Integrated development environment for r [Computer software manual]. Boston, MA.
Ruby, E.G. and Lee, K.-H. (1998) The Vibrio fischeri-Euprymna scolopes light organ association: current ecological paradigms. *Appl. Environ. Microbiol.* **64**: 805–812.
Sacristán-Soriano, O., Becerro, M.A., and Becerro, M.A. (2016) Publication impact in

sponge chemical and microbial ecology. *Sci. Mar.* **80**: 555.
Sata, N.U., Matsunaga, S., Fusetani, N., and van Soest, R.W.M. (1999) Aurantosides D, E, and F: New Antifungal Tetramic Acid Glycosides from the Marine Sponge Siliquariaspongia j aponica. *J. Nat. Prod.* **62**: 969–971.
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., Hentschel, U., and Taylor, M.W. (2012) Deep sequencing reveals diversity and community structure of complex microbiota in five Mediterranean sponges. *Hydrobiologia* **687**: 341–351.
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.
Schneemann, I., Kajahn, I., Ohlendorf, B., Zinecker, H., Erhard, A., Nagel, K., et al. (2010) Mayamycin, a cytotoxic polyketide from a streptomyces strain isolated from the marine sponge halichondria panicea. *J. Nat. Prod.* **73**: 1309–1312.
Schneemann, I., Nagel, K., Kajahn, I., Labes, A., Wiese, J., and Imhoff, J.F. (2010) Comprehensive investigation of marine actinobacteria associated with the sponge halichondria panicea. *Appl. Environ. Microbiol.* **76**: 3702–3714.
Schönberg, C.H.L. and Barthel, D. (1997) Inorganic skeleton of the demosponge Halichondria panicea. Seasonality in spicule production in the Baltic Sea. *Mar. Biol.* **130**: 133–140.
Shanahan, F. (2002) The host–microbe interface within the gut. *Best Pract. Res. Clin. Gastroenterol.* **16**: 915–931.
Shigenobu, S. and Wilson, A.C.C. (2011) Genomic revelations of a mutualism: the pea aphid and its obligate bacterial symbiont. *Cell. Mol. Life Sci.* **68**: 1297–1309.
Siegl, A. and Hentschel, U. (2010) PKS and NRPS gene clusters from microbial symbiont cells of marine sponges by whole genome amplification. *Environ. Microbiol. Rep.* **2**: 507–513.
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. (2012) Sponge-specific clusters revisited: a comprehensive phylogeny of sponge-associated microorganisms. *Environ. Microbiol.* **14**: 517–524.
Simpson, T.L. (1984) Gamete, Embryo, Larval Development. In, *The Cell Biology of Sponges*. Springer New York, New York, NY, pp. 341–413.
Simpson, T.L. (2012) The cell biology of sponges Springer Science & Business Media.
Sipkema, D., Franssen, M.C.R., Osinga, R., Tramper, J., and Wijffels, R.H. (2005) Marine sponges as pharmacy. *Mar. Biotechnol.* **7**: 142–162.
Sipkema, D., Osinga, R., Schatton, W., Mendola, D., Tramper, J., and Wijffels, R.H. (2005) Large-scale production of pharmaceuticals by marine sponges: Sea, cell, or synthesis? *Biotechnol. Bioeng.* **90**: 201–222.
Sipkema, D., Van Wielink, R., Van Lammeren, A.A.M., Tramper, J., Osinga, R., and Wijffels, R.H. (2003) Primmorphs from seven marine sponges : formation and structure. *J. Biotechnol.* **100**: 127–139.
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**: 2465–2478.

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.

Sogawa, K., Watanabe, M., Sato, K., Segawa, S., Ishii, C., Miyabe, A., et al. (2011) Use of the MALDI BioTyper system with MALDI–TOF mass spectrometry for rapid identification of microorganisms. *Anal. Bioanal. Chem.* **400**: 1905–1911.

Sokolova, A.M., Pozdnyakov, I.R., Ereskovsky, A. V., and Karpov, S.A. (2019) Kinetid structure in larval and adult stages of the demosponges Haliclona aquaeductus (Haplosclerida) and Halichondria panicea (Suberitida). *Zoomorphology* 1–14.

Spor, A., Koren, O., and Ley, R. (2011) Unravelling the effects of the environment and host genotype on the gut microbiome. *Nat. Rev. Microbiol.* **9**: 279.

Staley, J.T. and Konopka, A. (1985) Measurement of in Situ Activities of Nonphotosynthetic Microorganisms in Aquatic and Terrestrial Habitats. *Annu. Rev. Microbiol.* **39**: 321–346.

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.

Steinert, G., Stauffer, C.H., Aas-valleriani, N., Borchert, E., Bhushan, A., Campbell, A., et al. (2018) BluePharmTrain : Biology and Biotechnology of Marine Sponges.

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.

Swofford, D.L. and Maddison, W.P. (1987) Reconstructing ancestral character states under Wagner parsimony. *Math. Biosci.* **87**: 199–229.

Tamura, K. and Nei, M. (1993) Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol. Biol. Evol.* **10**: 512–526.

Tatusov, R.L., Galperin, M.Y., Natale, D.A., and Koonin, E. V. (2000) The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* **28**: 33–36.

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. (2012) 'Sponge-specific' bacteria are widespread (but rare) in diverse marine environments. *Isme J.* **7**: 438.

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.

Thomas, T., Rusch, D., DeMaere, M.Z., Yung, P.Y., Lewis, M., Halpern, A., et al. (2010) Functional genomic signatures of sponge bacteria reveal unique and shared features of symbiosis. *ISME J.* **4**: 1557–1567.

Thomassen, S. and Riisgård, H.U. (1995) Growth and energetics of the sponge Halichondria panicea. *Mar. Ecol. Prog. Ser.* **128**: 239–246.

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 L. ophophysema eversa harbours potential symbionts responsible for the nutrient conversions of carbon, nitrogen and sulfur. *Environ. Microbiol.* **18**: 2481–2494.

Tindall, B.J., Rosselló-Móra, R., Busse, H.-J., Ludwig, W., and Kämpfer, P. (2010) Notes on the characterization of prokaryote strains for taxonomic purposes. *Int. J. Syst. Evol. Microbiol.* **60**: 249–266.

Toth, G.B. and Lindeborg, M. (2008) Water-soluble compounds from the breadcrumb sponge Halichondria panicea deter attachment of the barnacle Balanus improvisus. *Mar. Ecol. Prog. Ser.* **354**: 125–132.
Trager, W. (2012) Living together: the biology of animal parasitism Springer Science & Business Media.
Turnbaugh, P.J., Ley, R.E., Hamady, M., Fraser-Liggett, C.M., Knight, R., and Gordon, J.I. (2007) The Human Microbiome Project. *Nature* **449**: 804–810.
Vacelet, J. and Boury-Esnault, N. (1995) Carnivorous sponges. *Nature* **373**: 333–335.
Vacelet, J. and Donadey, C. (1977) Electron microscope study of the association between some sponges and bacteria. *J. Exp. Mar. Bio. Ecol.* **30**: 301–314.
Venn, A.A., Loram, J.E., and Douglas, A.E. (2008) Photosynthetic symbioses in animals. *J. Exp. Bot.* **59**: 1069–1080.
Voultsiadou, E. (2010) Therapeutic properties and uses of marine invertebrates in the ancient Greek world and early Byzantium. *J. Ethnopharmacol.* **130**: 237–247.
Waldron, F.M., Stone, G.N., and Obbard, D.J. (2018) Metagenomic sequencing suggests a diversity of RNA interference-like responses to viruses across multicellular eukaryotes. *PLoS Genet.* **14**: e1007533.
Wang, Q., Garrity, G.M., Tiedje, J.M., and Cole, J.R. (2007) Naive Bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. *Appl. Environ. Microbiol.* **73**: 5261–7.
Warnecke, F., Luginbühl, P., Ivanova, N., Ghassemian, M., Richardson, T.H., Stege, J.T., et al. (2007) Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. *Nature* **450**: 560–565.
Webster, N.S. (2007) Sponge disease: a global threat? *Environ. Microbiol.* **9**: 1363–1375.
Webster, N.S., Cobb, R.E., Soo, R., Anthony, S.L., Battershill, C.N., Whalan, S., and Evans-Illidge, E. (2011) Bacterial Community Dynamics in the Marine Sponge Rhopaloeides odorabile Under In Situ and Ex Situ Cultivation. *Mar. Biotechnol.* **13**: 296–304.
Webster, N.S. and Taylor, M.W. (2012) Marine sponges and their microbial symbionts: Love and other relationships. *Environ. Microbiol.* **14**: 335–346.
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. and Thomas, T. (2016) The Sponge Hologenome. *MBio* **7**: e00135-16.
Webster, N.S., Wilson, K.J., Blackall, L.L., and Hill, R.T. (2001) Phylogenetic Diversity of Bacteria Associated with the Marine Sponge. *Appl. Environ. Microbiol.* **67**: 434–444.
Weinstock, G.M. (2012) Genomic approaches to studying the human microbiota. *Nature* **489**: 250–256.
Wichels, A., Würtz, S., Döpke, H., Schütt, C., and Gerdts, G. (2006) Bacterial diversity in the breadcrumb sponge Halichondria panicea (Pallas). *FEMS Microbiol. Ecol.* **56**: 102–118.
Wicke, C., Hüners, M., Wray, V., Nimtz, M., Bilitewski, U., and Lang, S. (2000) Production and Structure Elucidation of Glycoglycerolipids from a Marine Sponge-Associated Microbacterium Species.
Wickham, H. (2009) ggplot2: Elegant Graphics for Data Analysis Springer-Verlag. *New York*.
Wilkinson, C.R. (1978) Microbial associations in sponges. I. Ecology, physiology and microbial populations of coral reef sponges. *Mar. Biol.* **49**: 161–167.
Wille, N. (1910) Algologische Notizen XVI–XXI. *Nyt Mag. Naturvid.* **48**: 298–302.
Williams, K.P., Gillespie, J.J., Sobral, B.W.S., Nordberg, E.K., Snyder, E.E., Shallom, J.M.,

and Dickerman, A.W. (2010) Phylogeny of gammaproteobacteria. *J. Bacteriol.* **192**: 2305–14.

Witte, U., Barthel, D., and Tendal, O. (1994) The reproductive cycle of the sponge Halichondria panicea Pallas (1766) and its relationship to temperature and salinity. *J. Exp. Mar. Bio. Ecol.* **183**: 41–52.

Woese, C.R. (1987) Bacterial evolution. *Microbiol. Rev.* **51**: 221.

Woese, C.R. and Fox, G.E. (1977) Phylogenetic structure of the prokaryotic domain: the primary kingdoms. *Proc. Natl. Acad. Sci. U. S. A.* **74**: 5088–90.

Wolfrath, B. and Barthel, D. (1989) Production of faecal pellets by the marine sponge Halichondria panicea Pallas (1766). *J. Exp. Mar. Bio. Ecol.* **129**: 81–94.

Wu, M. and Eisen, J.A. (2008) A simple, fast, and accurate method of phylogenomic inference. *Genome Biol.* **9**: R151.

Wu, M. and Scott, A.J. (2012) Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. *Bioinformatics* **28**: 1033–1034.

Wulff, J.L. (2006) Ecological interactions of marine sponges. *Can. J. Zool.* **84**: 146–166.

Yarza, P., Richter, M., Peplies, J., Euzeby, J., Amann, R., Schleifer, K.-H., et al. (2008) The All-Species Living Tree project: A 16S rRNA-based phylogenetic tree of all sequenced type strains. *Syst. Appl. Microbiol.* **31**: 241–250.

Yilmaz, P., Parfrey, L.W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., et al. (2014) The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res.* **42**: D643–D648.

Yin, Z., Zhu, M., Davidson, E.H., Bottjer, D.J., Zhao, F., and Tafforeau, P. (2015) Sponge grade body fossil with cellular resolution dating 60 Myr before the Cambrian. *Proc. Natl. Acad. Sci. U. S. A.* **112**: E1453-60.

Yoon, S.-H., Ha, S.-M., Kwon, S., Lim, J., Kim, Y., Seo, H., and Chun, J. (2017) Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int. J. Syst. Evol. Microbiol.* **67**: 1613–1617.

Yoon, S.-H., Ha, S., Lim, J., Kwon, S., and Chun, J. (2017) A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* **110**: 1281–1286.

Yoon, S.S., Kim, E.-K., and Lee, W.-J. (2015) Functional genomic and metagenomic approaches to understanding gut microbiota–animal mutualism. *Curr. Opin. Microbiol.* **24**: 38–46.

Original papers

Paper I Knobloch, S., Jóhannsson, R., Marteinsson, V. 2019 **Bacterial diversity in the marine sponge *Halichondria panicea* from Icelandic waters and host-specificity of its dominant symbiont "*Candidatus* Halichondribacter symbioticus"**, *FEMS Microbiology Ecology*, 95 (1). fiy220, https://doi.org/10.1093/femsec/fiy220.

Paper II Knobloch, S., Jóhannsson, R., Marteinsson, V. **Genome analysis of sponge symbiont "*Candidatus* Halichondribacter symbioticus" shows genomic adaptation to a host-dependent lifestyle.** Unpublished Manuscript

Paper III Knobloch, S., Jóhannsson, R., Marteinsson, V. **Co-cultivation of the marine sponge *Halichondria panicea* and its associated microorganisms.** Unpublished Manuscript

Paper IV Knobloch, S., Daussin, A., Jóhannsson, R., Marteinsson, V. ***Pelagibaculum spongiae* gen. nov., sp. nov., isolated from a marine sponge in South-West Iceland.** Unpublished Manuscript

Paper V Knobloch, S., Jóhannsson, R., Marteinsson, V. ***Tenacibaculum islandicus* sp. nov., isolated from a marine sponge contains secondary metabolite gene clusters.** Unpublished Manuscript

Paper VI Knobloch, S., Jóhannsson, R., Marteinsson, V. ***Endozoicomonas halichondricola* sp. nov., isolated from the marine sponge *Halichondria panicea*.** Unpublished Manuscript

Paper I

I

Bacterial diversity in the marine sponge *Halichondria panicea* from Icelandic waters and host-specificity of its dominant symbiont "*Candidatus* Halichondribacter symbioticus"

Stephen Knobloch, Ragnar Jóhannsson, Viggó Þór Marteinsson

Published in *FEMS Microbiology Ecology*

Author contributions:

Conceived and designed the study: SK, RJ, VM
Performed fieldwork: SK
Performed laboratory work: SK
Analysed data: SK
Wrote the manuscript: SK
Reviewed and edited the manuscript: SK, RJ, VM
Correspondence (publication process): SK

FEMS Microbiology Ecology, 95, 2019, fiy220

doi: 10.1093/femsec/fiy220
Advance Access Publication Date: 10 November 2018
Research Article

RESEARCH ARTICLE

Bacterial diversity in the marine sponge *Halichondria panicea* from Icelandic waters and host-specificity of its dominant symbiont "*Candidatus* Halichondribacter symbioticus"

Stephen Knobloch[1,2,*,†], Ragnar Jóhannsson[3] and Viggó Marteinsson[1,4]

[1]Microbiology Group, Department of Research and Innovation, Matís ohf., Vinlandsleid 12, 113 Reykjavik, Iceland, [2]Faculty of Life and Environmental Sciences, University of Iceland, Saemundargata 2, 101 Reykjavík, Iceland, [3]Marine and Freshwater Research Institute, Hafrannsóknastofnun, Skúlagata 4, 101 Reykjavik, Iceland and [4]Faculty of Food Science and Nutrition, University of Iceland, Saemundargata 2, 101 Reykjavik, Iceland

[*]**Corresponding author:** Matis ohf. Vinlandsleid 12, 113 Reykjavik, Iceland. Tel: +354 7806179; Fax: +354 4225001; E-mail: stephen@matis.is

One sentence summary: The breadcrumb sponge (*Halichondria panicea*) from Icelandic waters hosts a dominant bacterial symbiont.

Editor: Julie Olson

[†]Stephen Knobloch, http://orcid.org/0000-0002-0930-8592

ABSTRACT

Marine sponges can harbour diverse bacteria that contribute to host metabolism and defence. Identifying these stable members of sponge bacterial communities remains a necessary step in understanding their ecological roles and underlying co-evolutionary processes. In this study, we applied high-throughput sequencing of 16S rRNA gene amplicons, ribosomal nucleotide variant analysis and fluorescence *in situ* hybridisation to characterise the core members of the bacterial community in the marine sponge *Halichondria panicea* from Icelandic waters. We show that the core bacterial community across all samples consisted of a single, dominant bacterial taxon, for which we propose a candidate status '*Candidatus* Halichondribacter symbioticus'. Comparison against public databases showed that '*Ca*. H. symbioticus' is both a highly abundant specialist in *H. panicea* and a low abundant opportunist in other sponge species. Additionally, *H. panicea* with and without '*Ca*. H. symbioticus' co-exist in similar locations in the North Atlantic. This dichotomy paired with the presence of geographically distinct ribosomal sequence variants of the symbiont make *H. panicea* an interesting sponge species for studying sponge-symbiont co-evolution and functional interactions.

Keywords: sponge; symbiosis; cultivation; marine; *Alphaproteobacterium*; 16S rRNA gene

INTRODUCTION

Marine sponges (phylum *Porifera*) are among the oldest extant metazoans (Feuda *et al.* 2017). Increased attention has been placed on their microbial communities due to their complex, sponge-specific associations and the biotechnological potential of their secondary metabolites (Hentschel *et al.* 2002; Wang 2006). Being perhaps the oldest of all animal symbiosis these associations have also attracted interest as models for understanding early animal evolution and host–microbe coevolution (Taylor *et al.* 2007b).

Sponges inhabit aquatic habitats from freshwater sites to the deep sea (Chambers 2003). Over 8500 species have been described so far, most of which occur in the marine environment

Received: 21 August 2018; **Accepted:** 7 November 2018

(Van Soest *et al*. 2012). Sponges are sessile filter feeders actively pumping water through their aquiferous system while filtering and ingesting dissolved nutrients and food particles such as microorganisms from the water column (Riisgård and Larsen 1995; De Goeij *et al*. 2008). Apart from feeding on microorganisms, some sponge species host a large abundance of microorganisms within their mesohyl, in some cases accounting for up to 40% of the mesohyl volume or several billion bacteria per gram of sponge wet weight (Wilkinson 1978; Gloeckner *et al*. 2014). Our understanding of the taxonomic and functional diversity of these sponge-associated microorganisms has rapidly increased over the past decades, in part, due to the advance of modern sequencing technology. We now know that some sponges host a remarkably large diversity of microorganisms (Webster *et al*. 2010; Schmitt, Hentschel and Taylor 2012), some of which are highly sponge-specific (Hentschel *et al*. 2002; Simister *et al*. 2012; Taylor *et al*. 2012; Thomas *et al*. 2016). For instance, members of the candidate phylum '*Poribacteria*' have been nearly exclusively found in sponges, both across different species and geographical locations (Lafi *et al*. 2009; Kamke *et al*. 2013). Many sponge-associated microorganisms possess functions to maintain the symbiotic relationship to their host (Thomas *et al*. 2010; Díez-Vives *et al*. 2017; Slaby *et al*. 2017) and are involved in potentially mutualistic behaviour such as defence (Flatt *et al*. 2005), host metabolism (Erwin and Thacker 2007; Freeman and Thacker 2011; Tian *et al*. 2016) and the removal of contaminants or metabolic waste products (Schläppy *et al*. 2010; Siegl *et al*. 2011; Keren *et al*. 2017). However, the ecological functions of most sponge symbionts are not yet fully understood, nor do we know what role microbial symbionts play in sponge-symbiont co-evolution (Webster and Taylor 2012; Webster and Thomas 2016).

Apart from sponge species that host a large diversity and abundance of sponge-associated microorganisms, other species contain less diverse and abundant microbial assemblages, a phenomenon described as the high microbial abundance (HMA), low microbial abundance (LMA) sponge dichotomy (Hentschel, Usher and Taylor 2006; Gloeckner *et al*. 2014). LMA sponges often harbour single dominant bacterial taxa (Giles *et al*. 2013) and these core taxa are often sponge species specific (Erwin *et al*. 2015). It has therefore been proposed that LMA sponges could function as models for understanding sponge-microbe interactions due to the lower complexity of their microbial communities (Croué *et al*. 2013). Identifying the stable members of sponge-associated microbial communities, however, remains a necessary first step in understanding these interactions (Taylor *et al*. 2007a).

The LMA sponge *Halichondria panicea* inhabits coastal areas around the globe and is a widely studied sponge species in terms of its biology and ecophysiology (c.f.) (Barthel 1986; Barthel and Wolfrath 1989; Witte, Barthel and Tendal 1994; Thomassen and Riisgård 1995; Mills *et al*. 2014; Riisgård, Kumala and Charitonidou 2016). Its bacterial diversity has previously been studied using cultivation-dependent (Imhoff and Stöhr 2003; Schneemann *et al*. 2010) and -independent methods (Althoff *et al*. 1998; Wichels *et al*. 2006; Lee, Wong and Qian 2009; Naim *et al*. 2014; Steinert *et al*. 2017), yet a conclusive list of its permanent-associated bacteria from different geographical locations, as well as the nature of their association remains to be identified.

In this study, we analysed the bacterial community of wild and cultivated *H. panicea*, and seawater samples from different geographical locations around Iceland and across different seasons using a deep sequencing approach and 16S rRNA fluorescence *in situ* hybridisation (FISH) to identify and describe the core members of the sponge-associated bacterial community. In addition, analysis of near full length 16S rRNA gene variants was performed to detect differences in bacterial populations from geographically separate *H. panicea*.

MATERIAL AND METHODS

Sample collection and identification

Sponge and seawater samples

Samples of wild *H. panicea* for bacterial diversity analysis were collected from four intertidal or subtidal locations (depths between 5 and 15 m) in Iceland between 2013 and 2017 (Table 1, Supporting Information). 21 sponge samples and six seawater samples were collected from Grótta (*Gro*), (64°09′ N, 22°00′ W); 39 sponge samples and 13 seawater samples from Eyrarbakki (*Eyr*), (63°51′ N 21°09′ W); six sponge samples and one seawater sample from Hafnir (*Haf*) (63°55′ N 22°41′ W); and 6 sponge samples from Skútan (*Sku*) (65°44′ N 18°06′ W). Samples from at least three different sponge individuals were collected at each sampling occasion. Approximately, 1 cm^3 of sponge body was removed, taking care to include similar amounts of ectosome and choanosome between samples. Sponge samples were rinsed with sterile artificial seawater to remove bacteria attached to the outer layer of the sponge section and transported on ice to the laboratory. 1 l of seawater was collected in close proximity to the sponges and filtered through 0.2 μm Sterivex filters (Millipore). Filters containing residue and sponge samples were frozen at −20 to −80°C until further processed for DNA extraction.

Taxonomic identification of sponge samples and mitochondrial genome of H. panicea

Initial identification was performed according to the species description in Ackers, Moss and Picton (2007). In addition, the complete mitochondrial DNA of selected individuals (Table 1, Supporting Information) was extracted from a metagenome (data not shown) and the COI gene compared against the NCBI nucleotide database (NCBI Resource Coordinators 2017). Sequences with the highest identity belonged to *H. panicea* voucher 'BELUM<GBR>:Mc4070' (Genbank: KC869423) and *H. panicea* voucher 'SB16S' (Genbank: KY492588) with 100% and 99% sequence identity respectively. The full mitochondrial DNA sequences including the COI gene of *H. panicea* from *Gro* and *Sku* are deposited under DDBJ/ENA/GenBank accession numbers MH756603 and MH756604 respectively.

Ex situ cultivation of H. panicea

Eight *H. panicea* individuals attached to stones were collected from the intertidal area in Eyrarbakki and transferred to a recirculating seawater aquarium at a commercial aquaculture farm less than 200 m from the collection site. The tank set-up consisted of a bacterial moving bed biofilter, aeration for water movement and oxygenation, and artificial lighting set to 12 h dark and 12 h light. Water exchange rates were adjusted to replace the tank volume of 360 l twice per day and was supplied with sand-filtered seawater from the collection site. Active pumping of the sponges was examined at least twice weekly using food-grade dye injected into the water column directly above an osculum of each individual. Growth or decline of the sponge attachment area to the substrate was evaluated using photographic images. Samples for 16S rRNA gene amplicon sequencing were collected directly upon transfer, as well as after two, four and six months in the seawater aquarium. Additionally, two seawater samples from within the aquarium

Table 1. Specific and universal 16S rRNA FISH probes used in this study.

Probe	Sequence (5′–3′)	Specificity	Reference
Rhd473	AGC CGG GAC TTC TTC TAC TG	OTU1–'*Ca.* Halichondribacter symbioticus'	This study
Rhd1136	CTC TTT AGT GTC CCC AAC TG	OTU1–'*Ca.* Halichondribacter symbioticus'	This study
EUB338 I	GCT GCC TCC CGT AGG AGT	Eubacteria	(Amann *et al.* 1990)
EUB338 II	GCA GCC ACC CGT AGG TGT	Eubacteria	(Daims *et al.* 1999)
EUB338 III	GCT GCC ACC CGT AGG TGT	Eubacteria	(Daims *et al.* 1999)

were collected at the start of the experiment and after four months. Two sponges died during the six-month period and were removed as soon as necrosis was detected. Bacterial cell counts of coastal seawater and aquarium water were performed on 0.2 μm black polycarbonate membrane filters (Sigma) stained with DAPI according to Porter and Feig (1980).

16S rRNA gene amplicon sequencing and analysis

DNA extraction

Genomic DNA of sponge samples for 16S rRNA gene amplicon sequencing was purified using the MasterPure Complete DNA and RNA purification Kit (Epicentre). In short, sponge samples were homogenised in laboratory grade Millipore water by passing them through a 90-gauge syringe 10 times or until sponge cells had visually disaggregated. Approximately, 100 μl of homogenized cell suspension was used for DNA extraction according to the manufacturer's instructions for total DNA purification. DNA quantity and integrity where determined on a Nanodrop 1000 Spectrophotometer (Thermo Scientific) and by visual observation on a 1% agarose gel. Seawater samples were processed in the same manner, apart from that filters containing seawater bacteria were ground in lysis buffer prior to following the protocol for DNA extraction.

PCR amplification, library construction and sequencing

For bacterial diversity analysis, a region covering the V3-V4 region of the 16S rRNA gene was amplified using the universal bacterial primer pair S-D-Bact-0341-b-S-17 (5′-CCTACGGGNGGCWGCAG-3′) / S-D-Bact-0785-a-A-21 (5′-GACTACHVGGGTATCTAATCC-3′) (Klindworth *et al.* 2013). PCR reactions were set up as followed: 50 ng template DNA, 5 μl 5X Phusion HF Buffer (New England BioLabs), 200 μM dNTPs, 0.2 μM of forward and 0.2 μM reverse primers containing Illumina overhang adapters, 0.2 μl Phusion High-Fidelity DNA polymerase and nuclease-free water to a total reaction volume of 25 μl. The thermocycler conditions were set to: Initial Denaturation at 98°C for 30 s followed by 30 cycles of denaturation at 98°C for 10 s, annealing at 52°C for 30 s and extension at 72°C for 30 s. Final extension was set to 72°C for 5 min. The PCR products were checked on 1% agarose gel for the correct amplicon size. Libraries were multiplexed using Nextera XT v2 indices (Illumina) and normalised using *SequalPrep* Normalization Plates (ThermoFisher Scientific) according to the manufacturer's instructions. Normalised libraries were pooled, denatured and diluted according to the *MiSeq System Denature and Dilute Libraries Guide* (Illumina) and sequenced on a MiSeq System with v3 chemistry and 2 × 300 cycles across four separate sequencing runs.

Bioinformatic analysis and statistics

Bioinformatic analyses, taxonomic assignment and all subsequent statistical analyses were performed in R version 3.4.2 (R Core Team 2017). Sequence variants (SVs) were inferred using the R package DADA2 version 1.4 (Callahan *et al.* 2016) as follows: Demultiplexed and adapter-trimmed sequences were quality filtered using the *filterAndTrim* command and following setting: truncLen = c(260,220), maxEE = 2, trimLeft = 15, truncQ = 2. Error rates were learned and SVs inferred for samples from each run independently using the *learnErrors* command with a subset of 1 million reads and the *dada* command on both forward and reverse reads. Both read ends were merged using the *mergePairs* command with default parameters. Sequences 50 bp smaller than the expected amplicon size of 445 bp were removed and chimeras detected using the *removeBimeraDenovo* command with *method* = *consensus*. Taxonomy was assigned with the *assignTaxonomy* command against a training set of the SILVA database version 128 (Quast *et al.* 2013). In order to infer relatedness of SVs, they were clustered into OTUs (Operational Taxonomic Units) at 99% identity threshold using the *cluster_smallmem* command of the UCLUST algorithm (Edgar 2010), sorted by read abundance. The SV and OTU tables were imported into the *R* package *phyloseq* (McMurdie and Holmes 2013). Singleton SVs and SVs assigned as Chloroplasts were removed from the sequence table. Plots were created using R package *ggplot2* (Wickham 2009). Dissimilarity matrices were calculated using the R package *vegan* (Oksanen *et al.* 2017) with method 'bray'. Hierarchical cluster plots were created using *hclust* of the R package *stats* with method 'average'.

Reference databases

In order to detect the presence of bacteria from the sponge samples in offshore seawater, a database with partial 16S rRNA gene sequences was created from seawater samples previously collected in conjunction with the annual research cruise *Bjarni Sæmundsson* carried out in May 2011, 2012 and 2013 (unpublished data). Raw sequences generated from a 454 GS-FLX (Roche) were quality filtered, trimmed and denoised in QIIME (Caporaso *et al.* 2010). The resulting 517.275 sequences where used to build a nucleotide database using the *makeblastdb* command of BLAST version 2.4.0+ (Camacho *et al.* 2009).

In addition, the to-date largest collection of sponge-associated bacteria reported in Moitinho-Silva *et al.* (2017a) was used as a reference database to detect the presence of sponge-associated bacteria from other study sites. The dataset (*Deblur.final.withtax.biom*) was retrieved from *Giga*DB and a BLAST database was constructed from the representative sequences. Sample locations were plotted onto a map using the R package *maps* (Becker *et al.* 2018) and *ggplot2*.

Single nucleotide variants of '*Ca.* H. symbioticus' 16S rRNA gene

DNA from three sponge individuals (Table 1, Supporting Information) at each sampling location and from different times in the year (apart from *Sku* where only samples from one

time point were available), was extracted as mentioned above. The near full length 16S rRNA gene was amplified using primer pair F9 (5′-GAGTTTGATCCTGGCTCAG-3′) and R1510 (5′-GGTTACCTTGTTACGACTT-3′) (Baker, Smith and Cowan 2003; Reynisson *et al.* 2009) and the same PCR reaction and thermocycler conditions as for 16S rRNA gene amplicon sequencing was used, apart from that the extension time was increased to 1 min. Purified amplicons were fragmented, amplified and indexed using the Nextera XT kit (Illumina) according to the manufacturer's description, and sequenced on a MiSeq (Illumina) with v3 chemistry. Reads were merged and assembled in Geneious v. 9.1.5 (Biomatters) and near full length 16S rRNA sequences which matched the '*Ca.* H. symbioticus' 16S rRNA V3-V4 region with ≥ 99% sequence identity were selected for alignment with additional sequences retrieved from Genbank. Alignments was performed with MUSCLE (Edgar 2004) and visualised in Boxshade (Hofmann and Baron 1996). Phylogenetic trees were constructed in Mega version 7 (Kumar, Stecher and Tamura 2016). A Maximum-likelihood (Felsenstein 1981; Tamura and Nei 1993) tree was built from alignments and evaluated by bootstrap analysis based on 1000 resamplings (Felsenstein 1985).

Histology and 16S rRNA fluorescence *in situ* hybridisation

In order to detect the abundance and location of '*Ca.* H. symbioticus' and other bacteria inside *H. panicea*, six wild individuals from different sampling events and six individuals from the *ex situ* cultivation experiment were subjected to 16S rRNA FISH. '*Ca.* H. symbioticus' specific probes Rhd473 and Rhd1136 (Table 1) where designed in ARB (Ludwig *et al.* 2004) using the full 16S rRNA gene sequence of '*Ca.* H. symbioticus' and checked against the NCBI nr/nt database for specificity. Sponge samples were fixed in freshly prepared 4% paraformaldehyde (PFA) in 1 x PBS for 12 h at 4°C directly after sample collection. Fixed samples were washed in PBS to remove residual PFA and stored in 70% ethanol for up to one year. Samples were rehydrated and submerged in 15% and subsequently 30% sucrose solution in 1 × PBS until saturation. Samples were covered with tissue freezing medium (Leica Biosystems), frozen on crushed dry ice and stored at −80°C until sectioning. Approximately, 5 µm thin sponge sections were cut on a Leica CM1800 cryotome set to −35°C with MX35 Ultra microtome blades (Thermo scientific). Cut sections on microscope slides were dehydrated in successive baths of 50%, 70%, 85% and 100% ethanol and subsequently air dried. Sponge sections were hybridized according to Pernthaler *et al.* (2001) in hybridisation buffer (900 mM NaCl, 20 mM Tris-HCl, 60% formamide and 0.01% SDS in deionized water) with Cy3 labelled probes Rhd473 and Rhd1136 and Alexa488 labelled universal bacterial probes EUB338 I—III (Table 1) at concentrations of 5 ng DNA μl^{-1} for 2 h at 46°C. Hybridised sponge sections were transferred to a wash buffer (14 mM NaCl, 20 mM Tris-HCl, 5 mM EDTA and 0.01% SDS in deionized water) at 48°C for 15 min, subsequently rinsed in deionized water, covered with Fluoroshield antifade containing DAPI (Sigma) and observed under a model BX51 epifluorescence microscope (Olympus) equipped with filters U-MNB2, U-MNG2 and U-MWU2. Epifluorescence images were processed in daime v. 2.1 (Daims, Lücker and Wagner 2006). Square zones, 0.01 mm^2 in size, from 36 FISH images were used to count bacterial and sponge cell abundance for wild and *ex situ* cultivated sponges. Cell counts are given as mean ± standard deviation.

Sequence deposits

The DDBJ/ENA/GenBank accession number for the 16S rRNA gene sequences of OTU1—'*Candidatus* Halichondribacter symbioticus' from South-West (Gro) and North (Sku) Iceland are deposited under MH734183 and MH734529, respectively. Raw 16S rRNA amplicon reads are deposited in the Sequence Read Archive under BioProject ID PRJNA495906.

RESULTS

Bacterial diversity in *H. panicea* from North and South-West Iceland

In order to identify the core sponge-associated members of the bacterial community in *H. panicea*, samples from four different location around Iceland and across different points of time were collected and subjected to high-throughput sequencing of the partial 16S rRNA gene. Approximately, 1.5 million quality filtered reads were generated from 72 *H. panicea* samples. Using the single-nucleotide resolution pipeline DADA2 a total of 6351 SVs were inferred which clustered into 5156 OTUs at 99% identity level and were taxonomically assigned to 44 phyla (Fig. S1, Supporting Information). For 20 seawater samples a total of 3903 SVs were inferred from approximately 0.3 million quality-filtered reads, clustering into 3112 OTUs and assigned to 39 phyla. The bacterial diversity in *H. panicea* samples assigned at class level did not differ markedly across sampling sites. The majority of reads were assigned to the class *Alphaproteobacteria*, followed by the classes *Gammaproteobacteria*, *Flavobacteriia*, *Planctomycetacia*, *Cyanobacteria* and *Verrucomicrobiae* (Fig. 1). One OTU, designated OTU1 and assigned to the class *Alphaproteobacteria*, was the most abundant OTU in sponge samples from all geographical locations and the only OTU found in all sponge samples (Fig. S2, Supporting Information). In addition, only 18 OTUs were shared between half of the sponge samples and the large majority of OTUs (4872 out of 5156) were shared between less than 10% of the samples.

Despite this low core sponge bacterial community, the remaining bacterial diversity detected within the sponges was not an exact representation of the surrounding seawater community. Excluding the dominant bacterium OTU1 and adjusting the remaining read abundance to 100%, only 42 ± 17% of the reads in the sponge samples from site *Eyr* were also detected in the corresponding seawater sample and 62 ± 16% of the read abundance in the seawater samples was detected in the respective sponge samples (Fig. 2). In addition, hierarchical clustering based on Bray–Curtis dissimilarity placed all sponge samples from *Gro* and *Eyr* into a separate group from their corresponding seawater samples. After excluding OTU1 from the dataset only two sample sets collected in *Eyr* in December and *Gro* in March 2015 clustered with its corresponding seawater sample (Fig. 2). Sample set Haf_Apr'14, collected in the subtidal area at depths between 5 and 15 m, fell within the same group of sponge samples collected in the intertidal zone and, after excluding OTU1 from the dataset, clustered into a separate group with sample set Haf_Nov'14, collected in the intertidal zone from the same site albeit at a different time of the year. The only sample set collected in North Iceland (Sku_Nov'13), also from a subtidal habitat, clustered separately to the sponge samples collected in South-West Iceland.

16S rRNA FISH with OTU1 specific und universal bacterial probes showed an even distribution of OTU1 throughout the sponge body of wild specimens (Fig. 3A–C). Other bacteria within

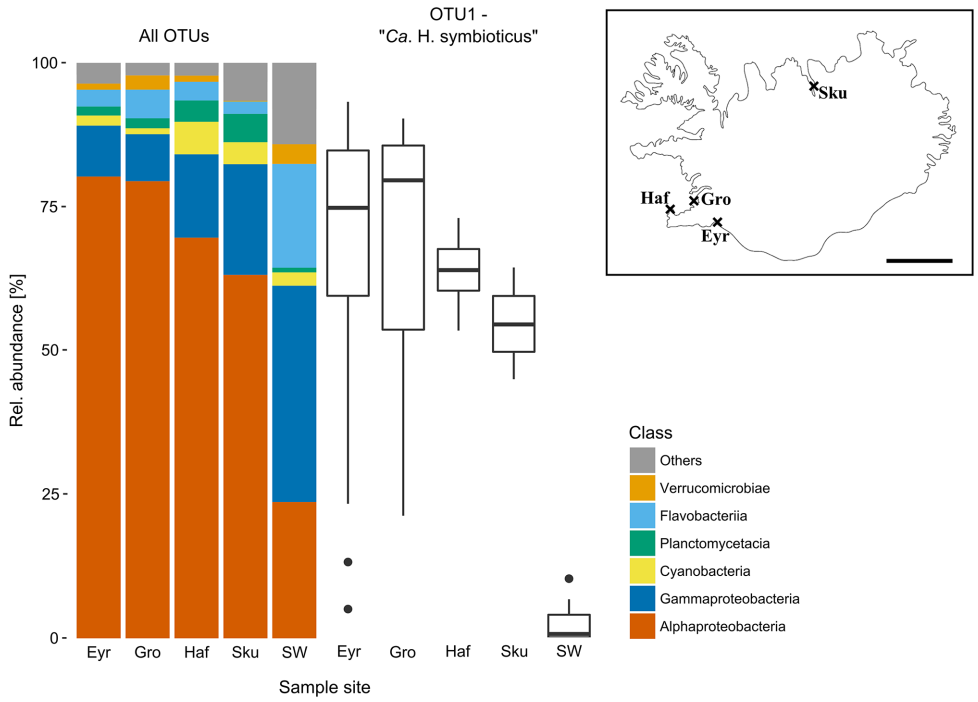


Figure 1. Relative abundance of OTUs assigned to taxonomic classes and of OTU1–'*Ca.* H. symbioticus', the dominant taxa, in *H. panicea* across sampling sites and in the surrounding seawater (SW). Two sponge samples collected in Eyr in September 2015 had a relative abundance of less than 20% for OTU1 and both samples are marked as outliers in the box plot. Sample numbers per site were: Eyr = 39, Gro = 21, Haf = 6, Sku = 6. Map of Iceland showing the four sampling sites (bar = 100 km).

the mesohyl appeared less abundant and were often present in small clusters (Fig. 3B). Cell counts of FISH images, showed an average OTU1 to sponge cell ratio of approximately 8:1 (450 ± 124 OTU1 to 58 ± 19 sponge cells) and OTU1 to other bacteria ratio of approximately 90:1 (450 ± 124 OTU1 to 5 ± 4 other bacterial cells) (Fig. 4), thereby confirming the dominance of OTU1 in the mesohyl of *H. panicea* and showing an underestimation of its relative abundance observed through 16S rRNA gene amplicon sequencing.

Bacterial diversity during *ex situ* cultivation of *H. panicea*

In order to evaluate the stability of the OTU1–*H. panicea* association, eight sponge individuals were transferred from the wild to an aquarium with filtered seawater for six months. Cell counts of microorganisms in the aquarium water ranged from 1.2 to 2.6 × 10^5 cells ml^{-1}, whereas seawater from the sponge collection site contained 5.3 to 20.7 × 10^5 cells ml^{-1} during the same time as the experiment. Throughout the duration of the experiment pumping could be detected in all living individuals. Of the initial sponges, two died before the end of the experiment, five decreased in observable attachment area and one increased in attachment area to the substrate (Fig. 5). In addition, all sponges exhibited a restructuring of the aquiferous system in conjunction with a widening of water canals and decreased body density, possibly in order to increase pumping activity due to the low presence of microbial food organisms.

Cell counts of 16S rRNA FISH images of sponge sections after six months in the seawater aquarium showed an OTU1 to sponge cell ratio of 1 to 9 (5 ± 9 OTU1 to 45 ± 8 sponge cells) representing a 69-fold reduction of OTU1 in the mesohyl compared to its wild counterpart (Fig. 4). A similar reduction of OTU1 was observed in all six sponge individuals. Additionally, 16S rRNA FISH images showed an uneven distribution of OTU1 throughout the sponge body restricted to sparsely distributed, small clusters of cells as depicted in Fig. 3D. Although the absolute abundance of OTU1 decreased during *ex situ* cultivation, a cell ratio of 5 to 2 (5 ± 9 OTU1 to 2 ± 5 other bacterial cells) between OTU1 and other bacterial cells demonstrated that OTU1 remained the dominant bacterium within the mesohyl of *H. panicea*. In contrast, partial 16S rRNA gene sequencing showed a reduction of OTU1 from an average relative abundance of 72% directly after transfer to the seawater aquarium to less than 2% of the relative abundance after two months (Fig. S3, Supporting Information), providing evidence that the reduction of OTU1 in the sponge mesohyl occurred within the first two months in the aquarium system, but also highlighting a strong discrepancy between the two methods used for analysing the relative abundance of OTU1 in the sponge mesohyl.

Ribosomal nucleotide variants and phylogenetics of the sponge symbiont

Near full length 16S rRNA gene sequences of OTU1 from all locations revealed two main ribosomal sequence variants between *H. panicea* collected in South-West Iceland (Eyr, Gro, Haf) and North Iceland (Sku) with 11 base mismatches across a 1421 bp region (Fig. S4, Supporting Information). A sequence comparison to the NCBI non-redundant nucleotide (nr/nt) database using BLAST closely matched the sequence to partial 16S rRNA gene sequences of uncultured bacteria in the host organism *H. panicea* with 98%–100% sequence similarity (GenBank accession numbers: AY948354, AY948358, KJ453525, Z88567, Z88569, Z88578) corresponding to the dominant *Alphaproteobacterium* found in the respective studies from the North Sea and Baltic Sea. A comparison between all near full-length sequences showed that the OTU1 population from South-West Iceland were genetically less

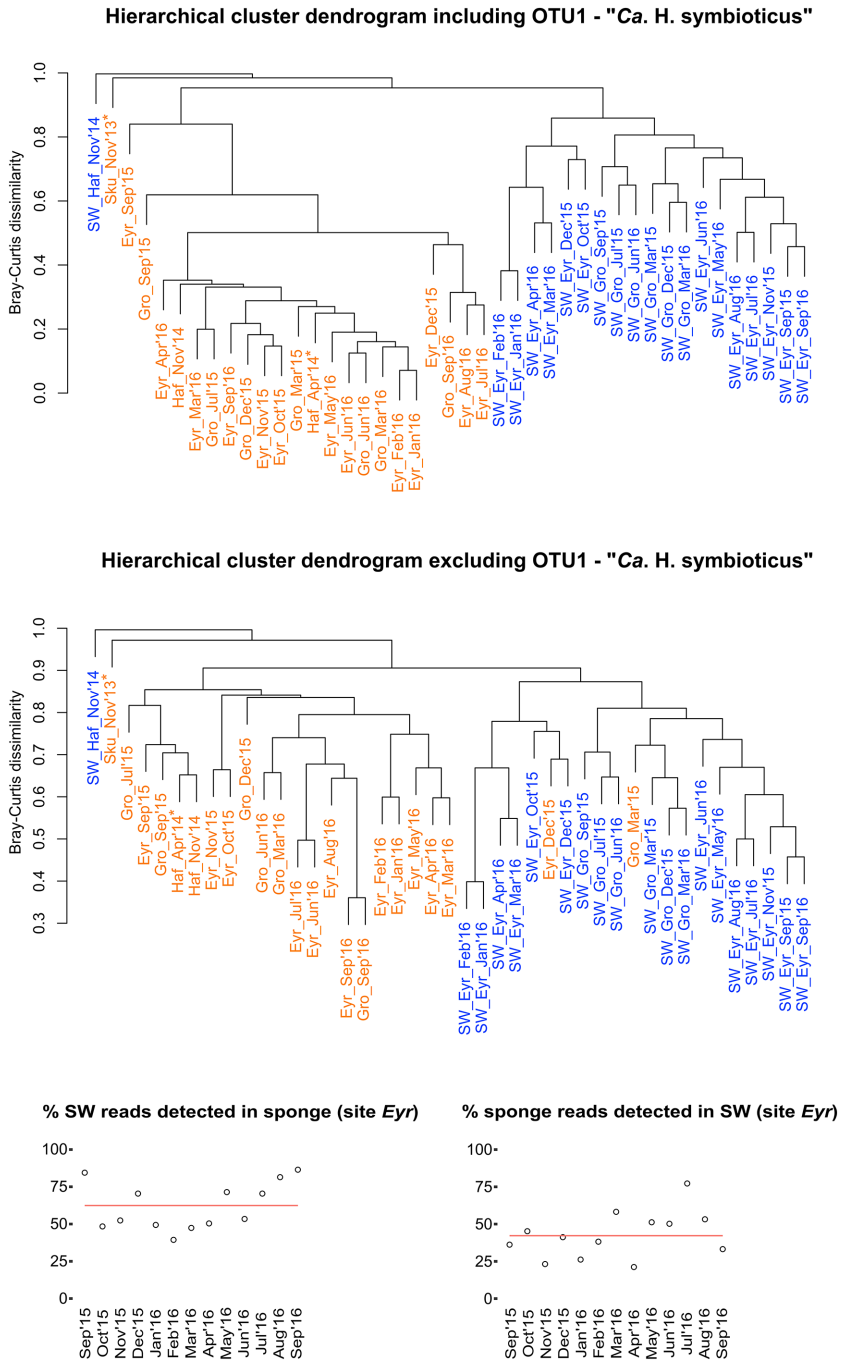


Figure 2. Top and middle: Hierarchical cluster dendrogram of sponge sample sets and seawater samples based on the Bray–Curtis index of dissimilarity including and excluding OTU1–'*Ca.* H. symbioticus'. Each sponge sample set is composed of samples from three sponge individuals (apart from Sku_Nov'13 which is composed of six). Sponge samples are marked orange and seawater sample are marked blue. Sample sets marked with an asterisk are from subtidal habitats. Bottom: Percentage of 16S rRNA gene sequence reads from one sample type detected in the other at site Eyr with average abundance (red line) across all samples. The dominant bacterium OTU1–'*Ca.* H. symbioticus' was removed from the dataset and the remaining read abundance adjusted to 100%.

related to the population in North Iceland and closer related to populations in the North Sea based on single nucleotide variations (Fig. S4, Supporting Information).

Phylogenetic analysis of the partial 16S rRNA gene to closely related cultivated and uncultivated bacteria indicated that the dominant *Alphaproteobacteria* in *H. panicea* from all available studies fall into a monophyletic group with OTU1 from this study, which is clearly distinguished from other closely related bacteria (Fig. 6). In addition, OTU1 from North Iceland represented a separate lineage within the monophyletic group with strong bootstrap support. The closest related, cultured bacteria to OTU1 were members of the genera *Amylibacter*, *Celeribacter*, *Confluentimicrobium*, *Thioclava*, *Paenirhodobacter* and *Pseudorhodobacter*, all belonging to the family *Rhodobacteraceae*. OTU1 had less than 94% sequence similarity to the closest related members of these genera indicating that it forms a novel genus within the family of *Rhodobacteraceae*.

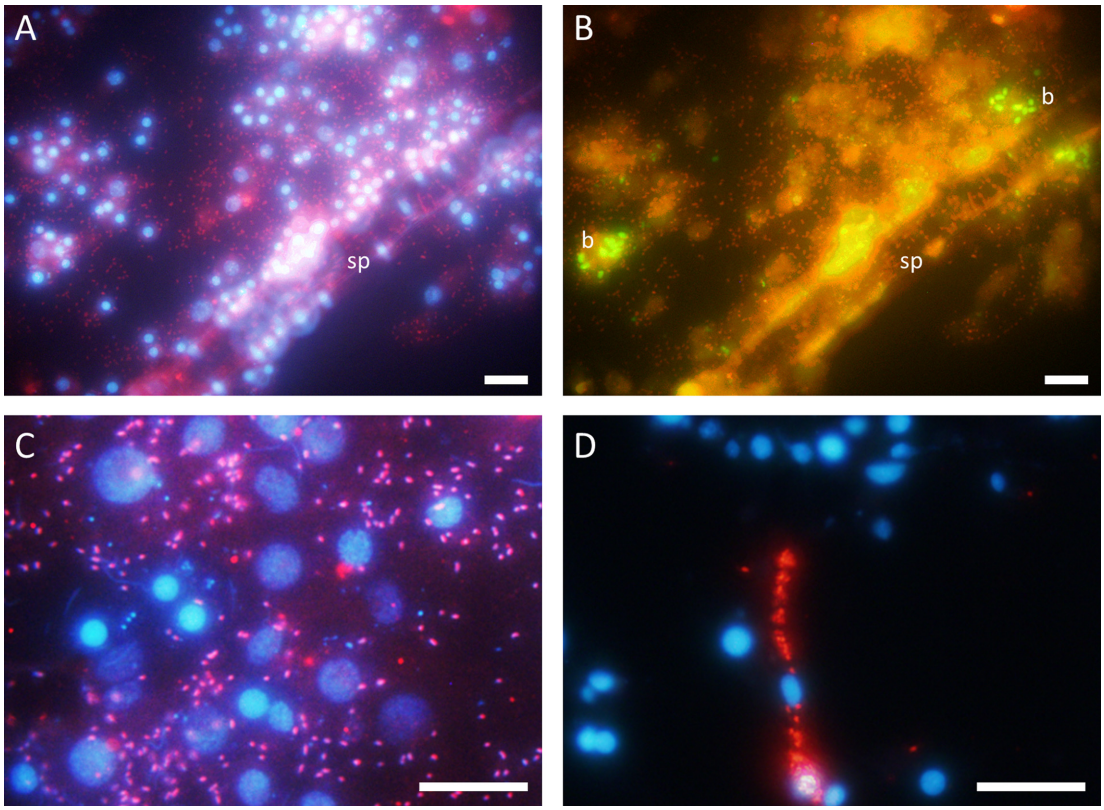


Figure 3. 16S rRNA FISH images of wild (A-C) and *ex situ* cultivated (D) *H. panicea* sections; (A), (C) and (D): Sponge cell nuclei are labelled with DAPI (blue) surrounded by OTU1–'*Ca.* H. symbioticus' labelled with Cy3 (red) and DAPI; DAPI signals in image D are weaker due to longer storage time of the sample; (B): Bacteria labelled with Alexa488 (green) and OTU1–'*Ca.* H. symbioticus' labelled with Cy3 and Alexa488; b: Bacteria labelled with universal bacterial probe and not with OTU1–'*Ca.* H. symbioticus' specific probe. sp: sponge spicule, bar = 10 μm.

Host specificity of the sponge symbiont

A comparison against the collection of sponge associated bacteria reported in Moitinho-Silva *et al.* (2017a) showed that *H. panicea* (samples 10533.PJS.SW.247, SW.245.1181828, SW.241.1181707 and SW.60.1181701 from Kosterfjorden, Sweden and sample 10533.MIIG0976 from Corranroo, Ireland) was the only sponge species in the collection that harbours a feature with 100% sequence similarity to the representative sequence of OTU1 at a high relative abundance of 66.0%–88.1% (Table 2, Supporting Information). It is necessary to mention that six sponges marked as *H. panicea* in the collection did not contain this or any closely related feature, although being from the same or similar sampling sites in the North Sea and North Atlantic. Additionally, 36 sponge samples belonging to other species than *H. panicea* and two seawater samples contained at least two reads of OTU1 at a relative abundance ranging between 0.7% and 0.0008% (Table 2, Supporting Information). These samples were collected from various geographical locations in the Atlantic, Pacific and Indian Ocean (Fig. 7). Four of these locations (New Zealand, Tanzania, Guam and Bonaire Island) are not reported to be inhabited by *H. panicea*, or contain possibly inaccurate reports of *H. panicea* according to the World Porifera database (Van Soest *et al.* 2017).

In addition, the partial 16S rRNA sequence of OTU1 was compared against a prokaryotic sequence database of seawater samples collected from 22 offshore stations around Iceland in May 2011–2013 using BLAST, in order to detect if OTU1 was present in offshore coastal waters around Iceland. No matching sequence was found with higher than 96% sequence similarity to the query sequence, indicating that OTU1 was either absent or present below the detection level in the water column of offshore seawater sites.

DISCUSSION

Halichondria panicea from Icelandic waters harbours a dominant and spatio-temporally stable bacterium: proposal of candidate status '*Candidatus* Halichondribacter symbioticus'

The presence of a dominant bacterial OTU reported here is in accordance with several studies on the bacterial community composition of *H. panicea* from the North Sea and Baltic Sea reporting the presence of a dominant *Alphaproteobacterium* albeit varying relative abundances of other bacterial classes between studies (Althoff *et al.* 1998; Wichels *et al.* 2006; Naim *et al.* 2014; Steinert *et al.* 2017). High sequence similarity of the dominant bacteria between these studies provides evidence that *H. panicea* can harbour the same bacterial species across different geographical locations, now extended to the Arctic Ocean, and supports the concept of a symbiotic sponge-bacterium relationship first proposed by Althoff *et al.* (1998). Using a deep-sequencing approach, we could show that the dominant OTU1 was the only bacterium present in all *H. panicea* samples from all geographical locations and sampling times collected in this study. The results from 16S rRNA FISH imaging of sponge sections further showed that the mesohyl was almost exclusively populated by this dominant bacterium.

Despite a low core sponge bacterial community composed of the dominant OTU1, the bacterial community detected in

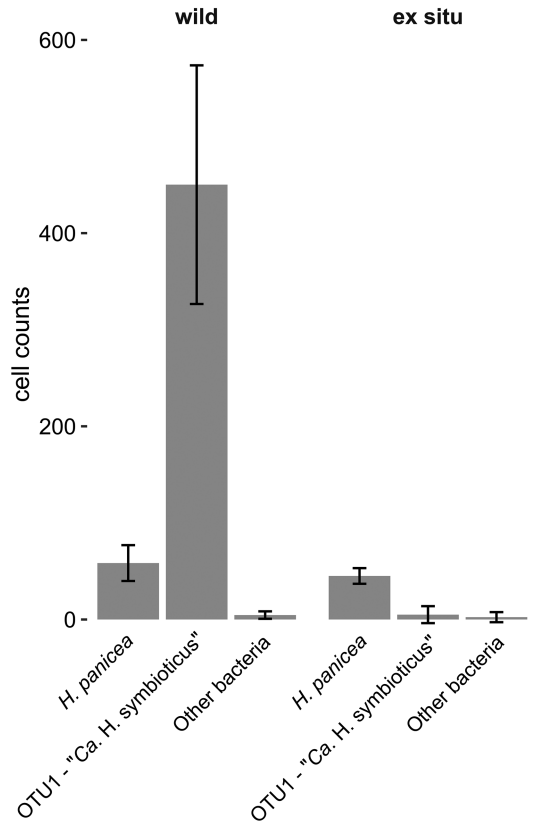


Figure 4. Cell counts of sponge sections from wild *H. panicea* (wild) and *H. panicea* after six months in a seawater aquarium (*ex situ*). Counts are based on 16S rRNA FISH images and are given as mean with standard deviation as error bar.

the sponge body was dissimilar to the surrounding seawater community at the OTU level. This suggests that *H. panicea* harbours additional bacterial taxa, albeit temporarily, geographically or inter-individually varying, that are associated with the sponge. Interestingly, two sponge sample sets, Eyr_Dec'15 and Gro_Mar'15, were more similar to their respective seawater samples than to other sponge samples based on hierarchical clustering of Bray–Curtis dissimilarity and after excluding OTU1 from the dataset. This could point towards periodic removal of the sponge bacterial community, possibly related to tissue sloughing (Barthel and Wolfrath 1989) or other cleaning mechanisms of the sponge to prevent fouling. However, concluding sponge association based on surrounding seawater samples alone remains difficult, as the sponge bacterial community could be influenced by other factors such as patterns of pumping activity (Kumala, Riisgård and Canfield 2017), or short term changes of the microbial diversity due to water currents or tidal flows, as for example shown for the coral *Coelastrea aspera* (Sweet *et al.* 2017). Compared against other sponge species analysed as part of the global Porifera microbiome survey, the microbiome of three *H. panicea* from Sweden, had a low inverse Simpson's index of diversity (*D*) (Thomas *et al.* 2016), likely due to the high relative abundance of the dominant symbiont sequence in these samples. It is hypothesised that microbial communities with a low *D* might be structured by host regulation or by the presence of competitively dominant symbionts (Easson and Thacker 2014). This is likely true for the *H. panicea*—OTU1 system, which might limit the permanent membership of other bacterial taxa to their community.

The dominant bacterium OTU1 appeared as a small, ca. 1 μm long and rod-shaped cell in 16S rRNA FISH images. Other bacteria were less abundant and frequently found in small clusters of adjacent bacterial cells, possibly resembling bacteria captured by choanocytes and transported to the mesohyl for phagocytosis (Wilkinson, Garrone and Vacelet 1984). *Halichondria panicea* has previously been classified as an LMA sponge based on TEM images by Moitinho-Silva *et al.* (2017b) and the low sponge-specific bacterial diversity combined with the bacteria to sponge cell ratio found in this study confirms this classification. However, the high abundance of OTU1 in the sponge mesohyl based on 16S rRNA FISH images is contrasting to the authors' observation of the mesohyl being largely devoid of microbial cells (Moitinho-Silva *et al.* 2017b), showing that *H. panicea* also exist without or with much lower abundances of this symbiont.

LMA sponges often host single dominant and sponge species specific bacterial taxa (Giles *et al.* 2013; Erwin *et al.* 2015). *Crambe crambe*, for instance, hosts a dominant *Betaproteobacterium* independent of habitat or exposure to pollutants (Croué *et al.* 2013; Gantt, López-Legentil and Erwin 2017). It is, however, not yet clear what role, mutualistic or otherwise, these single dominant bacterial symbionts play within the sponge holobiont. Apart from their role in host nutrition (Taylor *et al.* 2007b), some sponge symbionts have been associated with the production of cytotoxic compounds which might deter potential microbial competitors and thus help protect their host from pathogens (Wang 2006). Although cytotoxic properties have been found in extracts of *H. panicea* (Althoff *et al.* 1998; Ferreira *et al.* 2011), it is not yet clear if they originate from the sponge or its symbiont. Regarding the high abundance of OTU1 detected in the sponge body, it is possible that the symbiont simply out-competes other microorganism entering the sponge mesohyl thus protecting the sponge from disease (Selvin *et al.* 2010).

Due to its high abundance and close association to the sponge *H. panicea*, as well as its phylogenetic distinction, a candidate status is proposed for OTU1 and all members of its monophyletic clade, with the name '*Candidatus* Halichondribacter symbioticus' (Ha.li.chon.dri.bac'ter. N.L. fem. n. *Halichondria*, taxonomic name of a sponge genus, pertaining to source of the bacteria; L. masc. n. *bacter*, rod or staff; N.L. masc. n. *Halichondribacter*, bacteria found within a *Halichondria* species; sym.bio'ti.cus. Gr. prep. *syn.* together; Gr. n. *bios*, life; N.L. masc. adj. *symbioticus*, symbiotic).

'*Ca.* H. symbioticus' abundance decreases during *ex situ* cultivation but remains present within the sponge mesohyl

Exposure to *ex situ* environmental conditions in aquarium systems can alter the microbial community composition of sponges (Schippers *et al.* 2012). Some sponge species undergo large changes in their microbial community, whereas in others the sponge-associated microorganisms are largely retained (Webster and Taylor 2012). Here, we show that transfer of *H. panicea* to a seawater aquarium with low feed availability for six months led to a strong decrease in the overall bacterial abundance found within the sponge mesohyl. Nevertheless, '*Ca.* H. symbioticus' remained the dominant bacterium detected through 16S rRNA FISH after a six-month period, suggesting a certain stability of this sponge-bacterium association.

A high disparity between the relative abundance of '*Ca.* H. symbioticus' measured through 16S rRNA gene amplicon sequencing and 16S rRNA FISH was detected in all examined

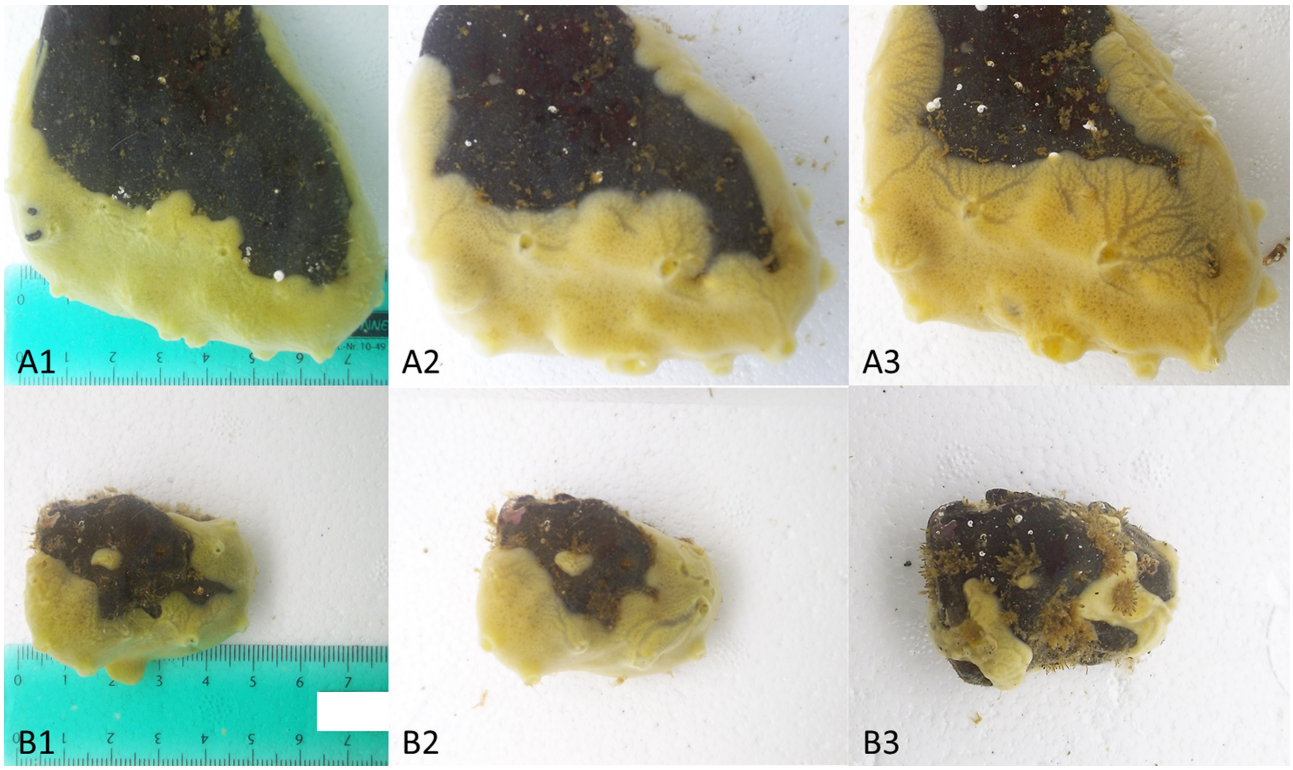


Figure 5. Images of two *H. panicea* individuals exhibiting an increase (A) or decrease (B) of the attachment area to the substrate within the six months of the *ex situ* cultivation trial. Images were taken after collection of sponges in the wild (1), after two months (2) and after six months (3) of *ex situ* cultivation in a recirculating seawater tank. Scale is given in cm.

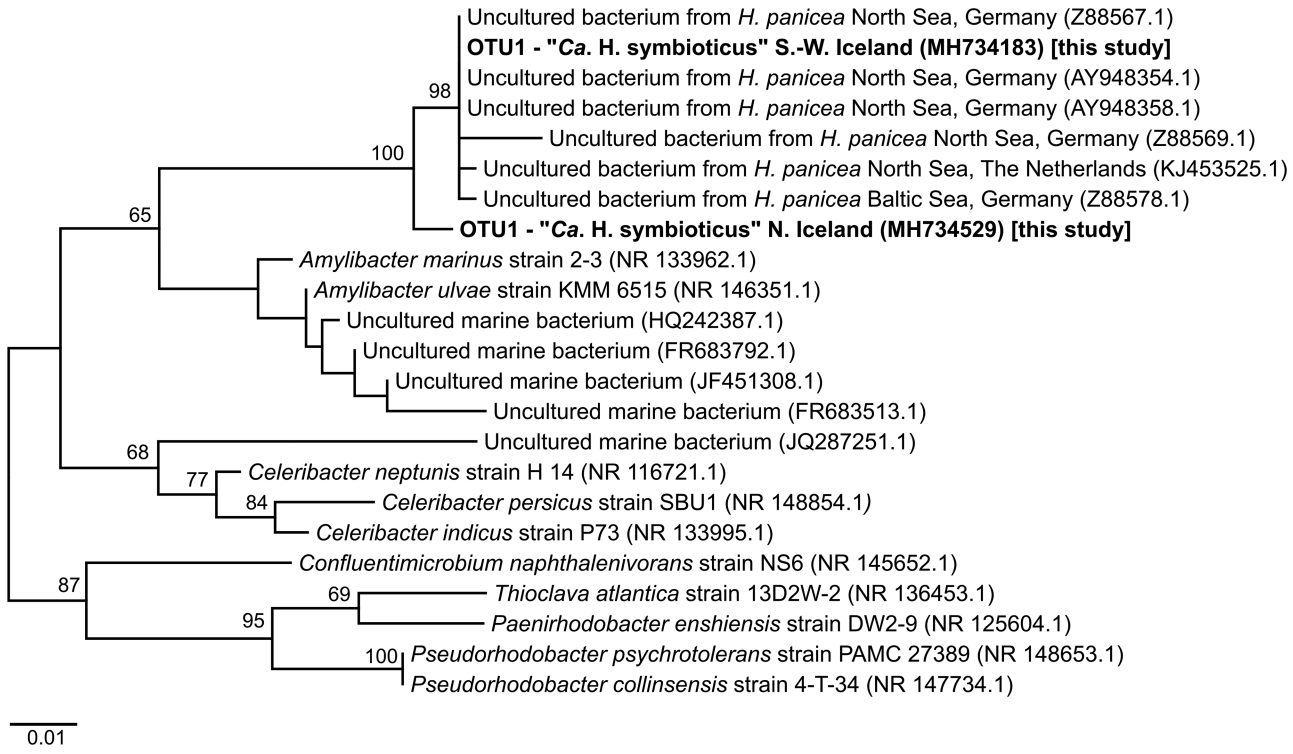


Figure 6. Maximum Likelihood tree of OTU1–'*Ca.* H. symbioticus' and closely related cultivated and uncultivated bacterial partial 16S rRNA gene sequences. Bootstrap values above 50% and based on 1000 resamplings are shown.

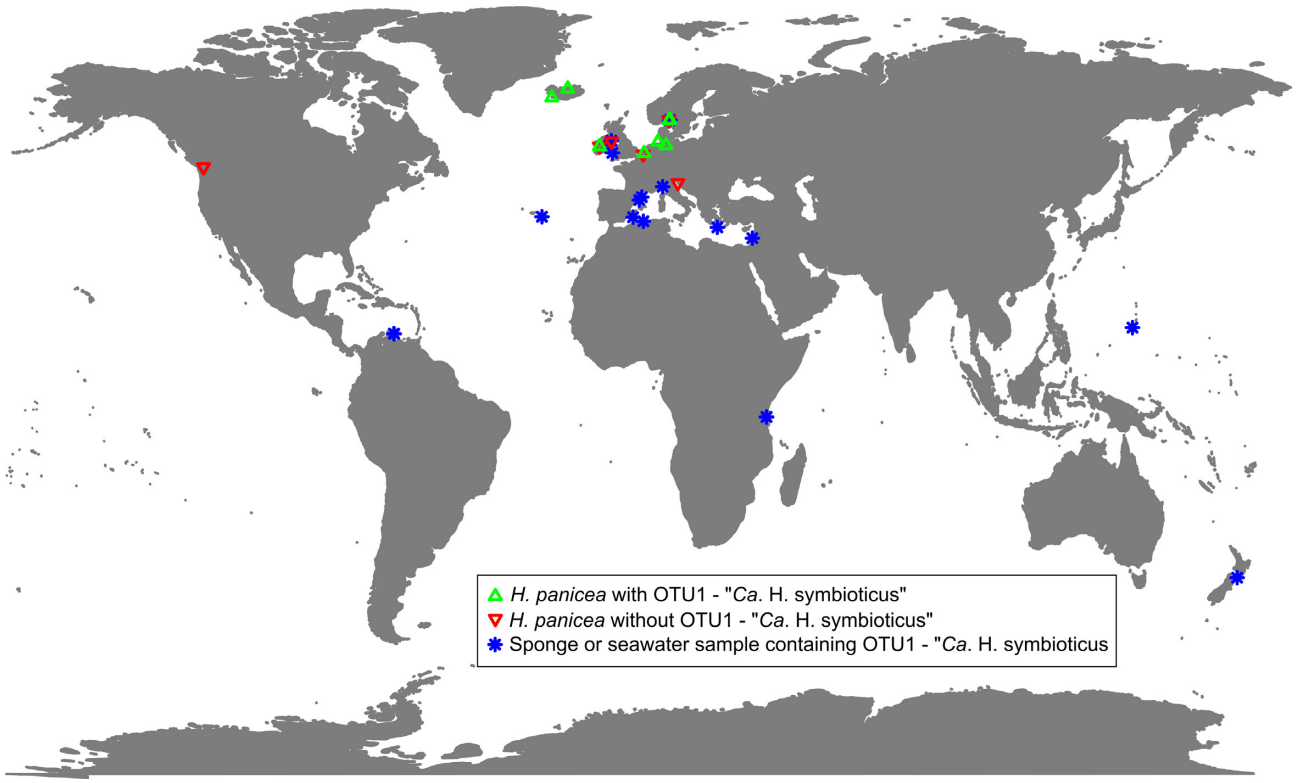


Figure 7. Global map showing distribution of *H. panicea* containing OTU1–'*Ca.* H. symbioticus' as the dominant bacterium (green triangle), *H. panicea* reported without detection of OTU1–'*Ca.* H. symbioticus' (red triangle), and other sponge species and seawater samples containing OTU1–'*Ca.* H. symbioticus' at low relative abundance (blue star). For detailed information see Table 2 (Supporting Information).

sponges, whereas this observation was more pronounced in the cultivated specimens where overall abundance of '*Ca.* H. symbioticus' was lower. This suggests that '*Ca.* H. symbioticus' was not sufficiently recovered during one of the steps leading to 16S rRNA gene amplicon sequencing, a possible result of insufficient DNA extraction due to the presence of extracellular matrices (Lee *et al.* 2003) or other inhibiting factors. Also, this could point towards an increased detection of bacteria present in the aquiferous system or on the sponge surface after sample collection (Schippers *et al.* 2012), which were likely removed by frequent washing steps during preparation for histology and FISH.

The *ex situ* cultivation experiment showed that the abundance of '*Ca.* H. symbioticus' in the sponge mesohyl decreased significantly due to environmental change. This could possibly be attributed to the low availability of particulate food and subsequently reduced host metabolism. Although seawater was supplied continuously from the sample collection site, a reduction or rapid exhaustion of dissolved nutrients in the aquarium water could alternatively have impacted the abundance of '*Ca.* H. symbioticus', yet this hypothesis has not been tested. It is intriguing to consider that '*Ca.* H. symbioticus' possess a mechanism to remain within *H. panicea* when faced with a less favourable environment, while thriving under normal conditions. Such mechanism could potentially be related to the vertical transmission of the bacterial symbiont during the larval stage of the sponge. Bacteria have previously been detected in larvae of *H. panicea*, a possible indicator for the vertical transmission of symbionts (Wichels *et al.* 2006). The cultivation experiment further showed that, although '*Ca.* H. symbioticus' is highly abundant in wild *H. panicea*, a high abundance is not necessary for restructuring of the sponge aquiferous system, increasing its attachment area, or for its survival over several months.

Global distribution and sequence variants of '*Ca.* H. symbioticus'

Analysing ribosomal sequence variants increases the sensitivity and specificity of marker-gene analysis and allows for a better discrimination of ecological patterns in bacterial communities (Acinas *et al.* 2004; Callahan, McMurdie and Holmes 2017). By analysing ribosomal nucleotide variations Eren *et al.* (2014), for instance, were able to show a strong differential distribution of two variants of a dominant *Gammaproteobacteria* found in two cryptic species of the deep-sea sponge *Hexadella*, allowing for a more detailed characterisation of the sponge species-specific microbial communities. Ribosomal sequence variants of '*Ca.* H. symbioticus' indicate that there are geographically distinct subpopulations of the sponge symbiont, with '*Ca.* H. symbioticus' from sponge samples in the North Sea and South-West Iceland being genetically closer related than to '*Ca.* H. symbioticus' from North Iceland. Samples from both intertidal and subtidal *H. panicea* and from depths of less than 15 m were compared, increasing the likelihood that ribosomal sequence variants detected in this study are due to geographical differences, as opposed to specific subpopulations adapted to temporary air exposure or depth-specific adaptation.

H. panicea has been considered a cosmopolitan sponge species (Barthel and Wolfrath 1989), yet molecular analysis of the COI gene suggests that sponges identified as *H. panicea* from the North-East Pacific are genetically distant to the North-East Atlantic species complex of *H. panicea* and its sympatric sister

species *Halichondria bowerbanki* (Erpenbeck *et al.* 2004). In line with this finding, '*Ca.* H. symbioticus' has been detected in *H. panicea* from various location in the North-East Atlantic and now also in the Arctic Ocean, whereas *H. panicea* from the North-East Pacific neither showed a reoccurring band between samples using TRFLP fingerprinting (Lee, Wong and Qian 2009), nor did publicly available sequences retrieved from DGGE bands match the 16S rRNA gene of '*Ca.* H. symbioticus'. The public collection of sponge-associated bacteria reported in Moitinho-Silva *et al.* (2017a) also contained samples identified as *H. panicea* without sequences related to '*Ca.* H. symbioticus', which suggests that, even in similar locations, *H. panicea* exist with and without this bacterial symbiont. This raises the question if physiological or molecular differences between *H. panicea* with or without '*Ca.* H. symbioticus' exist, or if instead other microorganisms are involved in a similar symbiotic relationship with the sponge. Althoff *et al.* (1998), for instance, report that *H. panicea* from the Adriatic Sea hosts a dominant *Alphaproteobacterium*, however, clone library sequences show less than 93% sequence similarity to '*Ca.* H. symbioticus', implying a possible alternative symbiont in Mediterranean *H. panicea*. Further studies from different geographical locations and with molecular evidence of a uniform *H. panicea* lineage are, however, needed to address these questions in detail.

In addition to its presence in *H. panicea*, '*Ca.* H. symbioticus' can be detected in low abundances in other sponge species from a wide variety of geographical locations. Whereas in some cases this might be attributed to the presence of the symbiont in the seawater surrounding *H. panicea*, also observed in this study (Fig. 1), many of these locations are not reported to contain *H. panicea*, suggesting that '*Ca.* H. symbioticus' could be an opportunistic sponge-associated bacterium which thrives exceptionally well in certain *H. panicea*. This, in turn, could be linked to host selection for benefits derived from the symbiont and drive niche differentiation, not only between host species (Easson and Thacker 2014), but also at the sub-species level.

CONCLUSION

With the current study, we show that the marine sponge *H. panicea* from different geographical locations in Icelandic waters hosts a core dominant bacterium, for which the candidate status '*Candidatus* Halichondribacter symbioticus' is proposed. Under *ex situ* conditions with low food availability for the sponge the absolute abundance of '*Ca.* H. symbioticus' decreased drastically, possibly due to a dependence on its host metabolism. Despite this reduction, '*Ca.* H. symbioticus' remained present within the sponge mesohyl for six months, pointing towards the presence of a mechanism that maintains the symbiotic relationship even under less favourable environmental conditions. Ribosomal sequence variants of '*Ca.* H. symbioticus' show that subpopulations of the symbiont exist between different geographical locations which could be used to unravel fine-scale evolutionary processes in the *H. panicea*—'*Ca.* H. symbioticus' relationship. However, sponge-species specificity of the sponge symbiont remains to be elucidated, as '*Ca.* H. symbioticus' can be found, at low abundance, in other sponge species from various geographical locations, while it is either absent or the dominant bacterium in *H. panicea* from similar sites in the North Atlantic. Apart from elucidating the ecological role of '*Ca.* H. symbioticus' within its host, future studies should therefore focus on its wider distribution and on the potential genetic and functional divergence of *H. panicea* with and without this symbiont.

SUPPLEMENTARY DATA

Supplementary data are available at *FEMSEC* online.

ACKNOWLEDGMENTS

We thank Eydís Einarsdóttir for providing sponge samples from North Iceland, Detmer Sipkema and Ronald Osinga for valuable discussion, Bernhard Schink for support with the nomenclature, and Ásgeir Guðnason for providing facilities and help with the cultivation experiments.

FUNDING

This work was supported by the EC Marie Curie Actions ITN Project BluePharmTrain [No. 607786].

Conflicts of interest. None declared.

REFERENCES

Acinas SG, Klepac-Ceraj V, Hunt DE *et al.* Fine-scale phylogenetic architecture of a complex bacterial community. *Nature* 2004;**430**:551–4.

Ackers RG, Moss D, Picton BE. *Sponges of the British Isles (" Sponge V"): A Colour Guide and Working Document, 1992 Edition, Reset with Modifications*, Ross-on-Wye, UK. Marine Conservation Society, 2007.

Althoff K, Schütt C, Steffen R *et al.* Evidence for a symbiosis between bacteria of the genus Rhodobacter and the marine sponge Halichondria panicea: Harbor also for putatively toxic bacteria? *Mar Biol* 1998;**130**:529–36.

Amann RI, Binder BJ, Olson RJ *et al.* Combination of 16S rRNA-targeted oligonucleotide probes with flow cytometry for analyzing mixed microbial populations. *Appl. Environ. Microbiol.* 1990; **56:** 1919–25

Baker GC, Smith JJ, Cowan DA. Review and re-analysis of domain-specific 16S primers. *J Microbiol Methods* 2003;**55**:541–55.

Barthel D, Wolfrath B. Tissue sloughing in the sponge Halichondria panicea: A fouling organism prevents being fouled. *Oecologia* 1989;**78**:357–60.

Barthel D. On the ecophysiology of the sponge Halichondria panicea in Kiel Bight. I. Substrate specificity, growth and reproduction. *Mar Ecol Prog Ser* 1986;**32**:291–8.

Becker RA, Wilks AR, Brownrigg R *et al.* maps: Draw Geographical Maps, R package version 3.3.0 https://CRAN.R-project.org/package=maps 2018.

Callahan BJ, McMurdie PJ, Holmes SP. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *Isme J* 2017;**11**:2639.

Callahan BJ, McMurdie PJ, Rosen MJ *et al.* DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Meth* 2016;**13**:581–3.

Camacho C, Coulouris G, Avagyan V *et al.* BLAST+: Architecture and applications. *BMC Bioinformatics* 2009;**10**:421.

Caporaso JG, Kuczynski J, Stombaugh J *et al.* QIIME allows analysis of high-throughput community sequencing data. *Nat Methods* 2010;**7**:335–6.

Chambers S. Systema porifera. In: Hooper JNA., van Soest RWM (eds). *A Guide to the Classification of Sponges*. Dordrecht: Kluwer Academic/Plenum Publishers , 2002, 1756pp. ISBN 0-306-47260-0. Aquat Conserv Mar Freshw Ecosyst 2003;**13**:461–2.

Croué J, West NJ, Escande M-L *et al.* A single betaproteobacterium dominates the microbial community of the crambescidine-containing sponge Crambe crambe. *Sci Rep* 2013;**3**:2583.

Daims H, Brühl A, Amann R *et al.* The domain-specific probe EUB338 is insufficient for the detection of all Bacteria: development and evaluation of a more comprehensive probe set. *Syst. Appl. Microbiol* 1999; **22**:434–44

Daims H, Lücker S, Wagner M. Daime, a novel image analysis program for microbial ecology and biofilm research. *Environ Microbiol* 2006;**8**:200–13.

De Goeij JM, Van Den Berg H, Van Oostveen MM *et al.* Major bulk dissolved organic carbon (DOC) removal by encrusting coral reef cavity sponges. *Mar Ecol Prog Ser* 2008;**357**:139–51.

Díez-Vives C, Moitinho-Silva L, Nielsen S *et al.* Expression of eukaryotic-like protein in the microbiome of sponges. *Mol Ecol* 2017;**26**:1432–51.

Easson CG, Thacker RW. Phylogenetic signal in the community structure of host-specific microbiomes of tropical marine sponges. *Front Microbiol* 2014;**5**:532.

Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;**32**:1792–7.

Edgar RC. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 2010;**26**:2460–1.

Eren AM, Morrison HG, Lescault PJ *et al.* Minimum entropy decomposition: Unsupervised oligotyping for sensitive partitioning of high-throughput marker gene sequences. *Isme J* 2014;**9**:968.

Erpenbeck DIRK, Knowlton AL, Talbot SL *et al.* A molecular comparison of Alaskan and Northeast Atlantic Halichondria panicea (Pallas 1766) (Porifera: Demospongiae) populations. *Boll Mus Ist Univ Genova* 2004;**61**:319–25.

Erwin PM, Coma R, López-Sendino P *et al.* Stable symbionts across the HMA-LMA dichotomy: Low seasonal and interannual variation in sponge-associated bacteria from taxonomically diverse hosts. *FEMS Microbiol Ecol* 2015;**91**:fiv115–.

Erwin PM, Thacker RW. Incidence and identity of photosynthetic symbionts in Caribbean coral reef sponge assemblages. *J Mar Biol Assoc UK* 2007;**87**:1683–92.

Felsenstein J. Confidence Limits on Phylogenies: An Approach Using the Bootstrap. *Evolution (N Y)* 1985;**39**:783–91.

Felsenstein J. Evolutionary trees from DNA sequences: Amaximum likelihood approach. *J Mol Evol* 1981;**17**:368–76.

Ferreira M, Cabado AG, Chapela M-J *et al.* Cytotoxic activity of extracts of marine sponges from NW Spain on a neuroblastoma cell line. *Environ Toxicol Pharmacol* 2011;**32**:430–7.

Feuda R, Dohrmann M, Pett W *et al.* Improved modeling of compositional heterogeneity supports sponges as sister to all other animals. *Curr Biol* 2017;**27**:3864–70. e4.a

Flatt PM, Gautschi JT, Thacker RW *et al.* Identification of the cellular site of polychlorinated peptide biosynthesis in the marine sponge Dysidea (Lamellodysidea) herbacea and symbiotic cyanobacterium Oscillatoria spongeliae by CARD-FISH analysis. *Mar Biol* 2005;**147**:761–74.

Freeman CJ, Thacker RW. Complex interactions between marine sponges and their symbiotic microbial communities. *Limnol Oceanogr* 2011;**56**:1577–86.

Gantt SE, López-Legentil S, Erwin PM. Stable microbial communities in the sponge Crambe crambe from inside and outside a polluted Mediterranean harbor. *FEMS Microbiol Lett* 2017;**364**:fnx105.

Giles EC, Kamke J, Moitinho-Silva L *et al.* Bacterial community profiles in low microbial abundance sponges. *FEMS Microbiol Ecol* 2013;**83**:232–41.

Gloeckner V, Wehrl M, Moitinho-Silva L *et al.* The HMA-LMA dichotomy revisited: An electron microscopical survey of 56 sponge species. *Biol Bull* 2014;**227**:78–88.

Hentschel U, Hopke J, Horn M *et al.* Molecular Evidence for a Uniform Microbial Community in Sponges from Different Oceans Molecular Evidence for a Uniform Microbial Community in Sponges from Different Oceans. *Appl Environ Microbiol* 2002;**68**:4431–40.

Hentschel U, Usher KM, Taylor MW. Marine sponges as microbial fermenters. *FEMS Microbiol Ecol* 2006;**55**:167–77.

Hofmann K, Baron MD. *Boxshade 3.21. Pretty Printing and Shading of Multiple-Alignment Files*. Lausanne: Kay Hofmann ISREC Bioinformatics Group, 1996.

Imhoff JF, Stöhr R. *Sponge-Associated Bacteria: General Overview and Special Aspects of Bacteria Associated with Halichondria panicea. Sponges (Porifera)*. Berlin: Springer, 2003, 35–57.

Kamke J, Sczyrba A, Ivanova N *et al.* Single-cell genomics reveals complex carbohydrate degradation patterns in poribacterial symbionts of marine sponges. *ISME J* 2013;**7**:2287–300.

Keren R, Mayzel B, Lavy A *et al.* Sponge-associated bacteria mineralize arsenic and barium on intracellular vesicles. *Nat Commun* 2017;**8**:14393.

Klindworth A, Pruesse E, Schweer T *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res* 2013;**41**:e1–e1.

Kumala L, Riisgård HU, Canfield DE. Osculum dynamics and filtration activity studied in small single-osculum explants of the demosponge Halichondria panicea. *Mar Ecol Prog Ser* 2017;**572**:117–28.

Kumar S, Stecher G, Tamura K. MEGA7 : Molecular Evolutionary Genetics Analysis version 7. 0 for bigger datasets. *Mol Biol Evol* 2016;**33**:1–11.

Lafi FF, Fuerst JA, Fieseler L *et al.* Widespread distribution of poribacteria in demospongiae. *Appl Environ Microbiol* 2009;**75**:5695–9.

Lee OO, Wong YH, Qian PY. Inter- and intraspecific variations of bacterial communities associated with marine sponges from San Juan Island, Washington. *Appl Environ Microbiol* 2009;**75**:3513–21.

Lee Y., Kim H., Liu C. *et al.* A simple method for DNA extraction from marine bacteria that produce extracellular materials. *J Microbiol Methods* 2003;**52**:245–50.

Ludwig W, Strunk O, Westram R *et al.* ARB: A software environment for sequence data. *Nucleic Acids Res* 2004;**32**:1363–71.

McMurdie PJ, Holmes S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One* 2013;**8**:e61217.

Mills DB, Ward LM, Jones C *et al.* Oxygen requirements of the earliest animals. *Proc Natl Acad Sci* 2014;**111**:4168–72.

Moitinho-Silva L, Nielsen S, Amir A *et al.* The sponge microbiome project. *Gigascience* 2017a;**6**:1–7.

Moitinho-Silva L, Steinert G, Nielsen S *et al.* Predicting the HMA-LMA Status in Marine Sponges by Machine Learning. *Front Microbiol* 2017b;**8**:752.

Naim MA, Morillo JA, Sørensen SJ *et al.* Host-specific microbial communities in three sympatric North Sea sponges. *FEMS Microbiol Ecol* 2014;**90**:390–403.

NCBI Resource Coordinators. Database Resources of the National Center for Biotechnology Information. *Nucleic Acids Res* 2017;**45**:D12–7.

Oksanen J, Blanchet FG, Friendly M *et al.* vegan: Community Ecology Package. R package version 2.4-4. https ://CRAN.R-project.org/package=vegan 2017.

Pernthaler J, Glöckner F-O, Schönhuber W *et al*. Fluorescence in situ hybridization (FISH) with rRNA-targeted oligonucleotide probes. *Methods in Microbiology* **30;** 2001: 207–26.

Porter KG, Feig YS. The use of DAPI for identifying and counting aquatic microflora1. *Limnol Oceanogr* 1980;**25**:943–8.

Quast C, Pruesse E, Yilmaz P *et al*. The SILVA ribosomal RNA gene database project: Improved data processing and web-based tools. *Nucleic Acids Res* 2013;**41**:D590–6.

R Core Team. *A language and environment for statistical computing*, Vienna: R Foundation for Statistical Computing, 2017.

Reynisson E, Lauzon HL, Magnússon H *et al*. Bacterial composition and succession during storage of North-Atlantic cod (Gadus morhua) at superchilled temperatures. *BMC Microbiol* 2009;**9**:250.

Riisgård HU, Kumala L, Charitonidou K. Using the F/R-ratio for an evaluation of the ability of the demosponge Halichondria panicea to nourish solely on phytoplankton versus free-living bacteria in the sea. *Mar Biol Res* 2016;**12**:907–16.

Riisgård HU, Larsen PS. Filter-feeding in marine macro-invertebrates: Pump characteristics, modelling and energy cost. *Biol Rev* 1995;**70**:67–106.

Schippers KJ, Sipkema D, Osinga R *et al*. Cultivation of Sponges, Sponge Cells and Symbionts. Achievements and Future Prospects. *Adv Mar Biol* 2012;**62**:273–337.

Schläppy M-L, Schöttner SI, Lavik G *et al*. Evidence of nitrification and denitrification in high and low microbial abundance sponges. *Mar Biol* 2010;**157**:593–602.

Schmitt S, Hentschel U, Taylor MW. Deep sequencing reveals diversity and community structure of complex microbiota in five Mediterranean sponges. *Hydrobiologia* 2012;**687**:341–51.

Schneemann I, Nagel K, Kajahn I *et al*. Comprehensive investigation of marine actinobacteria associated with the sponge halichondria panicea. *Appl Environ Microbiol* 2010;**76**:3702–14.

Selvin J, Ninawe AS, Seghal Kiran G *et al*. Sponge-microbial interactions: Ecological implications and bioprospecting avenues. *Crit Rev Microbiol* 2010;**36**:82–90.

Siegl A, Kamke J, Hochmuth T *et al*. Single-cell genomics reveals the lifestyle of Poribacteria, a candidate phylum symbiotically associated with marine sponges. *ISME J* 2011;**5**:61–70.

Simister RL, Deines P, Botté ES *et al*. Sponge-specific clusters revisited: A comprehensive phylogeny of sponge-associated microorganisms. *Environ Microbiol* 2012;**14**:517–24.

Slaby BM, Hackl T, Horn H *et al*. Metagenomic binning of a marine sponge microbiome reveals unity in defense but metabolic specialization. *ISME J* 2017;**11**:2465–78.

Steinert G, Rohde S, Janussen D *et al*. Host-specific assembly of sponge-associated prokaryotes at high taxonomic ranks. *Sci Rep* 2017;**7**:2542.

Sweet MJ, Brown BE, Dunne RP *et al*. Evidence for rapid, tide-related shifts in the microbiome of the coral Coelastrea aspera. *Coral Reefs* 2017;**36**:815–28.

Tamura K, Nei M. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Mol Biol Evol* 1993;**10**:512–26.

Taylor MW, Hill RT, Piel J *et al*. Soaking it up: The complex lives of marine sponges and their microbial associates. *ISME J* 2007a;**1**:187–90.

Taylor MW, Radax R, Steger D *et al*. Sponge-Associated Microorganisms: Evolution, Ecology, and Biotechnological Potential. *Microbiol Mol Biol Rev* 2007b;**71**:295–347.

Taylor MW, Tsai P, Simister RL *et al*. 'Sponge-specific' bacteria are widespread (but rare) in diverse marine environments. *Isme J* 2012;**7**:438.

Thomassen S, Riisgård HU. Growth and energetics of the sponge Halichondria panicea. *Mar Ecol Prog Ser* 1995;**128**:239–46.

Thomas T, Moitinho-Silva L, Lurgi M *et al*. Diversity, structure and convergent evolution of the global sponge microbiome. *Nat Commun* 2016;**7**:11870.

Thomas T, Rusch D, DeMaere MZ *et al*. Functional genomic signatures of sponge bacteria reveal unique and shared features of symbiosis. *ISME J* 2010;**4**:1557–67.

Tian R-M, Sun J, Cai L *et al*. The deep-sea glass sponge Lophophysema eversa harbours potential symbionts responsible for the nutrient conversions of carbon, nitrogen and sulfur. *Environ Microbiol* 2016;**18**:2481–94.

Van Soest RWM, Boury-Esnault N, Hooper JNA *et al*. World Porifera database. 2017; Accessed at http ://www.marinespecies.org/porifera on 2018-10-05.

Van Soest RWM, Boury-Esnault N, Vacelet J *et al*. Global Diversity of Sponges (Porifera). *PLoS One* 2012;**7**:e35105.

Wang G. Diversity and biotechnological potential of the sponge-associated microbial consortia. *J Ind Microbiol Biotechnol* 2006;**33**:545–51.

Webster NS, Taylor MW, Behnam F *et al*. Deep sequencing reveals exceptional diversity and modes of transmission for bacterial sponge symbionts. *Environ Microbiol* 2010;**12**:2070–82.

Webster NS, Taylor MW. Marine sponges and their microbial symbionts: Love and other relationships. *Environ Microbiol* 2012;**14**:335–46.

Webster NS, Thomas T. The Sponge Hologenome. *MBio* 2016;**7**:e00135–16.

Wichels A, Würtz S, Döpke H *et al*. Bacterial diversity in the breadcrumb sponge Halichondria panicea (Pallas). *FEMS Microbiol Ecol* 2006;**56**:102–18.

Wickham H. *ggplot2: Elegant Graphics for Data Analysis*, New York: Springer-Verlag, 2009.

Wilkinson CR, Garrone R, Vacelet J. Marine Sponges Discriminate between Food Bacteria and Bacterial Symbionts: Electron Microscope Radioautography and in situ Evidence. *Proc R Soc London Ser B Biol Sci* 1984;**220**:519–28.

Wilkinson CR. Microbial associations in sponges. I. Ecology, physiology and microbial populations of coral reef sponges. *Mar Biol* 1978;**49**:161–7.

Witte U, Barthel D, Tendal O. The reproductive cycle of the sponge Halichondria panicea Pallas (1766) and its relationship to temperature and salinity. *J Exp Mar Bio Ecol* 1994;**183**:41–52.

Paper II

Genome analysis of sponge symbiont "*Candidatus* Halichondribacter symbioticus" shows genomic adaptation to a host-dependent lifestyle

Stephen Knobloch, Ragnar Jóhannsson, Viggó Þór Marteinsson

Submitted to *Environmental Microbiology*

Author contributions:

Conceived and designed the study: SK, RJ, VM
Performed sampling: SK
Performed laboratory work: SK
Analysed data: SK
Wrote the manuscript: SK
Reviewed and edited the manuscript: SK, RJ, VM
Correspondence (publication process): SK

Summary

The marine sponge *Halichondria panicea* inhabits coastal areas around the globe and is a widely studied sponge species in terms of its biology, yet the ecological functions of its dominant bacterial symbiont "*Candidatus* Halichondribacter symbioticus" remain unknown. Here, we present the draft genome of "*Ca.* H. symbioticus" HS1 (2.8 Mbp, ca. 93% genome recovery) recovered from the sponge metagenome of *H. panicea* in order to study functions and symbiotic interactions at the genome level. Functional genome comparison of HS1 against closely related free-living seawater bacteria revealed a reduction of genes associated with carbohydrate transport and transcription regulation, pointing towards a limited carbohydrate metabolism, and static transcriptional dynamics reminiscent of other bacterial symbionts. In addition, HS1 was enriched in sponge symbiont specific gene families related to host-symbiont interactions and defence. Similarity in the functional gene repertoire between HS1 and a phylogenetically more distant symbiont in the marine sponge *Aplysina aerophoba*, suggest a convergent evolution of symbiont specific traits and general metabolic features, warranting further investigation into convergent genomic evolution of symbionts across different sponge species and habitats.

Originality-significance statement

The current study reports the first draft genome of the dominant symbiont in a widely studied and ecologically important sponge species. Genomic analysis places the symbiont into the context of its host environment based on its functional attributes and draws comparisons to other obligate sponge symbiont genomes. The study describes the merits of the *H. panicea* – "*Ca*. H. symbioticus" model system for studying early sponge-microbe symbiosis and lays the groundwork for future studies exploring these interactions through hypothesis-driven, experimental research. This is both relevant for biotechnological exploitation of sponge-microbe derived secondary metabolites and for the wider study of animal-microbe interactions in general.

Introduction

Sponges (phylum *Porifera*) are considered one of the oldest animal lineages (Feuda *et al.*, 2017), estimated to have evolved over 600 million years ago (Yin *et al.*, 2015). Despite having one of the simplest animal body plans, they can be found in most aquatic habitats and hold important ecological functions, from creating diversity-enhancing sponge grounds in the deep sea (Beazley *et al.*, 2013) to retaining nutrients in coral reefs (de Goeij *et al.*, 2013). Their association with symbiotic microorganisms was suggested over 40 years ago through electron microscopy studies of several marine sponges (Vacelet and Donadey, 1977). An increase in modern sequencing technology has since dramatically expanded our understanding of their microbial diversity (Thomas *et al.*, 2016). Many sponges can host an enormously diverse and abundant community of microorganisms, in some cases exceeding several billion microbial cells per gram of sponge wet weight (Gloeckner *et al.*, 2014). Sponge-associated microorganisms can be highly sponge-specific (Hentschel *et al.*, 2002) and play important roles in their host, ranging from chemical defence and host nutrition to removal of contaminants or metabolic waste products (Flatt *et al.*, 2005; Erwin and Thacker, 2007; Siegl *et al.*, 2011). In recent years, the study of sponge symbionts has received increased interest due to the biotechnological potential of their secondary metabolites (Hentschel *et al.*, 2012). At the same time, sponge-microbe models are fertile ground for studying early animal-microbe symbiosis that can further our understanding of host-microbe interactions through conserved mechanisms (Pita *et al.*, 2016). Since obligate sponge symbionts have not yet been cultivated, genomic analysis is performed through metagenomic and single-cell genomic approaches. These have so far demonstrated both genetic similarities and metabolic specialisations for a wide range of symbionts from diverse sponge species and across different habitats (Webster and Thomas, 2016). Shared features between obligate sponge symbionts include the presence of eukaryotic-like proteins which can modulate host-symbiont interactions (Nguyen *et al.*, 2014; Díez-Vives *et al.*, 2017), enrichment of mobile genetic elements enabling transfer of functional genes between symbionts (Fan *et al.*, 2012) and enrichment of genes associated with restriction-modification systems related to bacterial defence (Slaby *et al.*, 2017). Specialisation of genetic features related to a sponge symbiotic lifestyle include secondary metabolite biosynthesis (Siegl and Hentschel, 2010), functional adaptations to host-dependent lifestyles (Liu *et al.*, 2011; Gao *et al.*, 2014; Burgsdorf *et al.*, 2015), or general metabolic functions such as sulphur oxidation and reduction (Hoffmann *et al.*, 2005; Gauthier *et al.*, 2016), CO2-fixation (Siegl *et al.*, 2011), urea utilisation (Hallam *et al.*, 2006) and ammonium oxidation (Fan *et al.*, 2012). To date, only a small number of all sponge symbionts have been studied to the genome level and despite current advances in this field many questions remain, for instance, related to the role of host immunity in mediating their microbial symbionts or to the symbionts' impact on enhancing the environmental adaptation of their host (Webster and Thomas, 2016).

The bread crumb sponge *Halichondria panicea* is a globally distributed sponge, inhabiting coastal areas (Barthel and Wolfrath, 1989) and occasionally dominating the intertidal range (Knowlton and Highsmith, 2000). However, the genetical homogeneity of *H. panicea* is not yet fully understood with indications of genetically separate populations between distinct geographical locations (Erpenbeck *et al.*, 2004). *H. panicea* has been subject to a wide range of studies on its biology and ecophysiology, both due to its importance in benthic communities (Peattie and Hoare, 1981; Barthel, 1986) and its amenability to laboratory cultivation (Barthel and Theede, 1986; Barthel, 1988; Kumala *et al.*, 2017). In a recent study

we showed that *H. panicea* from Icelandic waters hosts a dominant bacterium, "*Candidatus* Halichondribacter symbioticus", which was also the only bacterial taxon present across all seasons and from all studied geographical locations (Knobloch *et al.* 2018). Furthermore, "*Ca*. H. symbioticus" had a widespread distribution within the sponge mesohyl, being eight times more abundant than sponge cells within the body of wild specimens. However, "*Ca*. H. symbioticus" decreased in abundance when sponges were transferred to a seawater tank with limited particulate nutrient availability, possibly connected to a lower metabolic rate of its host. Apart from being dominant in several *H. panicea* studied in the North Atlantic (Althoff *et al.*, 1998; Wichels *et al.*, 2006; Naim *et al.*, 2014; Steinert *et al.*, 2017) other sponges identified as *H. panicea*, even from similar locations in the North Sea, do not appear to contain "*Ca*. H. symbioticus" at all. Moreover, partial 16S rRNA gene sequences of "*Ca*. H. symbioticus" have been detected in a range of other sponge species from locations in the Atlantic, Pacific and Indian Ocean, albeit at much lower relative abundance than in *H. panicea* (Moitinho-Silva *et al.*, 2017). It therefore appears that "*Ca*. H. symbioticus" is a general sponge symbiont with either a high specialisation for the *H. panicea* ecosystem or providing a trait highly desirable to the sponge.

Here we present the draft genome of "*Ca*. H. symbioticus" HS1 extracted form a *H. panicea* metagenome and compare it to the genomes of closely related free-living bacteria and sponge symbionts. We further perform ancestral genome reconstruction and analysis of a publicly available transcriptomic dataset in order to generate first insights into the symbiotic relationship between "*Ca*. H. symbioticus" and its main host *H. panicea*.

Results and Discussion

Sponge metagenome and draft genome of "*Ca.* Halichondribacter symbioticus" HS1

Metagenomic DNA was obtained from the bacterial fraction of two *H. panicea* individuals collected in South-West Iceland. In order to analyse the relative abundance of "*Ca*. H. symbioticus" in the metagenomes, 16S rRNA gene sequence fragments were extracted from the assembled reads and assigned a taxonomic classification. In MG1 and MG2, 86 and 79 % of all 16S rRNA gene sequences belonged to the family *Rhodobacteraceae* of the class *Alphaproteobacteria* and closely matched the 16S rRNA gene of "*Ca*. H. symbioticus" (Genbank: MH734183 and MH734529) with higher than 99.0 % sequence similarity, respectively (Figure 1). This is in accordance with previous studies in which "*Ca*. H. symbioticus" was the dominant bacterium in *H. panicea* (Althoff *et al.*, 1998; Wichels *et al.*, 2006; Naim *et al.*, 2014; Steinert *et al.*, 2017; Knobloch *et al.*, 2018) and highlights the large relative abundance of this symbiont compared to other bacteria found within the sponge.

Metagenomic reads were binned from MG1 based on genome coverage, GC-content and tetranucleotide frequency. The extracted draft genome of "*Ca*. H. symbioticus", henceforth referred to as HS1, was 2.8 Mbp in size with a N50 value of 2152, GC content of 49.4% and an average genome coverage of 131. In total, 3175 coding gene sequences were predicted, which fell within 1170 Clusters of Orthologous Groups of proteins (COGs) in 20 COG categories, and 1414 KEGG Orthologues (KOs) (Table 1). Based on the presence of 99 out of 106 single copy marker genes the completeness of the draft genome was estimated to be approximately 93%. Although the full genome of HS1 was not recovered from the

metagenome, key functional pathways, such as DNA replication, the citrate cycle (TCA cycle) and peptidoglycan biosynthesis were complete (Table 1), indicating that the HS1 draft genome was of sufficient quality to allow analysis of functional properties.

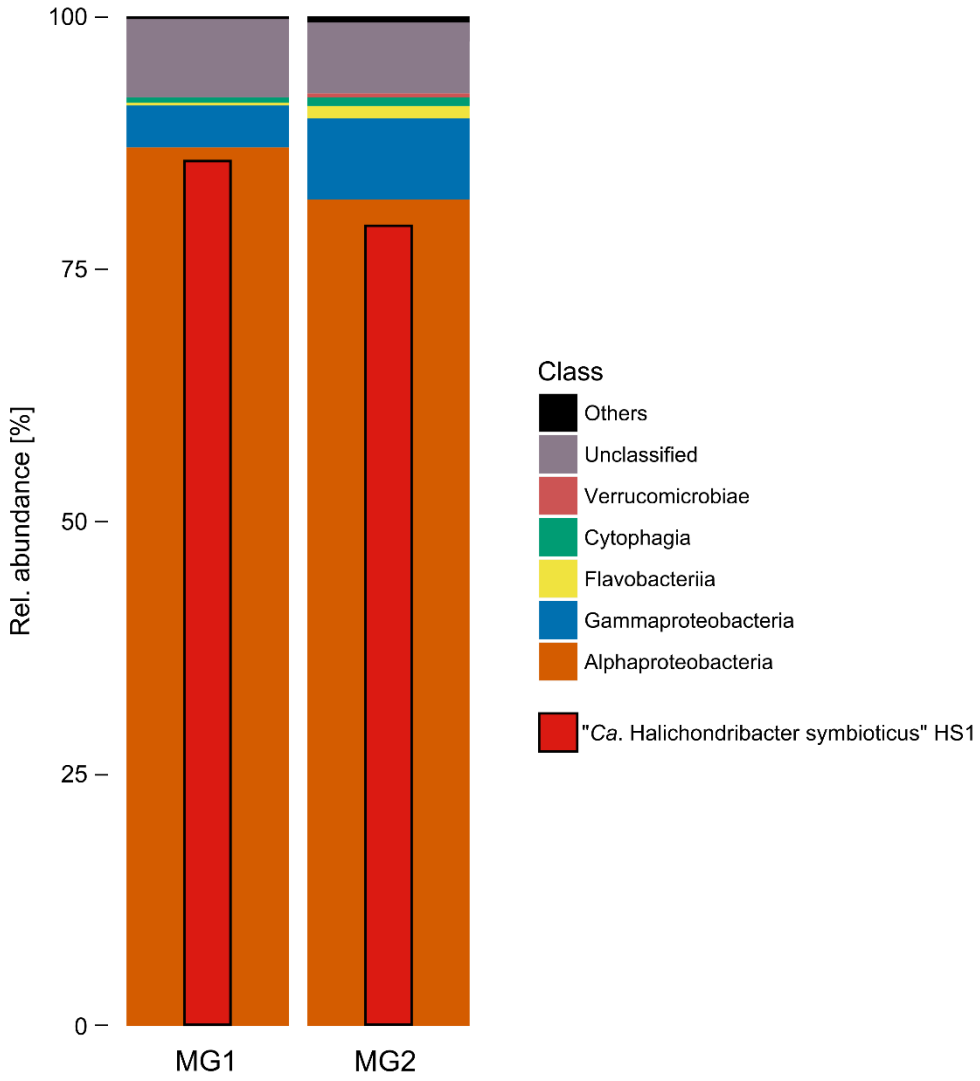


Figure 1: *Relative abundance of 16S rRNA gene tags in H. panicea metagenomes at class level assignment. MG1 and MG2 retrieved from two H. panicea individuals collected from South-West Iceland and used for differential abundance binning of "Ca. H. symbioticus" genome HS1. Black framed bars indicate the relative abundance of sponge symbiont "Ca. H. symbioticus".*

Phylogenomic inference of HS1 genome

Previous phylogenetic analysis based on the 16S rRNA gene placed "*Ca*. H. symbioticus" into a separate clade to members of the genus *Amylibacter* in the family *Rhodobacteraceae*, however with low bootstrap support (Knobloch *et al.*, 2018). Here, phylogenomic analysis based on 31 conserved proteins of closely related bacteria and selected sponge symbionts placed HS1 into a clade with two *Amylibacter* species and an unclassified *Rhodobacterales* bacterium (HTCC2255) with strong bootstrap support (Figure 2), indicating a closer relatedness between these species. Similar to the morphological description of "*Ca*. H.

symbioticus" (Knobloch *et al.*, 2018) all published *Amylibacter* spp. are short, rod-shaped, non-motile cells and were isolated from the marine environment (Teramoto and Nishijima, 2014; Nedashkovskaya *et al.*, 2016; Wang *et al.*, 2017; Feng *et al.*, 2018; Wong *et al.*, 2018). Based on a study of the sponge specificity of 16S rRNA gene sequence by Simister *et al.* (2012), published sequences of "*Ca*. H. symbioticus" (GenBank: Z88578, Z88567, Z88569, AY948354, AY948358) neither formed a sponge-specific sequence cluster with other sponge-associated sequences, nor were they closely related to another sponge-specific cluster. This, in addition to the phylogenomic tree reported here (Figure 2), indicates that HS1 evolved from a free-living seawater bacterial lineage related to extant members of the genus *Amylibacter*, as opposed to a sponge-specific lineage. Two species in the *Amylibacter* genus, *A. cionae* and *A. ulvae*, the genome of the latter not yet sequenced, were isolated from a sea squirt and macroalgae respectively (Nedashkovskaya *et al.*, 2016; Wang *et al.*, 2017). Though both *Amylibacter* species were readily grown on marine agar and could thus be characterised as free-living bacteria, as opposed to obligate symbionts, their isolation sources indicate a certain degree of association with benthic marine organisms.

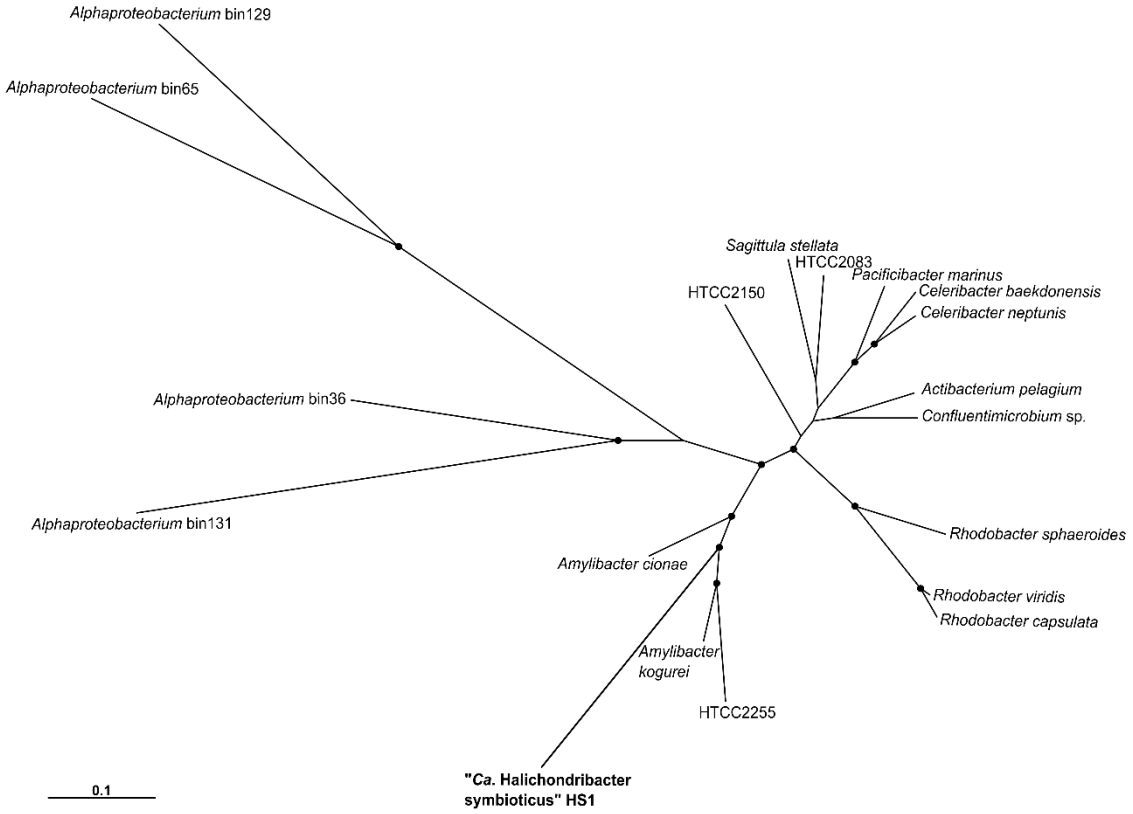


Figure 2: *Maximum-likelihood phylogenomic tree of HS1 and reference genomes. Phylogenetic tree based on 31 concatenated conserved proteins. Bootstrap values above 95% are marked with a black circle at the branching sites. Scale shows substitutions per nucleotide position.*

Based on the phylogenomic tree, the three closely related free-living bacteria, *A. cionae*, *A. kogruei* and *Rhodobacterales* bacterium HTCC2255, were selected for comparative genomic analysis, as well as four genomes described in a study by Slaby *et al.* (2017), as sponge symbiont representatives of the *Alphaproteobacteria* lineage found in the sponge *Aplysina aerophoba* (see Table 1).

Genome comparison of HS1 and ancestral gene content prediction

A comparison of orthologous groups of proteins showed that 983 COGs were shared between HS1 and the related free-living seawater bacterial genomes (Figure 3), indicating the presence of a core genome between these species. 51 COGs were only found in the HS1 genome, most of which were in categories S ('Function unknown'), R ('General function prediction only') and L ('Replication, recombination and repair') with 13, 12 and 9 COGs each (see Supplementary Table S1 for full list of COGs). In comparison, the three free-living reference genomes shared 157 COGs with each other that were not found in the HS1 genome which, however, might partially be accounted for by the genome incompleteness of HS1 (Figure 3). Compared against the reference genomes and selected sponge symbionts, HS1 most closely resembled the COG category distribution and relative abundance of HTCC2255 and the sponge symbiont bin131, albeit with markedly less COGs associated with category G ('Carbohydrate transport and metabolism'), suggesting a similar functional gene repertoire between the two sponge symbionts and the free-living bacterium (Table 1 and Figure 4). In addition, HS1, HTCC2255 and bin131 shared a similar number of COGs (Tables 1), which was significantly lower ($p < 0.005$, t-test) than in the other reference genomes.

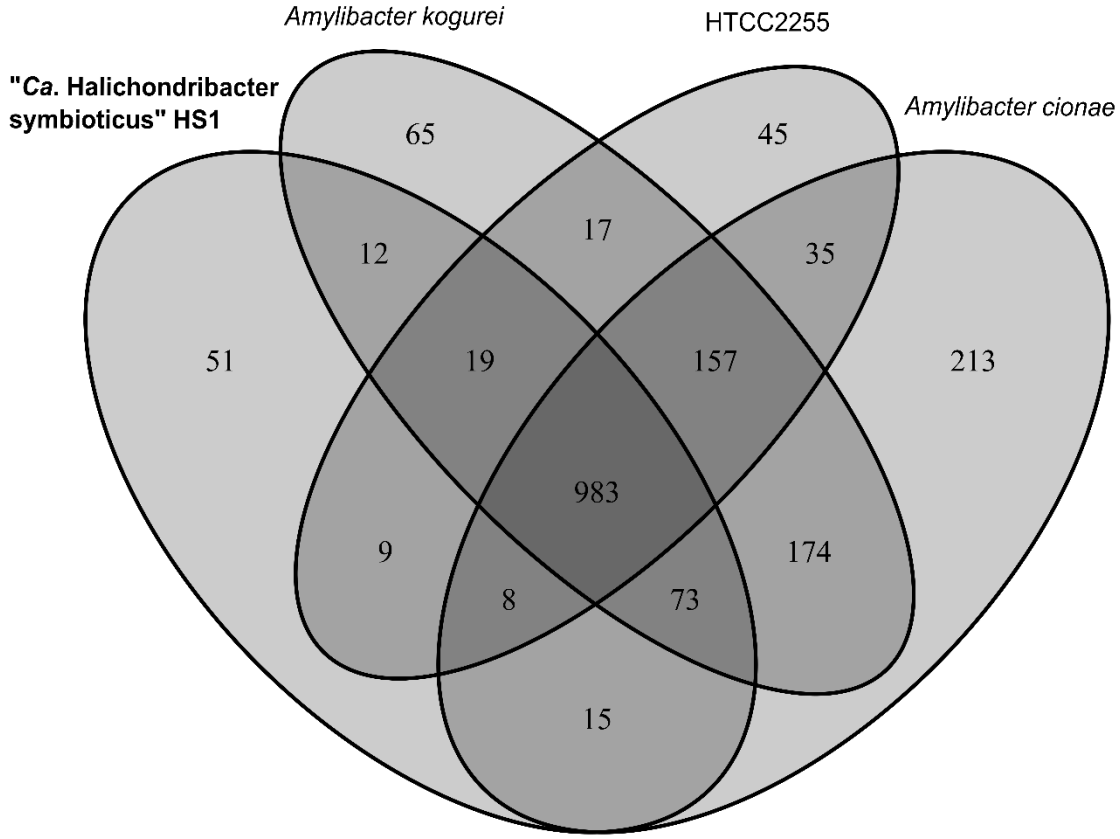


Figure 3: *Venn diagram of shared COGs between HS1 and closely related free-living reference genomes.*

***Table 1**: General properties of "*Ca*. H. symbioticus" draft genome HS1 and of selected reference genomes. *genome completeness from Slaby et al. (2017); + pathways complete; - [gene missing in pathway]*

	"*Ca*. H. symbioticus" HS1	**free-living seawater bacteria**			**sponge symbionts**			
		***R*. HTCC2255**	***A. kogurei***	***A. cionae***	**bin36**	**bin65**	**bin129**	**bin131**
RefSeq/Genbank	RWJS00000000	NZ_AATR00000000.1	NZ_MDGM00000000.1	NZ_PKOJ00000000.1	MPMP00000000.1	MPMX00000000.1	MPNO00000000.1	MPNP00000000.1
family	Rhodobacteraceae	Rhodobacteraceae	Rhodobacteraceae	Rhodobacteraceae	Rhodobacteraceae	Rhodospirillaceae	Rhodospirillaceae	Rhodobacteraceae
Genome size [Mbp]	2.89	2.30	2.98	4.29	5.6	4.29	4.83	3.74
GC-content [%]	49.4	38.9	48.8	56.7	58.0	66.2	69.6	42.0
CDS	3175	2211	2892	4001	5122	4036	4742	3392
COGs	1170	1273	1500	1658	1436	1502	1400	1229
KOs	1414	1474	1651	2123	2323	2088	2208	1497
SCMGs	99	103	105	105				
estimated completeness [%]	>93	>97	>99	>99	89.19*	91.89*	82.88*	93.69*
Number of genes in COG categories								
Transcription	62	114	161	287	144	154	143	78
Carbohydrate transport and metabolism	119	186	195	252	492	171	272	219

(Table continued from previous page)

Amino acid transport and metabolism	267	321	299	509	642	540	521	277
Number of genes in KEGG pathways								
DNA replication	14	14	14	14	14	14	12	14
TCA cycle	18	19	19	19	20	24	20	19
Pentose phosphate pathway	17	15	18	19	19	10	13	17
Peptidoglycan biosynthesis	15	15	16	17	15	16	15	15
Nitrogen metabolism	4	5	12	12	10	11	13	4
Sulphur metabolism	15	25	20	20	19	23	18	11
Amino acid biosynthesis pathways								
Serine biosynthesis	-[*serA*,*serC*,*serB*]	+	+	+	+	+	-[*serC*,*serB*]	+
Phenylalanine biosynthesis	-[*tyrB*]	+	+	+	+	-[*tyrB*]	-[*tyrB*]	+
Tyrosine biosynthesis	-[*tyrB*]	+	+	+	+	-[*tyrB*]	-[*tyrB*]	+

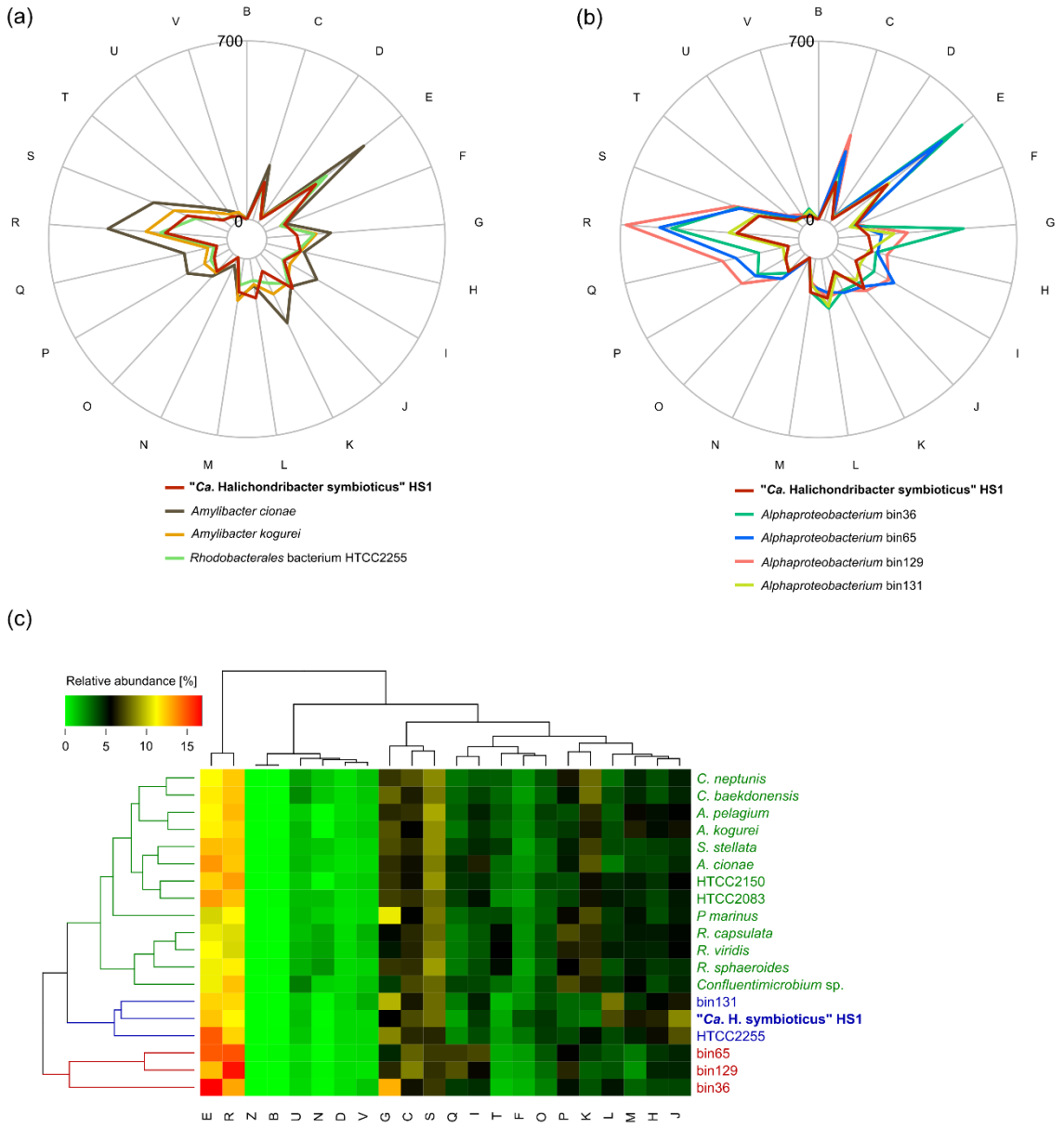


Figure 4: *Distribution and relative abundance of COG categories between HS1 and reference genomes. (a) Distribution of COG categories for HS1 and reference genomes of free-living seawater bacteria and (b) selected sponge symbionts in the Alphaproteobacteria class. (c) Relative abundance (percentage) of COG categories in relation to the total number of gene clusters in each genome. Row and column dendrograms based on UPGMA clustering; Genomes clustering together by COG category abundance could have similar metabolic features: green clade contains free-living seawater bacteria; blue clade contains "Ca. H. symbioticus" Hs1, sponge symbiont genome bin131 and free-living Rhodobacterales bacterium HTCC2255 genome; red clade contains other sponge symbionts of the Alphaproteobacteria class; COG functional categories: B, Chromatin structure and dynamics; C, Energy production and conversion; D, Cell cycle control, cell division, chromosome partitioning; E, Amino acid transport and metabolism; F, Nucleotide transport and metabolism; G, Carbohydrate transport and metabolism; H, Coenzyme transport and metabolism; I, Lipid transport and metabolism; J, Translation, ribosomal structure and biogenesis; K, Transcription; L, Replication, recombination and repair; M, Cell wall/membrane/envelop biogenesis; N, Cell motility; O, Post-translational modification, protein turnover, chaperones; P, Inorganic ion transport and metabolism; Q, Secondary metabolite biosynthesis, transport and catabolism; R, General function prediction only; S, Function unknown; T, Signal transduction mechanism; U, Intracellular trafficking, secretion and vesicular transport; V, Defence mechanisms.*

Ancestral genome reconstruction based on asymmetric Wagner parsimony (Swofford and Maddison, 1987; Csűrös, 2008) showed a strong gene loss event at the ancestral branching point AN2, between *A. cionae* and the other related reference genomes (Figure 5). Apart from category S and R, these included a loss of genes in categories N ('Cell motility'), P ('Inorganic ion transport and metabolism'), U ('Intracellular trafficking, secretion and vesicular transport') and Q ('Secondary metabolite biosynthesis, transport and catabolism'), largely associated with flagellar biosynthesis, bacterial type IV and VI secretion systems (Supplementary Table S2). Since motility was not observed in cultivated members of *A. cionae*, although flagellar biosynthesis genes were present in its genome, their absence at node AN2 and subsequently in its descendants is likely associated with a loss of bacterial type III secretion systems expressed by genes related to the flagellar apparatus (Galán and Collmer, 1999). Bacterial secretion systems enable bacteria to secrete proteins across their phospholipid membranes and are an essential component in many pathogenic bacteria allowing them, for instance, to secrete virulence factors into host target cells, but can also be found in both mutualistic and commensal symbionts where they are important in mediating interactions with their hosts (Pallen *et al.*, 2003; Tseng *et al.*, 2009, Green and Mecsas, 2016). Although secretion systems were found to be abundant in other sponge symbionts (Fan *et al.*, 2012), their loss in the HS1 ancestor and the low abundance in the reference *Alphaproteobacteria* sponge symbionts (Supplementary Table S2) suggests that they do not constitute a prerequisite for sponge-bacterial symbiosis, but rather point towards a more limited scope of interacting with their host.

After ancestral node AN3, gene loss in HS1 and HTCC2255 occurred independently from each other in similar functional categories (Figure 5). Over half of the 238 predicted COGs lost in HTCC2255 were also lost in HS1 but present in both *A. cionae* and *A. kogurei* (Supplementary Table S2), pointing towards a convergent evolution of many traits in the HS1 and HTCC2255 genomes. Prior to the genome availability of members of the *Amylibacter* genus, studies had shown that the HTCC2255 genome had undergone significant gene loss, branching off early from other members of its *Roseobacter* clade (Newton *et al.*, 2010; Luo *et al.*, 2013). Although many symbiotic bacteria have reduced genomes due to the loss of functional genes related to their host-dependent lifestyle (Kuwahara *et al.*, 2007; Moya *et al.*, 2008; McCutcheon and Moran, 2012), the similar gene loss events in the free-livingg HTCC2255 reveal that gene loss in HS1 cannot be completely attributed to its symbiotic lifestyle. However, acquisition of HS1 by its host could have been favoured by an already partially reduced genome.

In comparison to gene loss, gene gain was modest in HS1 with most gained genes associated with COG categories S, R and L. HS1 and HTCC2255 did not share the same gained COGs indicating that genes gained were specific to each lineage and possibly adapted to their respective ecological niche. However, gene gain could also have been underreported due to the gene gain penalty used in calculating the prediction and the low number of genomes at this branching point.

***Table 2**: Number of genes in COGs with over- or underrepresentation in HS1 compared to the reference genomes. COG descriptions are partially summarised. Where COGs were assigned to several categories, secondary categories are given in parentheses.*

Description	Category	COGs	"*Ca.* H. symbioticus" *HS1*	free-living seawater bacteria			sponge symbionts			
				R. HTCC2255	*A. kogurei*	*A. cionae*	bin36	bin65	bin129	bin131
Carbohydrate transport and metabolism										
TRAP-type C4-dicarboxylate transport systems	G	COG1593, COG1638, COG3090	11	76	27	56	55	60	130	25
ABC-type transport systems	G, (M)	COG3839, COG0395, COG1653, COG1175, COG1129, COG1879, COG1682, COG1134, COG4213, COG4214, COG4209	34	28	39	46	224	3	7	79
Coenzyme transport and metabolism										
Cobalamin biosynthesis	H	COG2109, COG1797, COG2087	10	3	3	2	3	3	3	3

(continued from previous page)

Transcription										
Transcriptional regulators	K, (E)	COG0583, COG1522, COG1802, COG1309, COG1414, COG2186, COG1167	4	44	53	130	43	39	29	10
Predicted transcriptional regulators	K	COG1396, COG0789, COG0640, COG1959	2	9	18	21	7	10	11	3
Response regulator	K, T	COG2197	0	2	4	8	2	0	2	0
Superfamily II DNA/RNA helicases, SNF2 family	K, L	COG0553	1	0	0	0	6	0	1	6
Replication, recombination and repair										
DNA methylase	L	COG0863, COG2189, COG0270, COG0338, COG2189	55	1	1	1	26	4	4	45
Superfamily I DNA and RNA helicases	L	COG0210	6	1	1	1	6	1	1	5
Predicted nuclease of the RecB family	L	COG1637	2	0	0	0	1	0	0	0

(continued from previous page)

Defence mechanisms										
Type II restriction enzyme, methylase subunits	V	COG1002	3	0	0	0	1	0	0	0
Restriction endonuclease	V	COG1715	1	0	0	0	1	0	1	2
Type I restriction-modification system	V	COG0286	2	0	2	1	7	1	1	8
General function prediction only										
TPR repeats	R	COG0790, COG0457	17	2	2	2	1	5	5	12
Ankyrin repeat	R	COG0666	9	0	0	0	2	1	0	4
Predicted helicase	R	COG4889	5	0	0	1	3	0	1	10
Function Unknown										
Toxin-antitoxin systems	S	COG1598, COG3177, COG3600	14	0	2	1	20	8	7	18

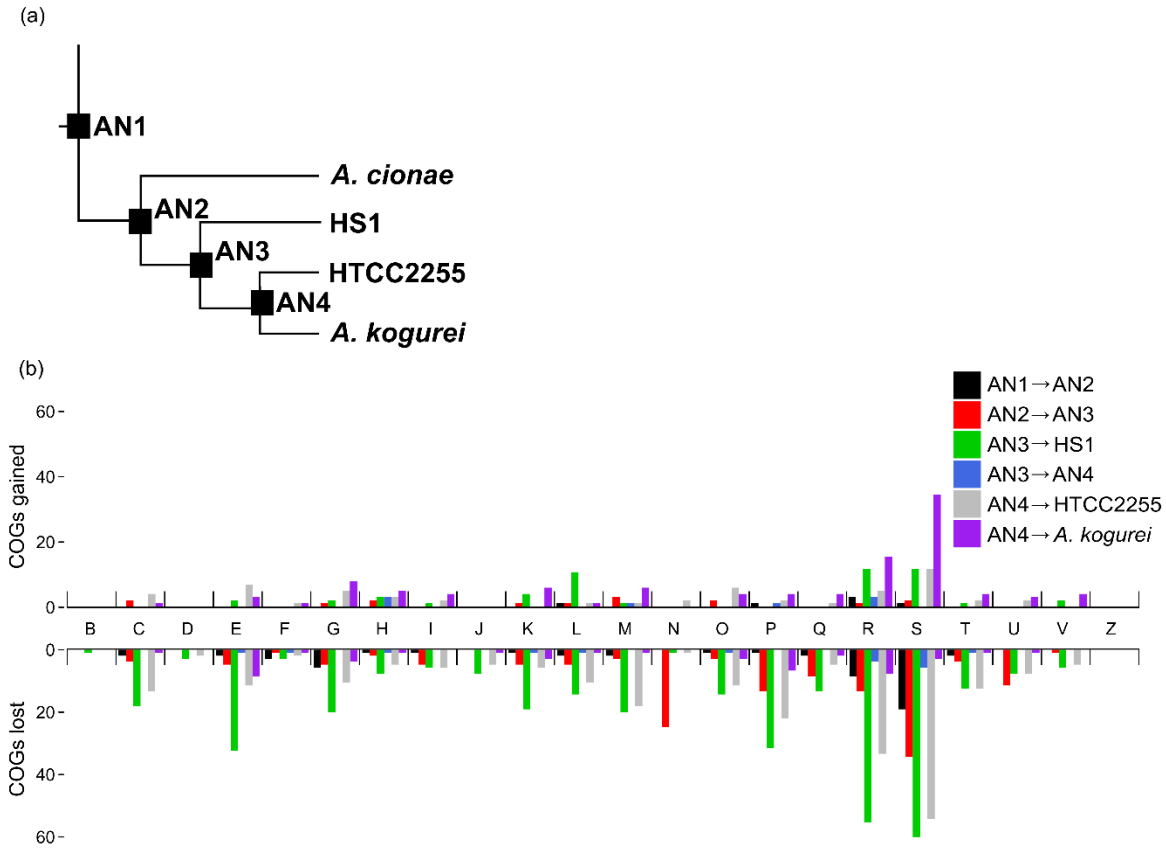


Figure 5. *Ancestral nodes and gene loss and gain events of HS1 and related lineages. (a) Section of phylogenomic tree from Figure 2 without reference sponge symbionts showing ancestral nodes between A. cionae, HS1, HTCC2255 and A. kogurei (AN2 - AN4) and all other reference genomes (AN1). (b) COG loss and gain events between ancestral nodes and extant genomes based on asymmetric Wagner parsimony.*

Sponge symbiotic features

In COG category K ('Transcription'), HS1 had a reduction of genes in COGs classified as transcription regulators, predicted transcription regulators and response regulator compared to the reference genomes (Table 2). Absence or reduction of transcriptional regulator genes has previously been reported in other sponge symbionts (Burgsdorf *et al.*, 2015; Slaby *et al.*, 2017), as well as in other symbiotic and parasitic bacterial genomes (Andersson *et al.*, 1998; Moran *et al.*, 2005). In the pea aphid symbiont *Buchnera aphidicola*, for instance, absence of transcriptional regulators points towards static transcriptional dynamics, while the main regulator of the symbiosis appears to be the host (Shigenobu and Wilson, 2011). Similarly, a loss of regulatory genes in HS1 could indicate the continuous production or consumption of metabolites in the host environment without the means of regulation by the symbiont. Many of the transcriptional regulators were present in the HTCC2255 genome, but similarly reduced in the sponge symbiont bin131 (Table 2), showing that this is likely a result of their host-dependent adaptation.

The HS1 genome was enriched in DNA methylase and nuclease associated with restriction-modification (R-M) systems, as well as in toxin-antitoxin (T-A) systems (Table 2, COG category L, V, S). These systems have previous been found to be enriched in other sponge symbionts and it has been hypothesised that they act as defence against foreign DNA, or stabilise mobile genetic elements in the chromosome and thus allow the exchange of DNA between sponge symbionts that share these R-M systems (Thomas *et al.*, 2010; Fan *et al.*, 2012; Burgsdorf *et al.*, 2015; Slaby *et al.*, 2017). In view of the ancestral gene content prediction, these R-M systems (mainly in COG category L) were one of the distinguishing gene gain events compared to the related free-living bacterium HTCC2255, showing that R-M systems appear to be an important component of this sponge-symbiont relationship.

HS1 was also enriched in genes associated with tetratricopeptide repeats (TPRs) (COG0790 and COG0457) and ankyrin repeats (ARs) (COG0666) (Table 2). These genes were rare in the free-living reference genomes and partially present in the selected sponge symbiont genomes. TPRs and ARs are domains found predominantly in eukaryotic proteins and are involved in mediating protein-protein interactions (Schmitz-Esser *et al.*, 2010). In some intracellular pathogens, such as *Legionella pneumophila*, they are necessary for replication within the eukaryotic host (Al-Khodor *et al.*, 2008). Previous studies have shown that TPRs, ARs and other eukaryotic-like proteins are enriched in sponge symbionts and are thought to be involved in sponge-symbiont interactions (Kamke *et al.*, 2010; Thomas *et al.*, 2010; Liu *et al.*, 2011; Siegl *et al.*, 2011; Fan *et al.*, 2012; Gao *et al.*, 2014). A study by Nguyen, Liu and Thomas (2014) further showed that AR proteins from a sponge symbiont, expressed in *Escherichia coli*, interfered with amoebal phagocytosis, generating a potential explanation on how sponge symbionts can prevent being digested by their host (Nguyen *et al.*, 2014). As *H. panicea* can utilize bacteria as feed source (Riisgård, Kumala and Charitonidou 2016) and "*Ca*. H. symbioticus" is highly abundant within the sponge mesohyl (Knobloch *et al.*, 2018), the presence of eukaryotic-like proteins could explain the apparent lack of being phagocytosed by the sponge.

Transcriptome of "*Ca*. H. symbioticus"

RNA-seq data from 37 wild *H. panicea* individuals from Scotland was retrieved from the public Sequence Read Archive repository under BioProject PRJNA394213. These sponges had been collected from an intertidal area and kept for 12 hours in 5-litre containers at 4 °C prior to pooled RNA extraction (Waldron *et al.*, 2018, and personal communication). "*Ca*. H. symbioticus" represented approximately 33% of the non-depleted bacterial 16S rRNA sequences in the dataset showing that it was highly represented in the sponges, though less abundant than in the Icelandic specimens. In total, 5.5% of all quality filtered reads mapped to 2163 gene coding sequences of the HS1 genome and were associated with 1071 COGs (Supplementary Table S3).

Category S had the highest relative expression (Figure 6), mainly due to two COGs (COG5457 and COG5572) of uncharacterised proteins DUF1127 and DUF2282. Although these are domains of unknown function, a gene encoding DUF1127 was reported as being co-transcribed with sRNA in *Rhodobacter sphaeroides* during various stress conditions (Hess *et al.*, 2014; Billenkamp *et al.*, 2015). Similarly, a gene encoding DUF2282 is expressed in *Caulobacter crescentus* under heavy metal stress (Kohler *et al.*, 2012) and is part of a GIG operon up-regulated in *Legionella pneumophila* in the presence of metal ions (Jwanoswki *et al.*, 2017). As the sponges were removed from their habitat prior to RNA extraction, a high expression of these genes could provide indications as to the symbiont's

stress response to either less favourable environmental conditions, such as accumulations of waste products in the storage containers, or host stress responses following tissue damage.

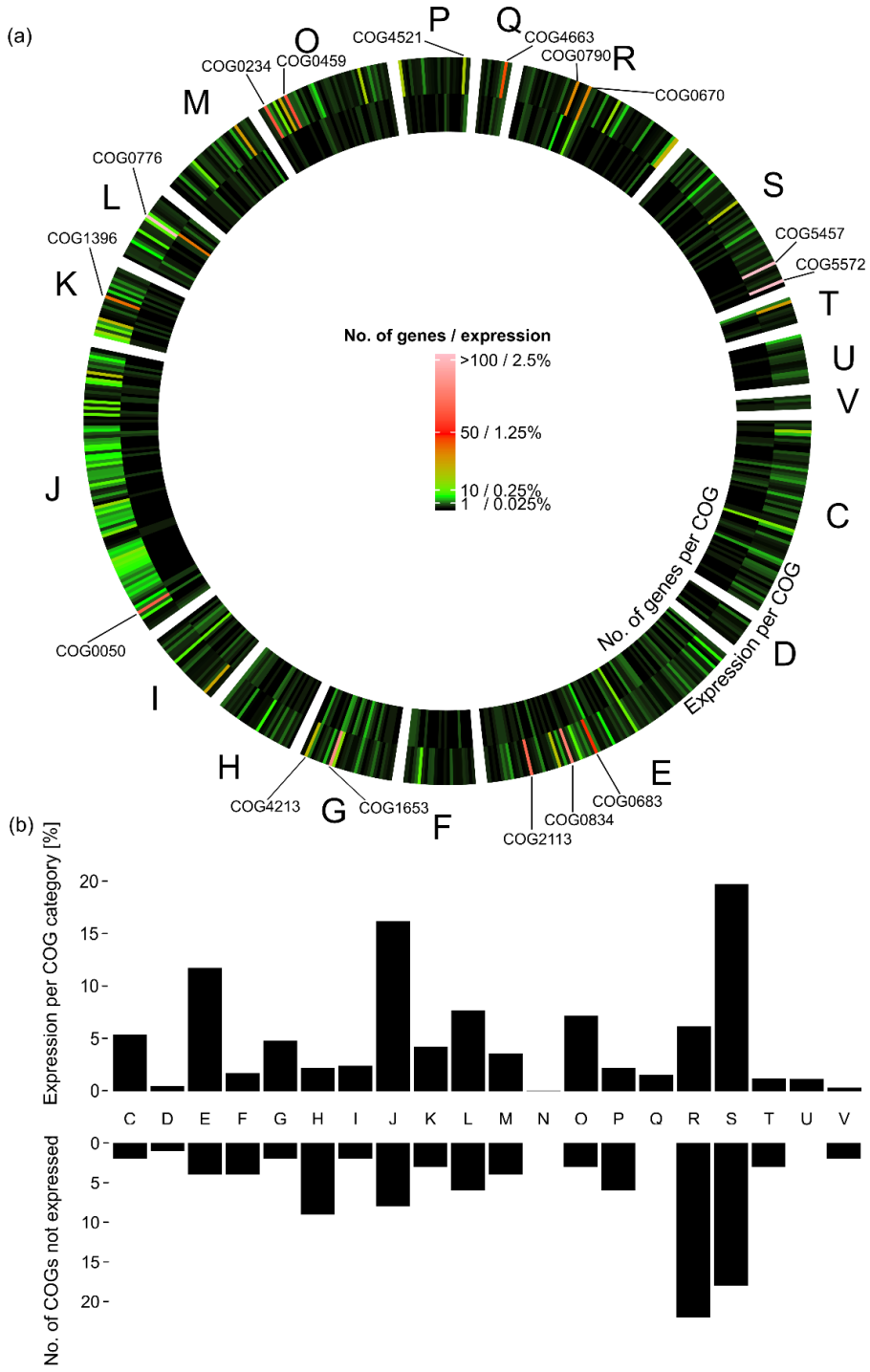


Figure 6: *Relative expression of genes in "Ca. H. symbioticus". (a) Number of genes per COG (inner ring). Relative expression of COGs as coverage per gene in each COG (outer ring). COGs in multiple categories were only counted once. Only COGs with a relative expression above 0.025% are shown and highly expressed genes are labelled. (b) Relative expression per COG category and number of COGs nor expressed.*

Category J ('Translation') had the second highest relative expression, mainly due to genes encoding ribosomal proteins and translation elongations factors, followed by category E ('Amino acid transport and metabolism'). In category E, three COGs of ABC-type amino acid transport systems (COG0683, COG0834 and COG2113) were highly expressed (Figure 6). The associated protein sequences showed highest similarity to branched-chain amino acids transporters, general L-amino acid transporters and glycine betaine-binding periplasmic proteins, possibly associated with amino acids exchange between host and symbiont. Other highly expressed COGs included an ABC-type sugar transport system (COG1653), a translation elongation factor (COG0050), a transcriptional regulator (COG1396) with high sequence similarity to antitoxin HipB, a DNA-binding protein (COG0776), a chaperonin and co-chaperonin (COG0234 and COG0459), a TRAP-type mannitol/chloroaromatic compound transport system (COG4663), an ABC-type taurine transport system (COG4521), and TPRs (COG0790). Interestingly, a previous metatranscriptomic study on planktonic seawater samples showed that the same ABC-type transporters of amino acids and sugars were also highly expressed in HTCC2255-like taxonomic groups (Ottesen *et al.*, 2011). This provides evidence that despite different habitats HS1 and HTCC2255 use similar nutrient acquisition strategies.

The 99 COGs which were not actively expressed belonged largely to uncharacterized and predicted proteins in COG category S and R (see Figure 6 and Supplementary Table S3). Interestingly, the type I R-M system (COG0286) that was also present in the free-living reference genomes was not expressed, indicating that those R-M systems present and expressed in HS1 are highly sponge-symbiont specific. Five ABC-type transport systems for thiamine, arginine/histidine, spermidine/putrescine, tungstate and ferric ion transport were not expressed, as well as a serine kinase of the HPr protein (COG1493) involved in carbon catabolite repression (Reizer *et al.*, 1998). Pathway reconstruction showed that only the first carbon oxidation of the citrate cycle, from oxaloacetate to 2-oxoglutarate, were expressed, the second carbon oxidation lacking expression of a succinate dehydrogenase / fumarate reductase. In addition, the ammonia channel protein AmtB (COG0004) was not expressed.

"*Ca*. H. symbioticus" HS1 metabolism and putative role in *H. panicea*

Compared to the reference genomes, HS1 had similar numbers of genes in carbohydrate and amino acid metabolism pathways (Figure S1 and S2). However, a reduction of genes was apparent in COGs associated with carbohydrate transport systems, in particular TRAP-type C4-dicarboxylate transport (Table 2). These transport systems are possibly the only high-affinity transporters for C4-dicarboxylates (Kelly and Thomas, 2001), indicating that the use of substrates such as malate, succinate and fumarate from exogenous sources may be limited in HS1. Although similar in COG category abundance, the disparity between HS1 and sponge symbiont bin131 in COG category G was mainly attributed to a reduction of ABC-type transport systems, further demonstrating that HS1 is less capable of utilising a wide range of carbohydrates. In addition to these features, lack of expression of carbon catabolite repression genes further points to a limited metabolic versatility of HS1. However, high expression of ABC-type sugar transport systems and TRAP-type mannitol/chloroaromatic compound transport systems, could indicate the importance of specific carbohydrate metabolism pathways, possibly adapted to the carbohydrate availability in the sponge body (Figure 7). Most amino acid biosynthesis pathways found in the reference genomes could also be reconstructed in HS1, with the exception of tyrosine and phenylalanine biosynthesis which were lacking the *tyrB* gene, and serine biosynthesis which was lacking *serA*, *serC* and *serB* (Figure 7). However, a gene coding for serine hydroxymethyltransferase, an enzyme

that reversible catalyses serine from glycine, was present, displaying an alternative serine biosynthesis pathway in HS1.

Nitrogen metabolism and ammonia assimilations in sponge symbionts have been studied extensively, highlighting a potential mutualistic interrelation in which symbionts benefits from a stable availability of nitrogenous waste products, whereas the sponge benefits from its removal (Webster and Thomas, 2016). Genes for ammonia assimilation were present in HS1 through the GS-GOGAT cycle which converts glutamate to glutamine using ammonia followed by two glutamate molecules using 2-oxoglutarate (Merrick and Edwards, 1995). Other pathways for nitrogen assimilation, such as ammonium oxidisation and denitrification, previously found in other sponge symbionts (Hoffmann *et al.*, 2009; Siegl *et al.*, 2011), were not present in HS1. A lack of genes for glutamate dehydrogenase, which was present in the reference genomes and which poses an alternative ammonia assimilation pathway often under high ammonia concentrations (Tempest *et al.*, 1970), as well as absence of *amtB* expression, suggests that ammonia assimilation does not occur at an elevated rate and is dependent on ammonia concentrations in the surrounding environment. AmtB transports ammonia through the cell membrane and is induced under conditions of ammonia limitation (Soupene *et al.*, 1998; Javelle *et al.*, 2005). Absence of AmtB in HS1, would reduce the rate of ammonia transported into the cell under ammonia limited conditions, leaving it entirely to unspecific diffusion across the cell membrane. Ammonia is excreted by sponges as a metabolic waste product (Mohamed *et al.*, 2010) and its concentration in the sponge body is therefore likely connected to its metabolic rate. *H. panicea* can be periodically exposed to air in the intertidal zone and even in subtidal areas exhibits periods of pumping inactivity (Kumala and Canfield, 2018). Absence of the ability to assimilate sufficient nitrogen by the symbiont under such conditions could reduce amino acids synthesis and prevent excessive growth and oxygen consumption of HS1 under lower metabolic rates of its hosts, providing an effective means of symbiont regulation. Cultivation of *H. panicea* under low nutrient availability has previously shown to induce a reduction of "*Ca*. H. symbioticus" in the sponge body (Knobloch *et al.*, 2018). Although this has not been unambiguously linked to ammonia secretion in the sponge it provides a potential explanation to this observation.

HS1 further contained genes associated with peptidase and protease activity (COG1404, COG0616 and COG2856) which could allow breakdown of peptides in the sponge mesohyl, for instance from phagocytised foreign bacteria, and subsequent transport of amino acids into the cell. Animal hosts can provide valuable nutrients such as amino acids for the proliferation of their symbionts (Graf and Ruby, 1998) and conversely, symbionts can provide their hosts with essential amino acids such as in the pea amphid-*Buchnera* system (Shigenobu and Wilson, 2011). In this context, three highly expressed amino acid transport systems indicate a dynamic exchange of amino acids between the symbiont and its host environment.

Genes involved in cobalamin (vitamin B_{12}) synthesis were enriched in HS1 compared to the reference genomes (Table 2), and genes involved in thiamine, riboflavin and biotin synthesis were present although not more frequently than in the free-living genomes. Cobalamin is an essential cofactor for all animals and enrichment of genes involved in its synthesis has previously been reported for several sponge symbiont (Thomas *et al.*, 2010; Siegl *et al.*, 2011; Fan *et al.*, 2012; Tian *et al.*, 2016), suggesting that *H. panicea* benefits from this feature of the symbiotic relationship.

Sponges and their symbionts can produce a wide variety of bioactive secondary metabolites which act as a defence system against predators and epibionts (Hentschel *et al.*, 2012). Search for secondary metabolite gene clusters in HS1 using antiSMASH (Blin *et al.*, 2017) revealed a single bacteriocin gene cluster of the DUF692 domain. Bacteriocins are antibacterial proteins that display inhibitory activity against other, often closely related bacteria (Jack *et al.*, 1995). In *H. panicea* this could be an effective means of preventing the establishment of foreign or pathogenic bacteria, thus reducing competition for nutrients in the sponge body or protecting its host. However, the presence of the same bacteriocin gene cluster, as well as nine and seven additional secondary metabolite clusters in the *A. cionae* and *A. kogurei* genomes respectively (Supplementary Table S2) shows that chemical defence is not a pronounced feature in HS1. The lack of other bioactive secondary metabolite gene clusters also renders in unlikely that HS1 can directly protect *H. panicea* from predators or other eukaryotic fouling organisms.

HS1 shared functional similarities with the phylogenetically only distantly related sponge symbiont bin131, both based on COG category abundance and enrichment or depletion of gene families. In the original research by Slaby et al. (2017) it was suggested that the symbionts of the sponge *Aplysina aerophoba* could be classified into three guilds based on their metabolic specialisation, whereas bin131 belonged to the guild of nutritional generalists. In line with this concept, the genome of HS1 did not contain an extraordinary number of genes that indicated the utilisation of specific metabolites such as carnitine or sulphated polysaccharides as did members of the other two guilds respectively. Other specialised metabolic functions found in sponge symbionts such as sulphur oxidation (Gauthier *et al.*, 2016), sulphate reduction (Hoffmann *et al.*, 2005), utilisation of creatinine, pyrimidines or 5-oxoproline (Fan *et al.*, 2012), carbon monoxide dehydrogenase (Thomas *et al.*, 2010), or urea utilisation (Siegl *et al.*, 2011) were also not detected in HS1. This lack of specialisation and possession of more general metabolic features also present in free-living bacteria might explain why some *H. panicea,* even from similar locations, do not contain "*Ca*. H. symbioticus" and thus do not appear to rely on its symbiotic features. The loss of bacterial secretion systems and lack of regulator genes in HS1, however, suggests that the acquisition was instigated by the sponge and the symbiotic relationship mainly regulated by it, and as such strongly points towards a mutualistic relationship between the "*Ca*. H. symbioticus" and *H. panicea*.

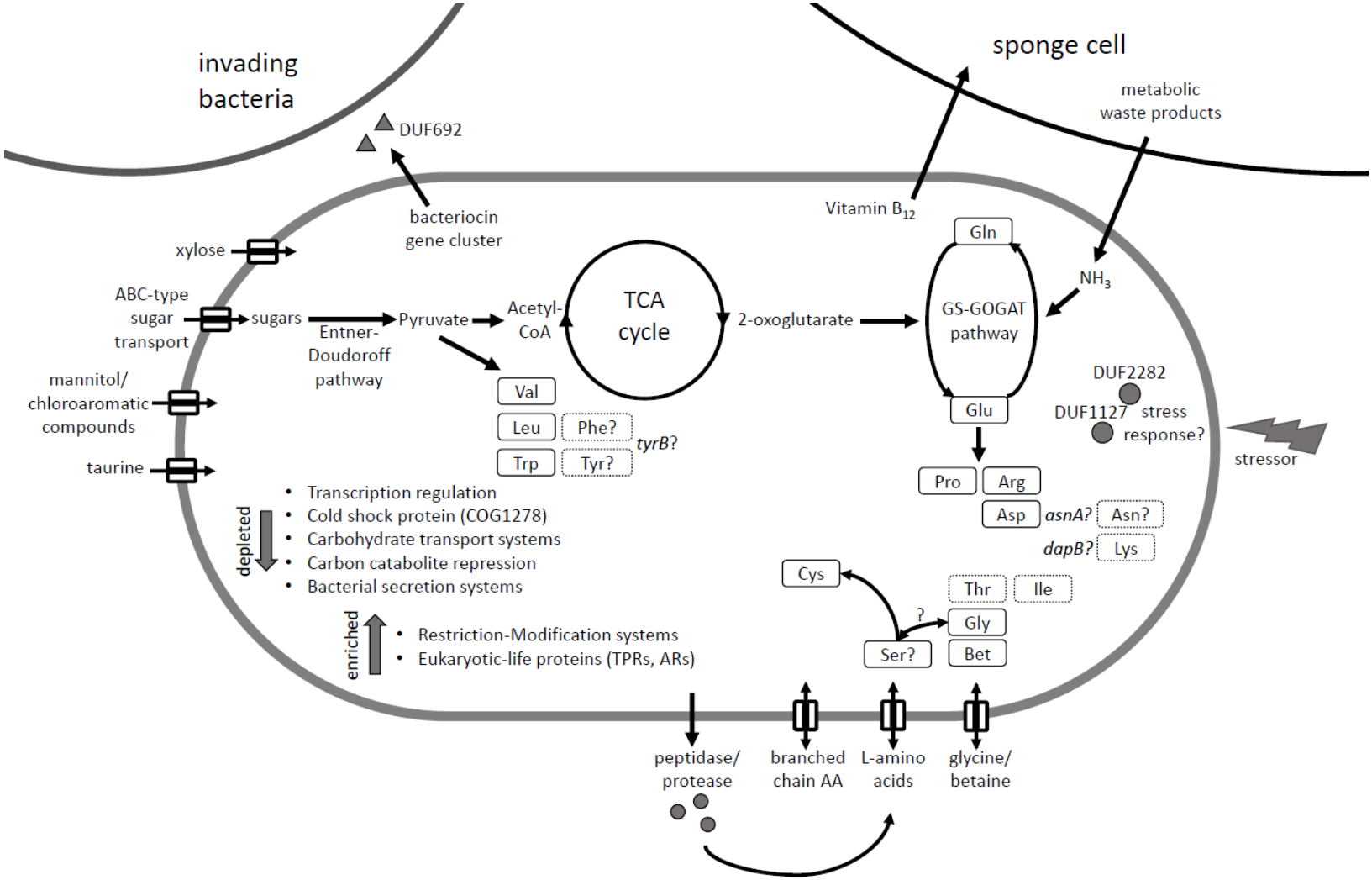


Figure 7: *Schematic figure showing functional traits of the sponge symbiont "*Ca. *H. symbioticus". Functions related to its symbiotic life-style were deduced from the genomic analysis of the draft genome HS1 and gene expression of a publicly available transcriptomic study. Black-rimmed boxes indicate full amino acid biosynthesis pathways present. Boxes with dotted frames represent incomplete pathways or pathways missing precursors. Missing genes and inconclusive interactions contain a question mark.*

Conclusions

Bacterial symbionts are important members of the sponge ecosystem and the increasing number of studies on their functional properties is beginning to elucidate their functional diversity and metabolic specialisation while confirming the widespread presence of shared sponge symbiotic features. Here we report the first draft genome of the sponge symbiont "*Ca*. H. symbioticus" HS1, the dominant symbiont in the marine sponge *H. panicea*. Dynamic exchange of metabolites between the symbiont and the host environment, ammonia assimilation, vitamin synthesis and the production of a putative bacteriocin are features that could explain a mutualistic relationship in which the symbiont has access to a stable supply of nutrients and a protected habitat, while the host benefits from the removal of metabolic waste products, availability of essential nutrients and cofactors, as well as growth inhibition of foreign bacteria. HS1 was enriched in sponge symbiotic features and shared a similar COG category distribution and relative abundance to a previously described sponge symbiont in the Mediterranean sponge *Aplysina aerophoba*. This underpins the convergent evolution of functional genomic traits among spatially separate sponge symbionts, previously suggested for sponge symbionts (Fan *et al.*, 2012; Ribes *et al.*, 2012) and other symbiotic bacteria (McCutcheon *et al.*, 2009).

The genomic insights into HS1 presented here are thought to increase our understanding of sponge-bacterial symbiosis. The *H. panicea* – "*Ca*. H. symbioticus" system in particular presents an interesting model for studying sponge-microbe interactions due to the presence of a dominant bacterial symbiont and the amenability of *H. panicea* towards laboratory cultivation. Future studies on gene expression under different environmental conditions and elucidation of complementary genomic features in the *H. panicea* genome are likely to reveal a more in-depth characterisation of this symbiotic system and of early animal-microbe symbiosis in general.

Experimental procedures

Sample collection and identification

Samples of *H. panicea* were collected in South-West Iceland (64°09' N, 22°00' W) for metagenomic analysis. Samples from two sponge individuals were ground in calcium- and magnesium-free artificial seawater to release intercellular bacteria, followed by centrifugation at 200 x g for 2 min at 4°C in microcentrifuge tubes to remove the majority of sponge cells and debris. The supernatant was centrifuged at 500 x g for 5 min to remove remaining sponge cells followed by centrifugation at 8000 x g for 10 min to pellet the bacterial fraction. The bacterial pellet was washed twice in sterile 1 x PBS and frozen at -80 °C until DNA extraction. Sponges were identified via spicule structure and COI gene analysis as described in Knobloch *et al.* (2018).

Metagenomic sequencing, assembly, binning and annotation

DNA extraction, library construction and sequencing

Genomic DNA from the bacterial pellets was extracted using the *MasterPure Complete DNA and RNA purification Kit* (Epicentre) according to the manufacturer's instructions. Paired

end libraries were constructed with the *Nextera XT DNA Library Preparation Kit* (Illumina) according the manufacturer's instruction without size selection. The libraries were sequenced on a MiSeq System using v3 chemistry and 2 x 300 cycles, resulting in approximately 3.8 million paired end reads.

Assembly, binning and annotation

Raw reads were trimmed and quality filtered using Trimmomatic (Bolger *et al.*, 2014) with settings LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:100 CROP:250 HEADCROP:10. Read coverage was normalized to 20x coverage while discarding reads with a depth under 2x using bbnorm of the BBMap package version 36.19 (https://sourceforge.net/projects/bbmap/). Coverage normalized reads which overlapped by more than 50 bp and less than 15% mismatches were merged with FLASH (Magoc and Salzberg, 2011) and both merged and unmerged reads were assembled using SPAdes version 3.11.0 (Bankevich *et al.*, 2012) running the -meta flag (Nurk *et al.*, 2017) and kmer sizes 21, 33 and 55.

In order to extract the genome of "*Ca*. H. symbioticus", contigs were binned using the Multi-metagenome workflow (http://madsalbertsen.github.io/multi-metagenome) described in Albertsen *et al.* (2013). In short, quality filtered paired end reads were mapped against the assembly with Bowtie2 version 2.3.2 (Langmead and Salzberg, 2012) and read coverage of each contig was calculated with the depth function of SAMtools version 1.2 (Li *et al.*, 2009). Tetranucleotide frequency patterns and GC content of the assembled contigs was generated with perl scripts from the Multi-metagenome workflow. ORFs were predicted using MetaProdigal version 2.6.2 (Hyatt *et al.*, 2012) and searched against hidden Markov models of 107 bacterial proteins with *HMMER* version 3.1b2 in order to identify essential single-copy maker genes (SCMGs). Taxonomic affiliation of essential SCMGs was identified by BLASTP (BLAST version 2.4.0+, Camacho *et al.*, 2009) against the NCBI refseq_protein database (Altschul *et al.*, 1990) and classified using MEGANs LCA algorithm (Huson *et al.*, 2016). Genome binning was performed manually in RStudio based on read coverage and further refined using principal component analysis of tetranucleotide frequency patterns. Contaminations were manually removed by searching ORFs against the NCBI refseq_protein database using BLASTP. Completeness of the binned draft genome of "*Ca*. H. symbioticus" HS1 was estimated based on the number of SCMGs detected in the bin and the expected number of SCMGs for the *Alphaproteobacteria* class (Albertsen *et al.*, 2013). Predicted ORFs of HS1 and reference genomes were annotated using RPS-BLAST+ against the 2014 release of the COG database (Galperin *et al.*, 2015) accessed from NCBI (https://ftp.ncbi.nlm.nih.gov/pub/mmdb/cdd, downloaded on 9.10.2017) with an e-value cut-off of 1e-5. Duplicate assignments were removed keeping the assignment with the higher e-value. COG category and functional descriptions were inferred using the cdd2cog.pl script (Leimbach, 2016) and reference databases (ftp://ftp.ncbi.nlm.nih.gov/pub/COG/COG/, downloaded on 9.10.2017). KO assignment of ORFs were made using the web service BlastKOALA (Kanehisa *et al.*, 2016) against the KEGG database (Kanehisa and Goto, 2000) (accessed 1.-14.11.2018).

Bacterial diversity and phylogenomic analysis

Partial 16S rRNA sequences were extracted from the quality-filtered, merged reads using Meta_RNA(H3) (Huang *et al.*, 2009). Predicted 16S rRNA gene fragments larger than 100 bp were classified against the SILVA v123 LTP database (Yilmaz *et al.*, 2014) using the SINTAX command of the USEARCH package version 10.0.240 (Edgar, 2010, 2016). The relative taxonomic abundance was analysed in R version 3.4.2 (R Core Team, 2017) with the phyloseq package (McMurdie and Holmes, 2013). Sequences were also aligned to the full 16S rRNA gene of "*Ca*. H. symbioticus" (Genbank: MH734183 and MH734529) in order to detect the relative abundance of the symbiont in the metagenomes. Reference genomes for phylogenomic analysis were selected based on 16S rRNA gene similarity to "*Ca*. H. symbioticus" using BLAST against the NCBI refseq_genomes database (see Supplementary Table S4 for full list and accession numbers). In addition, four genome bins of sponge symbiont in the *Alphaproteobacteria* class from the sponge *Aplysina aerophoba* (Slaby *et al.*, 2017) which showed closest 16S rRNA gene sequence similarity to "*Ca*. H. symbioticus" were used. 31 conserved proteins of HS1 and the reference genomes were predicted using Amphora2 (Wu and Scott, 2012) and aligned using the MUSCLE algorithm (Edgar, 2004) implemented in ARB (Ludwig *et al.*, 2004). Alignments were concatenated and a maximum-likelihood phylogenetic tree was then constructed using the PHYML algorithm (Guindon and Gascuel, 2003) as descripted previously (Wu and Eisen, 2008).

Comparative genomics and ancestral gene content prediction

Functional KEGG pathways and modules were reconstructed using the KEGG Mapper (https://www.genome.jp/kegg/mapper.html) to validate completeness of pathways or detect absence of genes of incomplete pathways of interest. Missing genes and COGs were searched against the ORF predictions of the entire MG1 and MG2 metagenome using BLASTP with the homologue sequences of the reference genomes. Matching sequences with an e-value cut-off below 1e-5 were searched against the NCBI refseq_protein database for taxonomic classification in order to ensure that genes were not falsely excluded from the HS1 genome during binning. Plots for comparative genome analysis were constructed in R version 3.4.2 using package ggplot2 (Wickham, 2009), a custom script for radial plots (http://pcwww.liv.ac.uk/~william/Geodemographic%20Classifiability/func%20CreateRadialPlot.r) and package VennDiagram (Chen and Boutros, 2011). A heatmap was created using the heatmap.2 function of the gplots package (Warnes *et al.*, 2016) and clustered according to UPMGA implemented in the hclust function of the R package stats. Ancestral gene content prediction was performed in the Count software (Csuos, 2010) using asymmetric Wagner parsimony (Swofford and Maddison, 1987; Csűrös, 2008) on COG distribution of HS1 and the free-living reference genomes. A gene gain penalty of 1.3 was used and compared against Dollo parsimony (Farris, 1977) which yielded comparable results ($R>0.79$, Pearson correlation).

Gene expression calculation

RNA-seq data was retrieved from the public Sequence Read Archive repository under BioProject PRJNA394213, accession SRX4378335. Sample collection, processing and sequencing are described in Waldron *et al.*, 2018. Relative abundance of "*Ca*. H. symbioticus" was measured as for the *H. panicea* metagenomes. For calculation of relative gene expression, sequences were first quality filtered using Trimmomatic with previously mentioned settings and mapped against HS1 ORFs using Bowtie2. Read coverage was

calculated with the depth function of SAMtools and summarised by assigned COG. Expression in COG categories was performed by summarising relative expression of COGs in each category, whereas COGs assigned to multiple categories were only counted once.

Nucleotide sequence deposits

The raw Illumina sequencing data has been deposited under BioProject PRJNA508373. The Whole Genome Shotgun project of HS1 has been deposited at DDBJ/ENA/GenBank under the accession RWJS00000000. The version described in this paper is version RWJS01000000.

Acknowledgements

This work was supported by the EC Marie Curie Actions ITN Project BluePharmTrain [No. 607786].

References

Al-Khodor, S., Price, C.T., Habyarimana, F., Kalia, A., and Abu Kwaik, Y. (2008) A Dot/Icm-translocated ankyrin protein of *Legionella pneumophila* is required for intracellular proliferation within human macrophages and protozoa. *Mol. Microbiol.* **70**: 908–923.

Albertsen, M., Hugenholtz, P., Skarshewski, A., Nielsen, K.L., Tyson, G.W., and Nielsen, P.H. (2013) Genome sequences of rare, uncultured bacteria obtained by differential coverage binning of multiple metagenomes. *Nat Biotech* **31**: 533–538.

Althoff, K., Schütt, C., Steffen, R., Batel, R., and Müller, W.E.G. (1998) Evidence for a symbiosis between bacteria of the genus *Rhodobacter* and the marine sponge *Halichondria panicea*: Harbor also for putatively toxic bacteria? *Mar. Biol.* **130**: 529–536.

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.

Andersson, S.G.E., Zomorodipour, A., Andersson, J.O., Sicheritz-Pontén, T., Alsmark, U.C.M., Podowski, R.M., et al. (1998) The genome sequence of Rickettsia prowazekii and the origin ofmitochondria. *Nature* **396**: 133–140.

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., et al. (2012) SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J. Comput. Biol.* **19**: 455–477.

Barthel, D. (1988) On the ecophysiology of the sponge Halichondria panicea in Kiel Bight. 2. Biomass, production, energy budget and integration in environmental processes. *Mar. Ecol. Prog. Ser. Oldend.* **43**: 87–93.

Barthel, D. (1986) On the ecophysiology of the sponge Halichondria panicea in Kiel Bight. I. Substrate specificity, growth and reproduction. *Mar. Ecol. Prog. Ser.* **32**: 291–298.

Barthel, D. and Theede, H. (1986) A new method for the culture of marine sponges and its application for experimental studies. *Ophelia* **25**: 75–82.

Barthel, D. and Wolfrath, B. (1989) Tissue sloughing in the sponge Halichondria panicea: a fouling organism prevents being fouled. *Oecologia* **78**: 357–360.

Beazley, L.I., Kenchington, E.L., Murillo, F.J., and Sacau, M. del M. (2013) Deep-sea sponge grounds enhance diversity and abundance of epibenthic megafauna in the Northwest Atlantic. *ICES J. Mar. Sci.* **70**: 1471–1490.

Billenkamp, F., Peng, T., Berghoff, B.A., and Klug, G. (2015) A cluster of four homologous small RNAs modulates C1 metabolism and the pyruvate dehydrogenase complex in

Rhodobacter sphaeroides under various stress conditions. *J. Bacteriol.* **197**: 1839–52.
Blin, K., Wolf, T., Chevrette, M.G., Lu, X., Schwalen, C.J., Kautsar, S.A., et al. (2017) antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res.* **45**: W36–W41.
Bolger, A.M., Lohse, M., and Usadel, B. (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120.
Burgsdorf, I., Slaby, B.M., Handley, K.M., Haber, M., Blom, J., Marshall, C.W., et al. (2015) Lifestyle evolution in cyanobacterial symbionts of sponges. *MBio* **6**: e00391-15.
Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009) BLAST+: architecture and applications. *BMC Bioinformatics* **10**: 421.
Chen, H. and Boutros, P.C. (2011) VennDiagram: a package for the generation of highly-customizable Venn and Euler diagrams in R. *BMC Bioinformatics* **12**: 35.
Csuos, M. (2010) Count: evolutionary analysis of phylogenetic profiles with parsimony and likelihood. *Bioinformatics* **26**: 1910–1912.
Csűrös, M. (2008) Ancestral Reconstruction by Asymmetric Wagner Parsimony over Continuous Characters and Squared Parsimony over Distributions. Springer, Berlin, Heidelberg, pp. 72–86.
Díez-Vives, C., Moitinho-Silva, L., Nielsen, S., Reynolds, D., and Thomas, T. (2017) Expression of eukaryotic-like protein in the microbiome of sponges. *Mol. Ecol.* **26**: 1432–1451.
Edgar, R. (2016) SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. *bioRxiv* 074161.
Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* **32**: 1792–1797.
Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.
Erpenbeck, D., Knowlton, A.L., Talbot, S.L., Highsmith, R.C., and Van Soest, R.W. (2004) A molecular comparison of Alaskan and Northeast Atlantic Halichondria panicea (Pallas 1766)(Porifera: Demospongiae) populations. *Boll Mus Ist Univ Genova* **61**: 319–325.
Erwin, P.M. and Thacker, R.W. (2007) Incidence and identity of photosynthetic symbionts in Caribbean coral reef sponge assemblages. *J. Mar. Biol. Assoc. UK* **87**: 1683–1692.
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. U. S. A.* **109**: E1878-87.
Farris, J.S. (1977) Phylogenetic Analysis Under Dollo's Law. *Syst. Biol.* **26**: 77–88.
Feng, T., Kim, K.H., Chun, B.H., and Jeon, C.O. (2018) Amylibacter lutimaris sp. nov., isolated from sea-tidal flat sediment. *Int. J. Syst. Evol. Microbiol.* **68**: 2088–2092.
Feuda, R., Dohrmann, M., Pett, W., Philippe, H., Rota-Stabelli, O., Lartillot, N., et al. (2017) Improved Modeling of Compositional Heterogeneity Supports Sponges as Sister to All Other Animals. *Curr. Biol.* **27**: 3864–3870.e4.
Flatt, P.M., Gautschi, J.T., Thacker, R.W., Musafija-Girt, M., Crews, P., and Gerwick, W.H. (2005) Identification of the cellular site of polychlorinated peptide biosynthesis in the marine sponge Dysidea (Lamellodysidea) herbacea and symbiotic cyanobacterium Oscillatoria spongeliae by CARD-FISH analysis. *Mar. Biol.* **147**: 761–774.
Galán, J.E. and Collmer, A. (1999) Type III Secretion Machines: Bacterial Devices for Protein Delivery into Host Cells. *Science (80-.).* **284**: 1322–1328.
Galperin, M.Y., Makarova, K.S., Wolf, Y.I., and Koonin, E. V. (2015) Expanded microbial

genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res.* **43**: D261–D269.

Gao, Z.-M., Wang, Y., Tian, R.-M., Wong, Y.H., Batang, Z.B., Al-Suwailem, A.M., et al. (2014) Symbiotic adaptation drives genome streamlining of the cyanobacterial sponge symbiont "Candidatus Synechococcus spongiarum" *MBio* **5**: e00079-14.

Gauthier, M.-E.A., Watson, J.R., and Degnan, S.M. (2016) Draft Genomes Shed Light on the Dual Bacterial Symbiosis that Dominates the Microbiome of the Coral Reef Sponge Amphimedon queenslandica. *Front. Mar. Sci.* **3**: 196.

Gloeckner, V., Wehrl, M., Moitinho-Silva, L., Gernert, C., Hentschel, U., Schupp, P., 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 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**: 108–10.

Graf, J. and Ruby, E.G. (1998) Host-derived amino acids support the proliferation of symbiotic bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **95**: 1818–22.

Green, E.R. and Mecsas, J. (2016) Bacterial Secretion Systems: An Overview. *Microbiol. Spectr.* **4**:.

Guindon, S. and Gascuel, O. (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst. Biol.* **52**: 696–704.

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.

Hentschel, U., Hopke, J., Horn, M., Anja, B., Wagner, M., Hacker, J., et al. (2002) Molecular Evidence for a Uniform Microbial Community in Sponges from Different Oceans Molecular Evidence for a Uniform Microbial Community in Sponges from Different Oceans. *Appl. Environ. Microbiol.* **68**: 4431–4440.

Hentschel, U., Piel, J., Degnan, S.M., and Taylor, M.W. (2012) Genomic insights into the marine sponge microbiome. *Nat. Rev. Microbiol.* **10**: 641–654.

Hess, W.R., Berghoff, B.A., Wilde, A., Steglich, C., and Klug, G. (2014) Riboregulators and the role of Hfq in photosynthetic bacteria. *RNA Biol.* **11**: 413–426.

Hoffmann, F., Larsen, O., Thiel, V., Rapp, H.T., Pape, T., Michaelis, W., and Reitner, J. (2005) An Anaerobic World in Sponges. *Geomicrobiol. J.* **22**: 1–10.

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.

Huang, Y., Gilna, P., and Li, W. (2009) Identification of ribosomal RNA genes in metagenomic fragments. *Bioinformatics* **25**: 1338–1340.

Huson, D.H., Beier, S., Flade, I., Górska, A., El-Hadidi, M., Mitra, S., et al. (2016) MEGAN Community Edition - Interactive Exploration and Analysis of Large-Scale Microbiome Sequencing Data. *PLOS Comput. Biol.* **12**: e1004957.

Hyatt, D., LoCascio, P.F., Hauser, L.J., and Uberbacher, E.C. (2012) Gene and translation initiation site prediction in metagenomic sequences. *Bioinformatics* **28**: 2223–2230.

Jack, R.W., Tagg, J.R., and Ray, B. (1995) Bacteriocins of gram-positive bacteria. *Microbiol. Rev.* **59**: 171–200.

Javelle, A., Thomas, G., Marini, A.-M., Krämer, R., and Merrick, M. (2005) In vivo functional characterization of the Escherichia coli ammonium channel AmtB: evidence for metabolic coupling of AmtB to glutamine synthetase. *Biochem. J.* **390**: 215–22.

Jwanoswki, K., Wells, C., Bruce, T., Rutt, J., Banks, T., and McNealy, T.L. (2017) The

Legionella pneumophila GIG operon responds to gold and copper in planktonic and biofilm cultures. *PLoS One* **12**: e0174245.

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.

Kanehisa, M. and Goto, S. (2000) KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **28**: 27–30.

Kanehisa, M., Sato, Y., and Morishima, K. (2016) BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. *J. Mol. Biol.* **428**: 726–731.

Kelly, D.J. and Thomas, G.H. (2001) The tripartite ATP-independent periplasmic (TRAP) transporters of bacteria and archaea. *FEMS Microbiol. Rev.* **25**: 405–424.

Knobloch, S., Jóhannsson, R., and Marteinsson, V. (2018) Bacterial diversity in the marine sponge Halichondria panicea from Icelandic waters and host-specificity of its dominant symbiont "Candidatus Halichondribacter symbioticus." *FEMS Microbiol. Ecol.*

Knowlton, A. and Highsmith, R. (2000) Convergence in the time-space continuum:a predator-prey interaction. *Mar. Ecol. Prog. Ser.* **197**: 285–291.

Kohler, C., Lourenço, R.F., Avelar, G.M., and Gomes, S.L. (2012) Extracytoplasmic function (ECF) sigma factor σF is involved in Caulobacter crescentus response to heavy metal stress. *BMC Microbiol.* **12**: 210.

Kumala, L. and Canfield, D.E. (2018) Contraction Dynamics and Respiration of Small Single-Osculum Explants of the Demosponge Halichondria panicea. *Front. Mar. Sci.* **5**: 410.

Kumala, L., Riisgård, H.U., and Canfield, D.E. (2017) Osculum dynamics and filtration activity studied in small single-osculum explants of the demosponge Halichondria panicea. *Mar Ecol Prog Ser* **572**: 117–128.

Kuwahara, H., Yoshida, T., Takaki, Y., Shimamura, S., Nishi, S., Harada, M., et al. (2007) Reduced Genome of the Thioautotrophic Intracellular Symbiont in a Deep-Sea Clam, Calyptogena okutanii. *Curr. Biol.* **17**: 881–886.

Langmead, B. and Salzberg, S.L. (2012) Fast gapped-read alignment with Bowtie 2. *Nat Meth* **9**: 357–359.

Leimbach, A. (2016) bac-genomics-scripts: Bovine E. coli mastitis comparative genomics edition.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079.

Liu, M.Y., Kjelleberg, S., and Thomas, T. (2011) Functional genomic analysis of an uncultured δ-proteobacterium in the sponge Cymbastela concentrica. *ISME J.* **5**: 427–435.

Ludwig, W., Strunk, O., Westram, R., Richter, L., Meier, H., Yadhukumar, et al. (2004) ARB: a software environment for sequence data. *Nucleic Acids Res.* **32**: 1363–1371.

Luo, H., Csuros, M., Hughes, A.L., and Moran, M.A. (2013) Evolution of divergent life history strategies in marine alphaproteobacteria. *MBio* **4**: e00373-13.

Magoc, T. and Salzberg, S.L. (2011) FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics* **27**: 2957–2963.

McCutcheon, J.P., McDonald, B.R., and Moran, N.A. (2009) Convergent evolution of metabolic roles in bacterial co-symbionts of insects. *Proc. Natl. Acad. Sci. U. S. A.* **106**: 15394–9.

McCutcheon, J.P. and Moran, N.A. (2012) Extreme genome reduction in symbiotic bacteria. *Nat. Rev. Microbiol.* **10**: 13–26.

McMurdie, P.J. and Holmes, S. (2013) phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One* **8**: e61217.

Merrick, M.J. and Edwards, R.A. (1995) Nitrogen control in bacteria. *Microbiol. Mol. Biol. Rev.* **59**:.

Mohamed, N.M., Saito, K., Tal, Y., and Hill, R.T. (2010) Diversity of aerobic and anaerobic ammonia-oxidizing bacteria in marine sponges. *ISME J.* **4**: 38–48.

Moitinho-Silva, L., Nielsen, S., Amir, A., Gonzalez, A., Ackermann, G.L., Cerrano, C., et al. (2017) The sponge microbiome project. *Gigascience* **6**: 1–7.

Moran, N.A., Dunbar, H.E., and Wilcox, J.L. (2005) Regulation of Transcription in a Reduced Bacterial Genome: Nutrient-Provisioning Genes of the Obligate Symbiont Buchnera aphidicola. *J. Bacteriol.* **187**: 4229–4237.

Moya, A., Peretó, J., Gil, R., and Latorre, A. (2008) Learning how to live together: genomic insights into prokaryote–animal symbioses. *Nat. Rev. Genet.* **9**: 218–229.

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.

Nedashkovskaya, O.I., Kukhlevskiy, A.D., Zhukova, N. V., and Kim, S.B. (2016) Amylibacter ulvae sp. nov., a new alphaproteobacterium isolated from the Pacific green alga Ulva fenestrata. *Arch. Microbiol.* **198**: 251–256.

Newton, R.J., Griffin, L.E., Bowles, K.M., Meile, C., Gifford, S., Givens, C.E., et al. (2010) Genome characteristics of a generalist marine bacterial lineage. *ISME J.* **4**: 784–798.

Nguyen, M.T.H.D., Liu, M., and Thomas, T. (2014) Ankyrin-repeat proteins from sponge symbionts modulate amoebal phagocytosis. *Mol. Ecol.* **23**: 1635–1645.

Nurk, S., Meleshko, D., Korobeynikov, A., and Pevzner, P.A. (2017) metaSPAdes: a new versatile metagenomic assembler. *Genome Res.* **27**: 824–834.

Ottesen, E.A., Marin, R., Preston, C.M., Young, C.R., Ryan, J.P., Scholin, C.A., and DeLong, E.F. (2011) Metatranscriptomic analysis of autonomously collected and preserved marine bacterioplankton. *ISME J.* **5**: 1881–1895.

Pallen, M.J., Chaudhuri, R.R., and Henderson, I.R. (2003) Genomic analysis of secretion systems. *Curr. Opin. Microbiol.* **6**: 519–527.

Peattie, M.E. and Hoare, R. (1981) The sublittoral ecology of the Menai Strait: II. The sponge Halichondria panicea (Pallas) and its associated fauna. *Estuar. Coast. Shelf Sci.* **13**: 621–635.

Pita, L., Fraune, S., and Hentschel, U. (2016) Emerging Sponge Models of Animal-Microbe Symbioses. *Front. Microbiol.* **7**: 2102.

R Core Team (2017) A language and environment for statistical computing. In, *R foundation for Statistical Computing*. Vienna, Austria.

Reizer, J., Hoischen, C., Titgemeyer, F., Rivolta, C., Rabus, R., Stülke, J., et al. (1998) A novel protein kinase that controls carbon catabolite repression in bacteria. *Mol. Microbiol.* **27**: 1157–1169.

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.

Schmitz-Esser, S., Tischler, P., Arnold, R., Montanaro, J., Wagner, M., Rattei, T., and Horn, M. (2010) The genome of the amoeba symbiont "Candidatus Amoebophilus asiaticus" reveals common mechanisms for host cell interaction among amoeba-associated bacteria. *J. Bacteriol.* **192**: 1045–57.

Shigenobu, S. and Wilson, A.C.C. (2011) Genomic revelations of a mutualism: the pea aphid and its obligate bacterial symbiont. *Cell. Mol. Life Sci.* **68**: 1297–1309.

Siegl, A. and Hentschel, U. (2010) PKS and NRPS gene clusters from microbial symbiont cells of marine sponges by whole genome amplification. *Environ. Microbiol. Rep.* **2**: 507–513.

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. (2012) Sponge-specific clusters revisited: a comprehensive phylogeny of sponge-associated microorganisms. *Environ. Microbiol.* **14**: 517–524.

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**: 2465–2478.

Soupene, E., He, L., Yan, D., and Kustu, S. (1998) Ammonia acquisition in enteric bacteria: physiological role of the ammonium/methylammonium transport B (AmtB) protein. *Proc. Natl. Acad. Sci. U. S. A.* **95**: 7030–4.

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.

Swofford, D.L. and Maddison, W.P. (1987) Reconstructing ancestral character states under Wagner parsimony. *Math. Biosci.* **87**: 199–229.

Tempest, D.W., Meers, J.L., and Brown, C.M. (1970) Synthesis of glutamate in Aerobacter aerogenes by a hitherto unknown route. *Biochem. J.* **117**: 405–7.

Teramoto, M. and Nishijima, M. (2014) Amylibacter marinus gen. nov., sp. nov., isolated from surface seawater. *Int. J. Syst. Evol. Microbiol.* **64**: 4016–4020.

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.

Thomas, T., Rusch, D., DeMaere, M.Z., Yung, P.Y., Lewis, M., Halpern, A., et al. (2010) Functional genomic signatures of sponge bacteria reveal unique and shared features of symbiosis. *ISME J.* **4**: 1557–1567.

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.

Tseng, T.-T., Tyler, B.M., and Setubal, J.C. (2009) Protein secretion systems in bacterial-host associations, and their description in the Gene Ontology. *BMC Microbiol.* **9**: S2.

Vacelet, J. and Donadey, C. (1977) Electron microscope study of the association between some sponges and bacteria. *J. Exp. Mar. Bio. Ecol.* **30**: 301–314.

Waldron, F.M., Stone, G.N., and Obbard, D.J. (2018) Metagenomic sequencing suggests a diversity of RNA interference-like responses to viruses across multicellular eukaryotes. *PLoS Genet.* **14**: e1007533.

Wang, D., Wei, Y., Cui, Q., and Li, W. (2017) Amylibacter cionae sp. nov., isolated from the sea squirt Ciona savignyi. *Int. J. Syst. Evol. Microbiol.* **67**: 3462–3466.

Warnes, G.R., Bolker, B., Bonebakker, L., Gentleman, R., Liaw, W.H.A., Lumley, T., et al. (2016) gplots: various R programming tools for plotting data. R package version 3.0. 1. *Compr. R Arch. Netw.*

Webster, N.S. and Thomas, T. (2016) The Sponge Hologenome. *MBio* **7**: e00135-16.

Wichels, A., Würtz, S., Döpke, H., Schütt, C., and Gerdts, G. (2006) Bacterial diversity in the breadcrumb sponge Halichondria panicea (Pallas). *FEMS Microbiol. Ecol.* **56**: 102–118.

Wickham, H. (2009) ggplot2: Elegant Graphics for Data Analysis Springer-Verlag. *New York*.

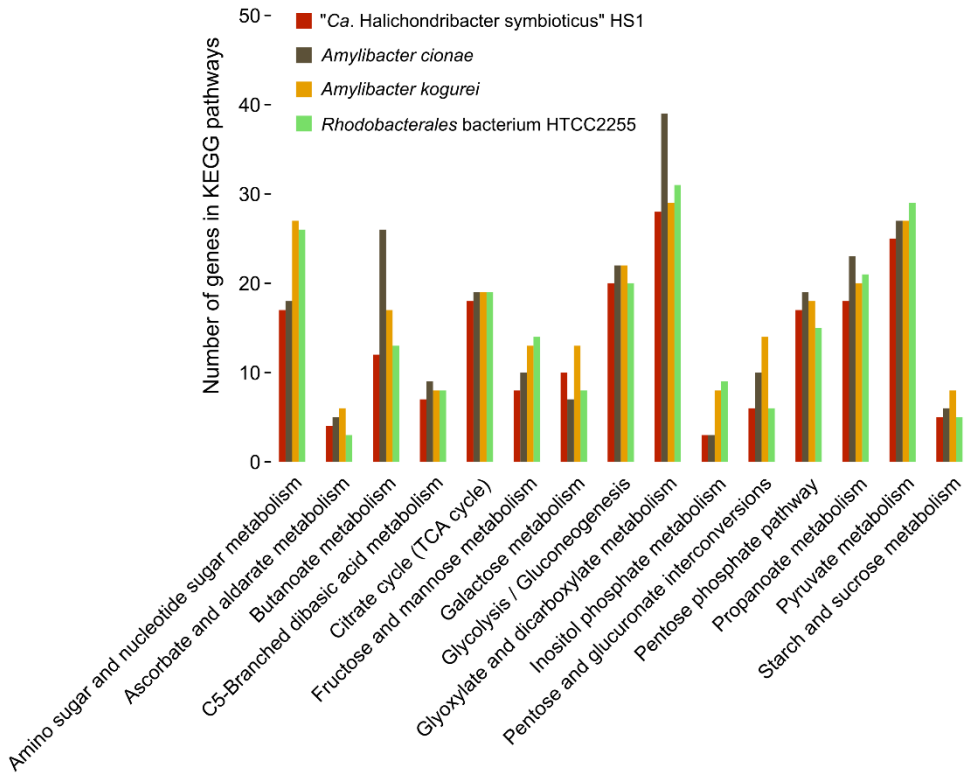
Wong, S.-K., Yoshizawa, S., Nakajima, Y., Cuadra, M.J., Nogi, Y., Nakamura, K., et al. (2018) Amylibacter kogurei sp. nov., a novel marine alphaproteobacterium isolated from the coastal sea surface microlayer of a marine inlet. *Int. J. Syst. Evol. Microbiol.*

Wu, M. and Eisen, J.A. (2008) A simple, fast, and accurate method of phylogenomic inference. *Genome Biol.* **9**: R151.

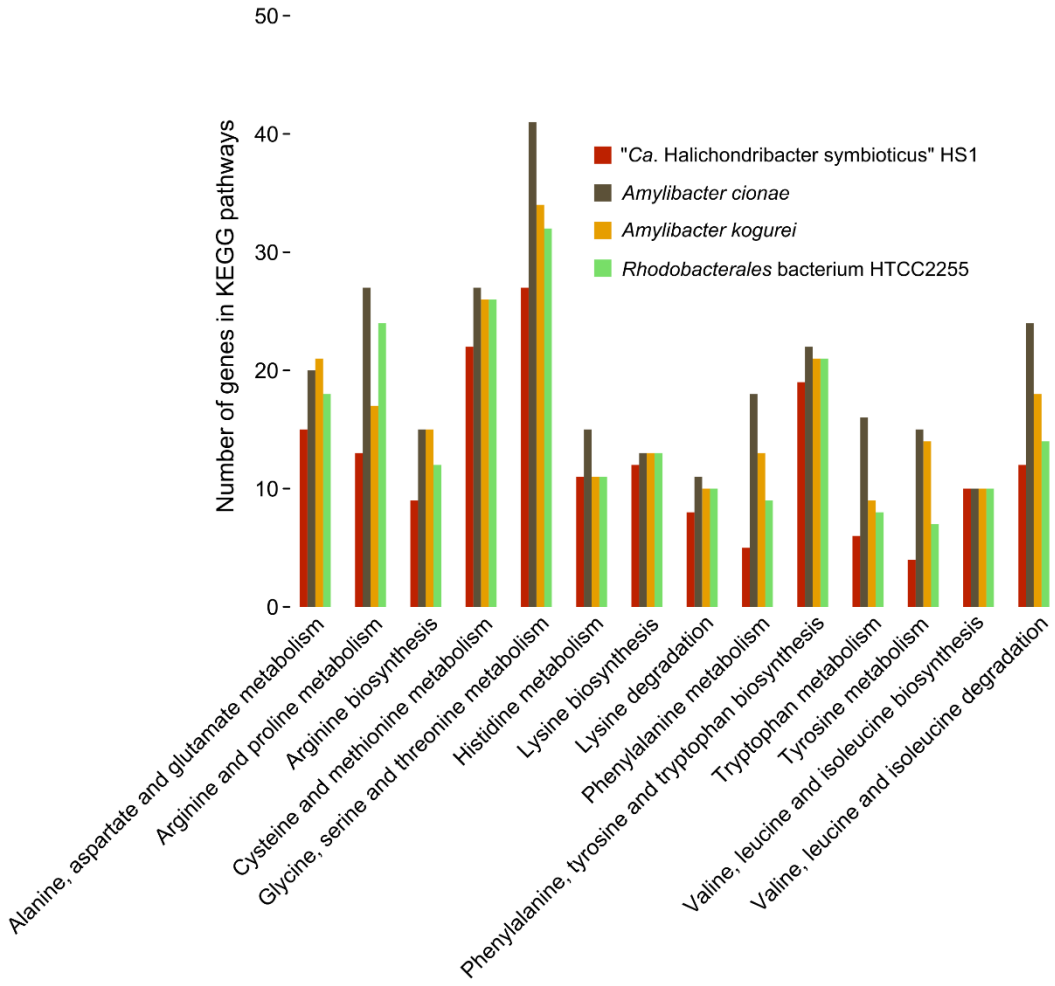
Wu, M. and Scott, A.J. (2012) Phylogenomic analysis of bacterial and archaeal sequences with AMPHORA2. *Bioinformatics* **28**: 1033–1034.

Yilmaz, P., Parfrey, L.W., Yarza, P., Gerken, J., Pruesse, E., Quast, C., et al. (2014) The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res.* **42**: D643–D648.

Yin, Z., Zhu, M., Davidson, E.H., Bottjer, D.J., Zhao, F., and Tafforeau, P. (2015) Sponge grade body fossil with cellular resolution dating 60 Myr before the Cambrian. *Proc. Natl. Acad. Sci. U. S. A.* **112**: E1453-60.



Supplementary Figure S1: *Number of genes in KEGG pathways for carbohydrate metabolism of "*Ca. *H. symbioticus" HS1 and related reference genomes.*



Supplementary Figure S2: *Number of genes in KEGG pathways for amino acid metabolism of "*Ca. *H. symbioticus" HS1 and related reference genomes.*

Supplementary Tables S1 – S4 are not added to the Annex due to their size but can be requested from authors and have been submitted along with the main manuscript for peer-review.

Paper III

Co-cultivation of the marine sponge *Halichondria panicea* and its associated microorganisms

Stephen Knobloch, Ragnar Jóhannsson, Viggó Þór Marteinsson

Submitted to *Scientific Reports*

Author contributions:

Conceived and designed the study: SK, RJ, VM
Performed sampling: SK
Performed experiments: SK
Performed laboratory work: SK
Analysed data: SK
Wrote the manuscript: SK
Reviewed and edited the manuscript: SK, RJ, VM
Correspondence (publication process): SK

Abstract

Marine sponges host bacterial symbionts with biotechnological potential, yet isolation of true sponge symbionts remains difficult due to their host dependency. Moreover, attempts to grow sponges for their pharmacologically-active compounds outside of their habitat often results in a shift of their microbial community. In this study we evaluate suitable sponge cultivation methods that allow maintenance of both the marine sponge *Halichondria panicea* and its associated bacteria in an *ex situ* environment. In addition, we present a method for co-cultivation of sponge explants and microbes separated by a membrane in a multi-chamber device. Tests on *ex situ* cultivation of *H. panicea* under different controlled conditions showed that only high water exchange rates in the aquarium enabled maintenance of its dominant symbiont "*Candidatus* Halichondribacter symbioticus" at a high relative abundance in the sponge body, a prerequisite for co-cultivation. The bacterial enrichment retrieved from co-cultivation contained bacteria from nine different classes in addition to sequences corresponding to "*Ca*. H. symbioticus". This represents an increase of the cultivable bacterial classes from *H. panicea* compared to standard isolation techniques on solid media plates. The current study provides insights into sponge-microbe maintenance under *ex situ* conditions and proposes a new method for the isolation of sponge-associated bacteria.

Introduction

Marine sponge (phylum *Porifera*) have been in the spotlight of marine derived natural product discovery for the past decades due to their rich inventory of pharmacologically-active compounds [1–3]. Despite this potential, sourcing sufficient quantities of a promising compound remains difficult due to the low concentrations of these chemicals in the sponge body. This issue, termed the "supply problem", limits the advance of preclinical trials and hinders exploitation of sponges for other biotechnological purposes [4]. Apart from compounds derived from the sponge itself, it has been shown that sponge-associated microorganisms can be the source of bioactive compounds with pharmaceutically interesting properties [5,6]. Isolating these sponge-associated microbes could therefore circumvent the "supply problem" [7]. Sponges contain diverse and highly specific microbial communities which can contribute to functional attributes of the sponge holobiont [8]. These include obligate sponge symbionts as well as opportunistic and generalistic sponge-associated bacteria [9].

It is estimated that less than 1 % of the bacterial diversity on earth is cultivable using traditional cultivation methods [10]. This has spawned new isolation and cultivation techniques ranging from microdroplet cultivation [11] to *in situ* cultivation using new devices such as the ichip [12] and other diffusion growth chambers [13]. Diffusion growth chambers consist of membranes with pore sizes small enough to trap microorganisms inside a chamber while allowing sufficient diffusion of metabolites between the chamber and the outside environment. The chambers are then inoculated with single microbial cells or microbial suspensions and are deployed back into their natural environment, thus providing the complex environmental conditions needed for growth [14]. Such systems have been, most notably, applied for the isolation of soil microorganisms [12,15], but have also been implanted in wild sponges, yielding novel so far uncultivated microorganisms [16]. Whereas this is a promising approach, higher throughput co-cultivation of sponges and microorganisms using growth diffusion chambers requires both the proximity to the sponge and easy accessibility to the experiment. E*x situ* cultivation of sponges provides easier access and, in addition, allows for a control of all environmental parameters [17]. However, sponges have been reported to undergo a shift in their microbial diversity when cultured under *ex situ* conditions [18], potentially detrimental to the isolation of sponge-associated microorganisms through a co-cultivation approach.

The marine sponge *Halichondria panicea* has been used as a model sponge species for a variety of sponge physiological experiment [19–22] and is well studied in terms of its biology and ecology in coastal areas [23–28]. Previous studies have found various bioactivities in extract of *H. panicea* and in microbes isolated from its body [29–34], including extracts with cytotoxic properties [35,36]. *H. panicea* from various locations in the North Atlantic host the dominant bacterial symbiont "*Candidatus* Halichondribacter symbioticus" which often constitutes over half of the relative bacterial abundance in the sponge body [36–40]. Other bacteria within the sponge body vary depending on geographical location and time of sampling [36,38]. Despite variability they exhibit a clear difference in diversity from the surrounding seawater bacterial community, showing that, apart from "*Ca*. H. symbioticus", *H. panicea* hosts other sponge-associated bacteria [37]. Bacterial isolation studies with *H. panicea* using standard cultivation methods have recovered strains associated with the taxa *Actinobacteria*, *Alpha-*, *Beta-* and *Gammaproteobacteria*, *Deinococcus*, *Cytophagia*, *Flavobacteria*, *Bacteroidetes* and low G+C Gram-positive bacteria [38,41], including many *Actinobacteria* with bioactive properties [42,43]. To date, the dominant symbiont "*Ca*. H. symbioticus" has not been cultivated outside

of its hosts. Being able to cultivate obligate sponge symbionts outside of, but in close proximity to, their host could allow valuable possibilities to study the mechanisms that lead to their host-dependency. Increasing the range of cultivable sponge-associated bacteria in general would be beneficial for biotechnological applications.

In this study we evaluate suitable sponge cultivation methods which allow maintenance of both the sponge and its associated bacteria in an *ex situ* environment, and present a new sponge-bacteria co-cultivation method for isolation of sponge-associated bacteria which can be used under laboratory setting. We apply 16S rRNA gene amplicon sequencing along with standard plating techniques to analyse the bacterial diversity in *H. panicea* during cultivation and compare the enriched bacterial fraction from co-cultivation to the results from standard plating.

Results

Ex situ cultivation of *Halichondria panicea*

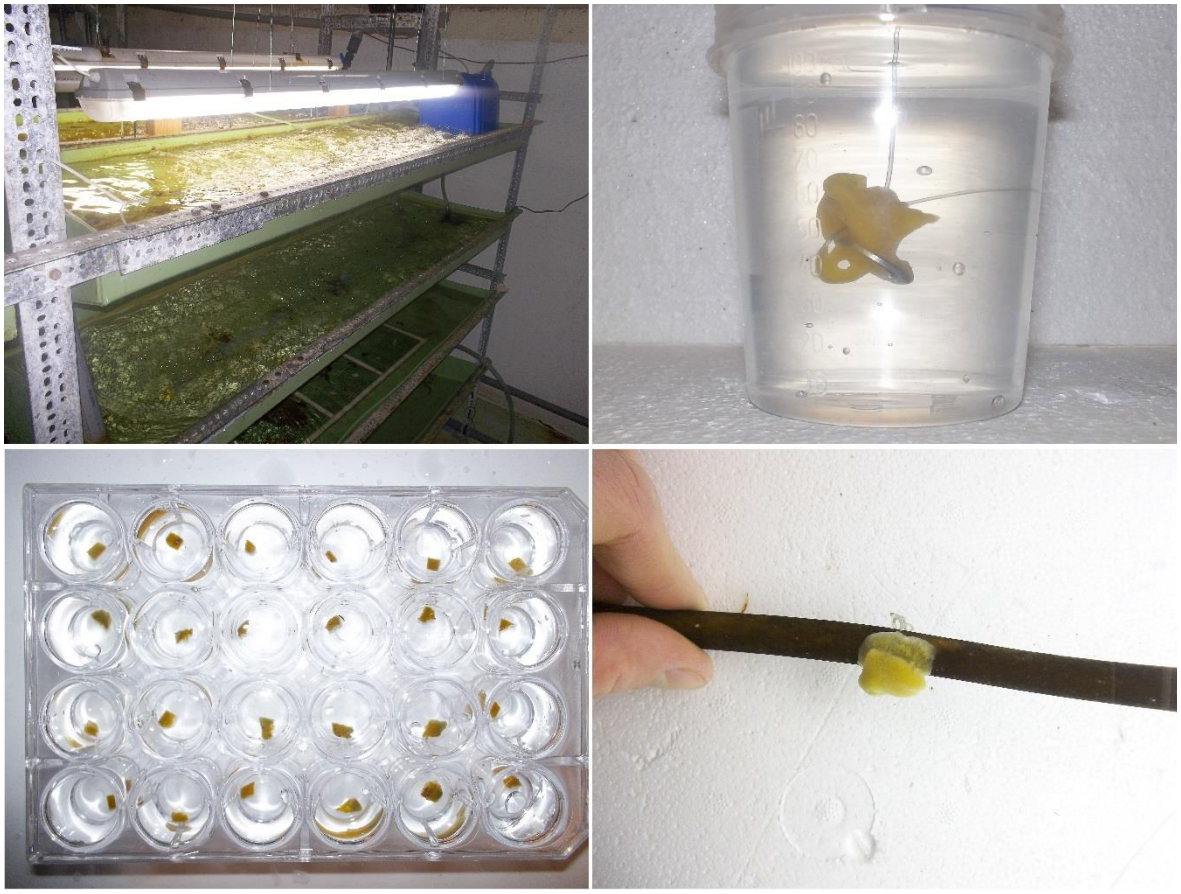


Fig. 1 *(top left, clockwise) Cultivation system for* H. panicea *in method 3 and 4; Sponge explant attached to nylon string and metal washer in cultivation method 2; Sponge explant attached to the thallus of Laminaria sp. in method 4; Sponge explants distributed in 24-well plate before being deployed in aquarium in method 5.*

To evaluate suitable methods for *ex situ* sponge cultivation five cultivation experiments were conducted with different aquarium set-ups (Fig. 1). Attachment to substrate, active pumping, growth as surface area expansion and survival were measured over the course of each

experiment. Additionally, samples before and after each experiment were collected for microbial diversity analysis through 16S rRNA gene amplicon sequencing to determine shifts of the sponge-associated microbial diversity and the relative abundance of the dominant bacterial symbiont "*Ca*. Halichondribacter symbioticus".

Attachment of explants to the substrate was visible within the first two weeks of the experiments, apart from in method 1 which used recirculated water and where over half of the explants showed no signs of attachment and eventually detached from the nylon string. Fastest attachment was visible in method 5 where all surviving explants attached to the surface within the first week of the experiment. Survival was above 50 % in all experiments (Table 1) apart from method 1 where of the initial 9 explants 7 detached from the substrate or decreased largely in size, eventually leading to mortality. In method 2, 3, 4 and 5 survival was 63, 75, 100 and 63 % of the explants or sponge individuals over the course of each experiment respectively.

Table 1 *Description of methods tested for ex situ cultivation of* H. panicea*; NA: not applicable (sponges in method 3 were not removed from their original substrate); † supplied from sand-filtered coastal seawater; * supplied from sand-filtered borehole seawater passed through a 3000 l public aquarium tank; +, positive; -, negative*

Method	**Tank volume (l)**	**Feeding**	**Water exchange (tank volume d^{-1})**	**Duration**	**Substrate**	**Explant size**	**Attachment**	**Survival (>50%)**
Method 1 (recirculating system)	60	yes	0	20 weeks	Nylon string	1 cm^3	-	-
Method 2 (flow-through)	60	yes	10†	20 weeks	Nylon string	1 cm^3	+	+
Method 3 (semi-recirculating)	360	no	2†	25 weeks	Rocks	NA	NA	+
Method 4 (semi-recirculating)	1360	no	4†	20 weeks	*Laminaria* sp.	1 cm^3	+	+
Method 5 (flow-through)	15	no	720*	18 weeks	24-well plate (PS)	0.125 cm^3	+	+

Within the first four weeks of all experiments, most sponges rearranged their aquiferous system and a single osculum, in the case of sponge explants, or multiple new oscula, in the case of sponge individuals kept on their original substrate, emerged and exhibited pumping activity. In method 5 oscula emerged earliest, within the first two weeks of the experiment, whereas in method 1 oscula on most explants were not observable. From all individuals used in the cultivation experiments growth, measured as an increase of the surface area, could only be observed in a single sponge individual in method 3. In most cases, sponge size did not vary between the beginning and end of the experiments and in some cases, especially in method 1, sponge size decreased during cultivation.

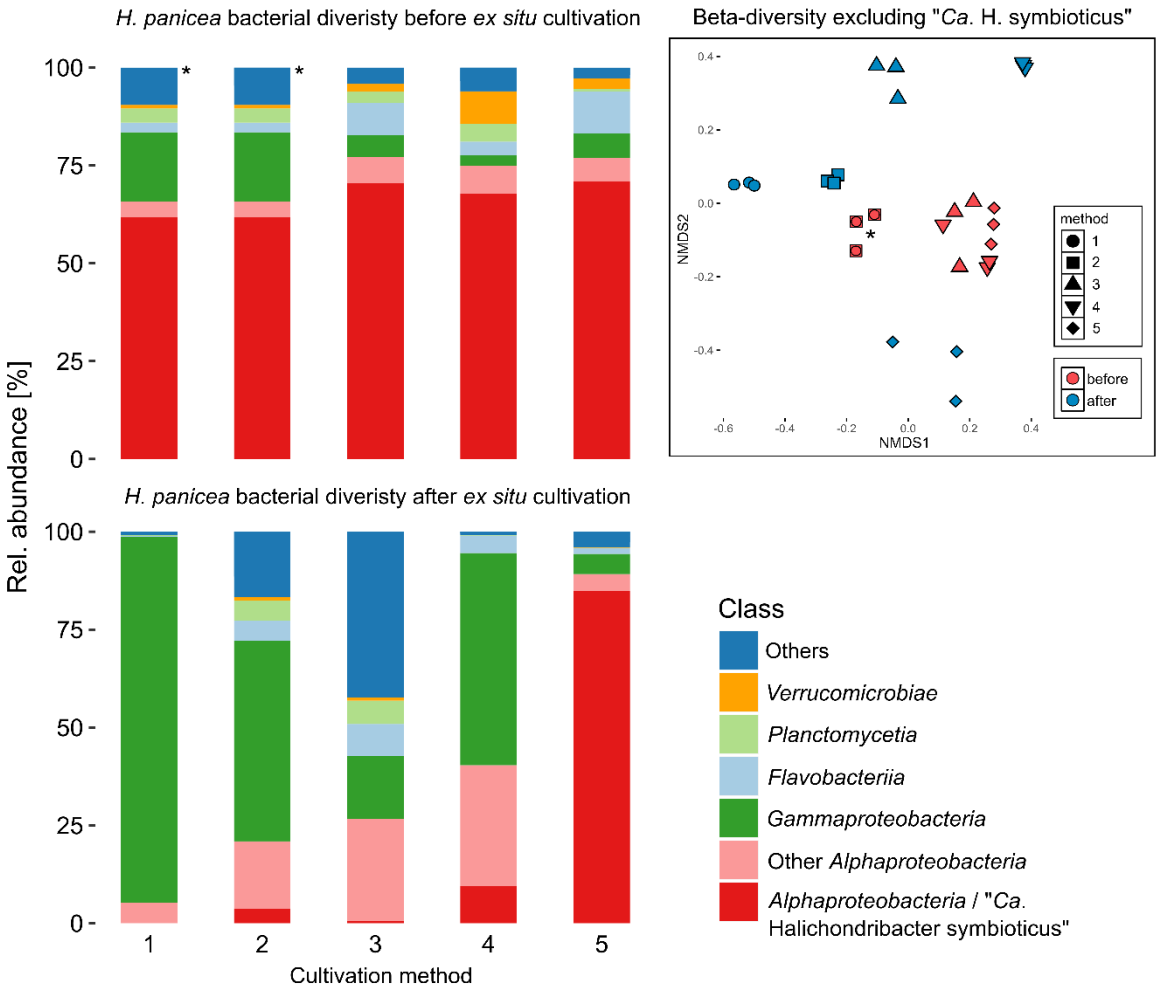


Fig. 2 *Relative abundance of the bacterial diversity in* H. panicea *before and after ex situ cultivation using five different cultivation methods. Each bar represents the mean relative abundance of three sponge samples or explants. Classes representing less than 0.7 % of the relative read abundance in the total dataset and unclassified taxa are grouped as "Others". Top right: NMSD Ordination plot of Bray-Curtis dissimilarity matrix. The same initial sponge was used for method 1 and 2, therefore sample sets marked with an asterisk are identical. Time between sample sets in weeks: 20 (method 1), 20 (method 2), 25 (method 3), 20 (method 4) and 18 (method 3).*

The bacterial diversity of wild *H. panicea* collected for all experiments contained taxa predominantly assigned to the classes *Alphaproteobacteria*, *Gammaproteobacteria*, *Flavobacteriia*, *Planctomycetia* and *Verrucomicrobiae* (Fig. 2) and, to a lesser degree, to 28 other bacterial and 4 archaeal classes. In addition, the dominant symbiont "*Ca*. H. symbioticus" made up over 60 % of the relative abundance in all samples. Although most sponges in methods 2 to 5 continued pumping and maintained their size over the course of the experiments, the bacterial diversity showed a marked shift between the beginning and end of the experiments (Fig. 2). The dominant symbiont "*Ca*. H. symbioticus" was reduced to less than 10 % in samples collected in methods 1 to 4. In method 1 the relative abundance was even reduced to an average of 0.004 % of the relative abundance. "*Ca*. H. symbioticus" only remained the dominant bacterial taxon in method 5 with an average of 85 % of the

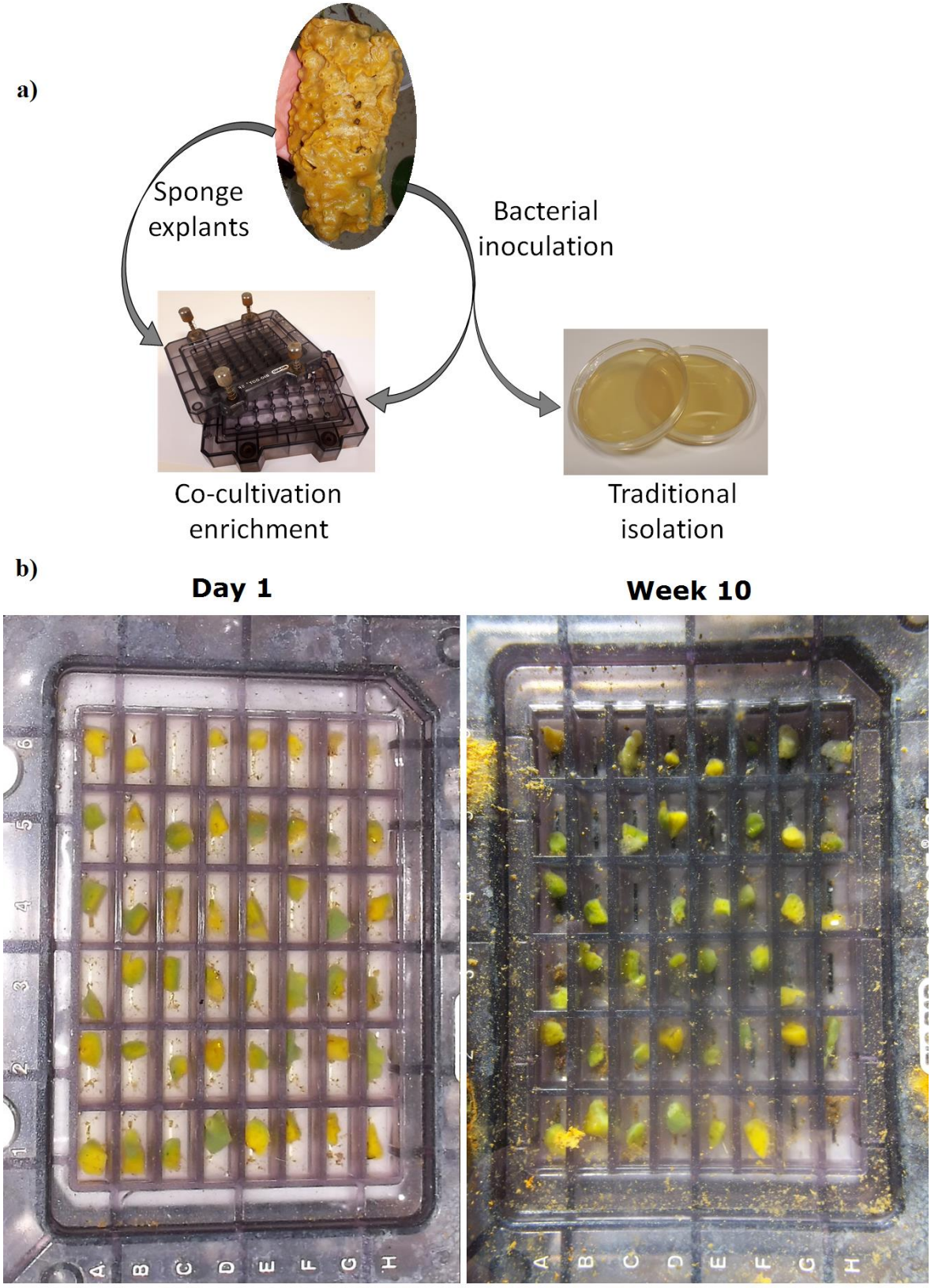


Fig.3 *(a) Schematically diagram of sponge-bacteria co-cultivation experiment; (b) sponge explants in co-cultivation device at day 1 and after week 10. Sponge explants collected for 16S rRNA gene amplicon sequencing and deceased explants were not replaced after the first week of the experiment*

relative microbial abundance. After excluding "*Ca*. H. symbioticus" from the dataset the remaining bacterial diversity in the samples after cultivation appeared more dissimilar to the diversity of the sample before cultivation of each respective method with the exception of the bacterial diversity in method 2 (Fig. 2). In line with this finding and depending on the cultivation method, between 68 and 90 % of the OTUs detected in the initial sponge samples were not detected in the samples after the experiments, showing that the majority of the initial bacterial taxa were lost during sponge cultivation.

Bacterial diversity of sponge explants during co-cultivation

Based on the previous trials, method 5 was chosen for the sponge-bacteria co-cultivation experiment. *H. panicea* explants were cut and distributed into 48 wells of a Bio-Dot SF microfiltration apparatus (BioHit), separated by a 0.2 µm membrane from bacterial inoculations generated through serial dilutions from the same sponge individual (Fig. 3). Within the first week of deployment in the aquarium set-up detailed in method 5, *H. panicea* explants regenerated the cut surface and started attaching to the substrate of the co-cultivation device. Displaced explants or those showing signs of necrosis were replaced within the first week. During the following ten weeks four explants died and were not replaced (Fig. 4). After ten weeks all remaining explants had developed at least one osculum which exhibited pumping activity when tested by placing food dye above the exhalent jet.

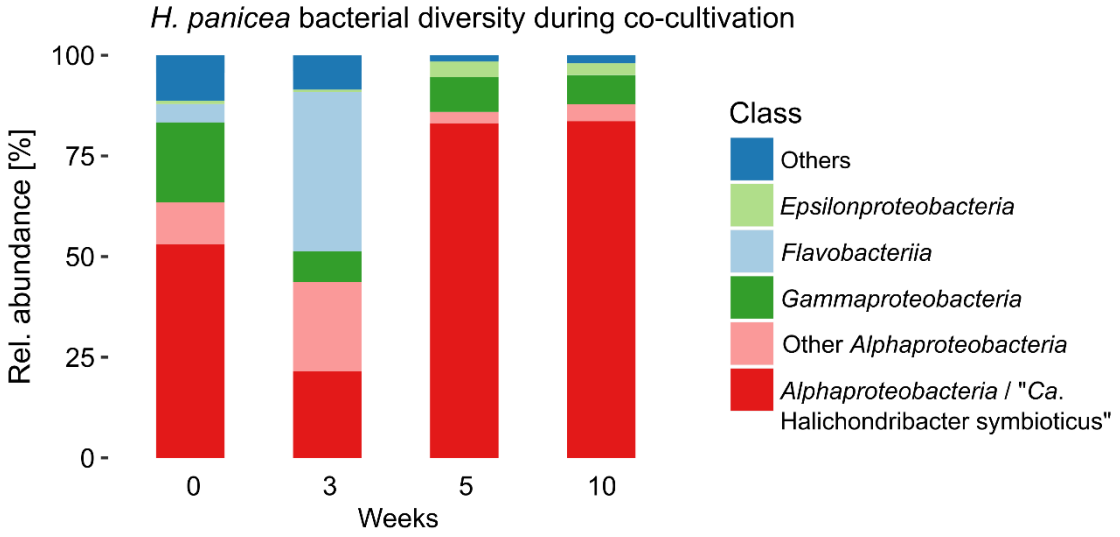


Fig. 4 *Relative abundance of the bacterial diversity in* H. panicea *at the start of the co-cultivation experiment and in* H. panicea *explants after 3, 5 and 10 weeks. Taxa are grouped by class level taxonomical association and classes representing less than 0.7 % of the relative read abundance in the total dataset and unclassified taxa are grouped as "Others".*

A sample from the original sponge used for bacterial inoculation as well as three healthy looking explants, removed after three, five and ten weeks, were subjected to 16S rRNA gene amplicon sequencing. The dominant bacterial taxon, corresponding to the symbiont "*Ca*. H. symbioticus", represented approximately 53 % of the relative abundance in the sponge collected from the wild at the beginning of the experiment (Fig. 4). After three weeks the relative abundance of "*Ca*. H. symbioticus" had dropped to 22 % of the relative abundance, whereas OTUs assigned to the class *Flavobacteriia* and other *Alphaproteobacteria* were more abundant than at day 0. In the explants analysed after five and ten weeks the relative

abundance of "*Ca*. H. symbioticus" had increased to 83 % and 84 % of the relative abundance respectively. Similar to the previous experiments, 67 % of the OTUs found in the sponge body before the start of the experiment were not detected in the explants anymore. Despite this shift in diversity, the microbial community in the explants was more similar to the original sponge sample than to the water samples collected from the tank at the same times, shown through hierarchical clustering of Bray-Curtis dissimilarities (Fig. 5).

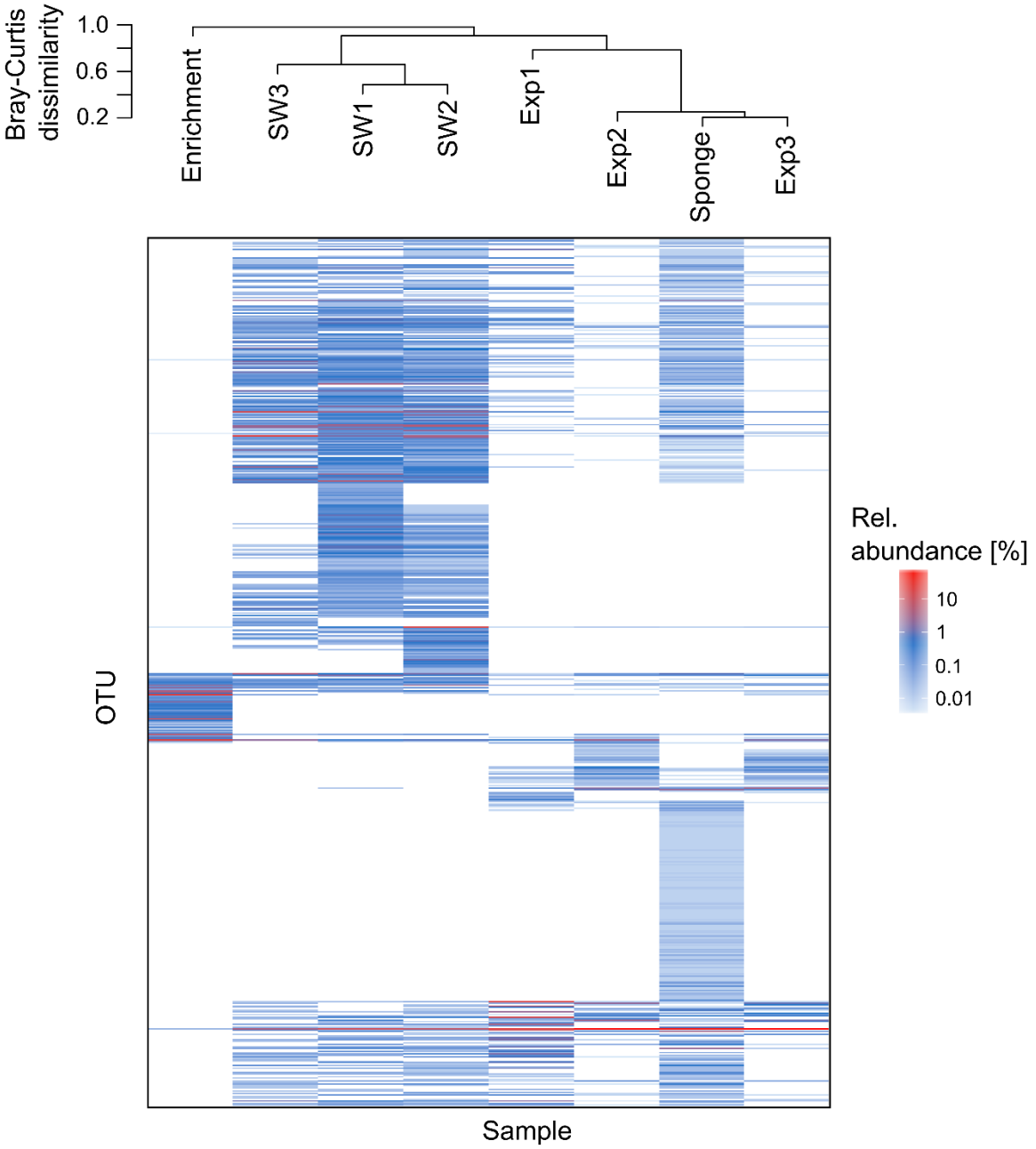


Fig.5 *OTU heatmap showing the presence and relative abundance of OTUs In the original sponge sample used for the co-cultivation experiment ("Sponge"), the sponge explants ("Exp"), seawater samples from the co-cultivation aquarium ("SW") and the bacterial enrichment pellet ("Enrichment"); Samples were ordered by hierarchical clustering of Bray-Curtis dissimilarities and method "average" of the hclust function in R.*

Comparison of bacterial diversity between co-cultivation enrichment and traditional isolation

After ten weeks the co-cultivation chamber was removed from the aquarium and the bacterial enrichment in form of visible pellets on the membrane was collected, pooled and subjected to 16S rRNA gene amplicon sequencing. The enrichment contained 61 OTUs in nine different bacterial classes, most of which assigned to the *Proteobacteria* (*Alpha-*, *Gamma-*, *Delta-*, *Epsilon-*) and *Clostridia* (Fig. 6). Only 14 OTUs in the enrichment were also detected in the original sponge used for bacterial inoculation (Fig. 5 and see Supplementary Table S1 for list of enrichment OTUs). In addition, these were mainly low abundant bacteria with less than 1 % of the read abundance in the original sponge, apart from "*Ca*. H. symbioticus", designated as Otu1, of which eight reads were detected in the enrichment. A comparison against the NCBI nr/nt database showed that 80 % of the OTUs showed closest sequence similarity to samples from marine origin and 33 % to samples from marine invertebrate.

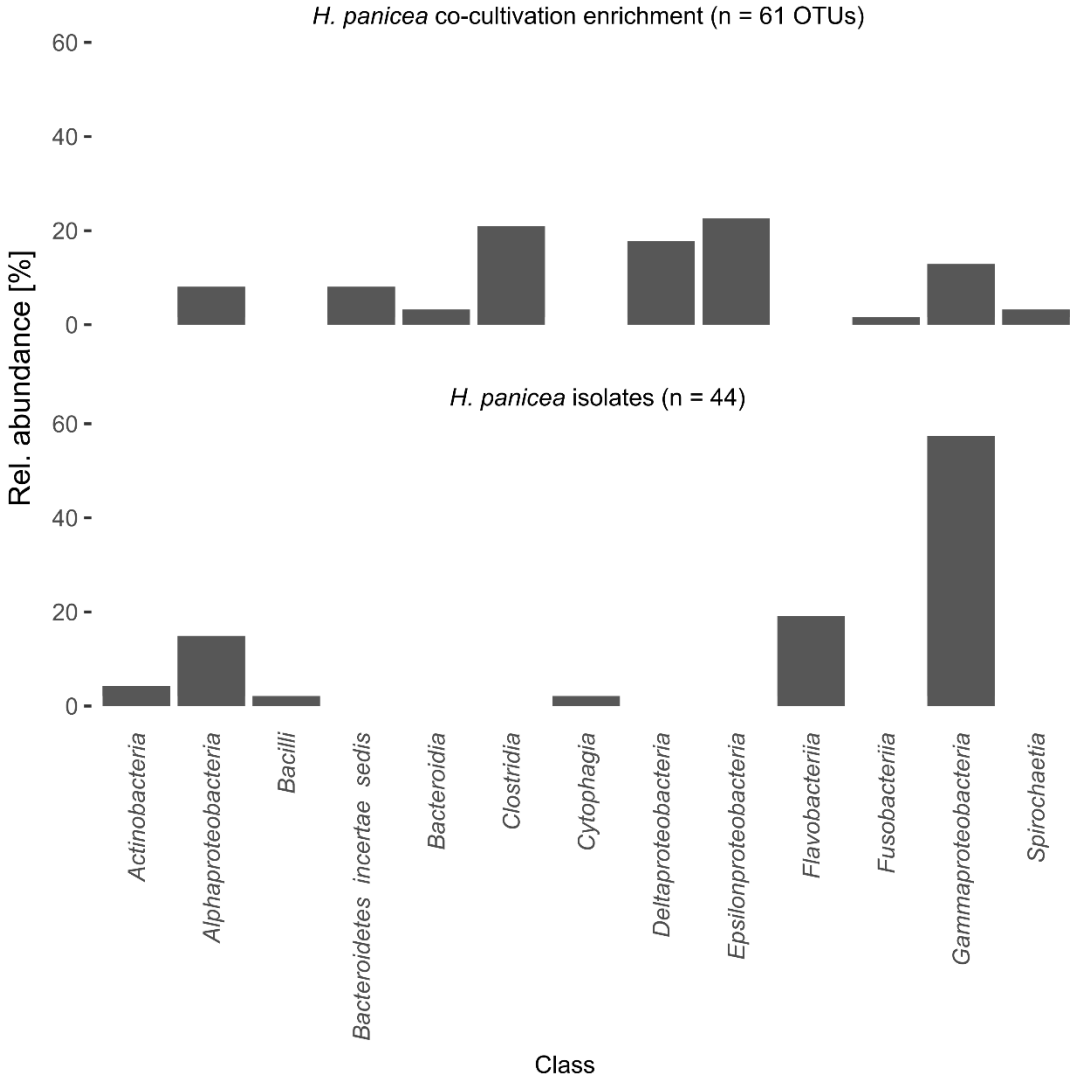


Fig. 6 *Relative abundance of OTUs enriched through sponge-bacterial co-cultivation (top) and strains isolated through standard plating method (bottom) by class level assignment*.

Using the traditional plating method, 253 colonies were isolated and screened using a MALDI-TOF biotyper. Out of these, 44 isolates showed a peptide mass fingerprint with sufficient dissimilarity to be classified as unique species and were selected for sequencing of the 16S rRNA gene. The Isolates were assigned to six different classes, whereas the majority of isolates belonged to the *Gammaproteobacteria* (Fig. 6). Similar to the co-cultivation enrichment, only 10 isolates showed high sequence similarity (> 97.0 %) to OTUs from the original inoculant and these were mostly low diversity OTUs (see Supplementary Table S1 for list of isolates). 84 % of all isolates showed highest sequence similarity to samples isolated from marine habitats in the NCBI nr/nt database, and 34 % were from marine invertebrates. In addition, two isolates showed highest sequence similarity to sequences previously isolated from *H. panicea*.

Discussion

Five methods for *ex situ* cultivation of the sponge *H. panicea* were tested of which four showed promising results for maintaining *H. panicea* explants over the course of several weeks. The only method which was not suitable due to high mortality was the recirculating system in method 1. Although feed was provided through two microalgae strains and toxic build-up of ammonia and nitrite was prevented through biofiltration, the size of the sponge explants decreased over time and high mortality occurred. Previous studies have shown that *H. panicea* can utilise cells such as the microalgae *Rhodomonas* sp. as feed source with high efficiency [22,44]. The microalgae provided as feed in this study were in a similar size range to *Rhodomonas* sp. and could thus be regarded as suitable feed for the sponge. However, in regards to the carbon requirements of *H. panicea* [45], insufficient feeding might have led to a decrease of explant size and eventually mortality. Further research needs to be conducted to evaluate suitable conditions for long-term cultivation of *H. panicea* in fully recirculating systems.

Despite successful maintenance of *H. panicea* using the other cultivation methods, only method 5 was suitable to maintain both the sponge and its dominant symbiont at a high relative abundance within the sponge body. The main difference between this method and the other methods was the high rate of water exchange within the tank, the water source and the size of the explants. As the water in method 5 was derived from the outlet of a public aquarium display tank and was supplied at a high rate compared to the sponge cultivation tank volume and the explant sizes, it is conceivable that enough feed was indirectly provided through the public aquarium to sufficiently nurture the sponge which, in turn, had a positive effect on the abundance of its symbiont. This is consistent with previous reports that report frequently detection of sponges, often unintentionally selected, in public aquaria, proving that they provide suitable cultivation conditions [4,46]. Conversely, a lack of feed provided by the other methods might have sustained the sponge over the cultivation period but triggered a reduction of its main symbiont.

Apart from "*Ca*. H. symbioticus", the remaining microbial diversity in the sponge body underwent a shift in its composition with a loss of most of the original bacterial taxa at the end of the cultivation period, regardless of the method applied. It has previously been reported that sponge kept under *ex situ* conditions undergo a shift in their microbial community [47–49] showing that changes in environmental conditions can have an impact on the sponge-associated microbial community. These changes could be due to the loss of

transient seawater bacteria not normally associated with the sponge ecosystem but nevertheless detected due to their close proximity, or to a change in environmental conditions making the sponge habitat less suitable for growth of many opportunistic bacterial taxa otherwise associated with sponges in the wild. Previous experiment on the bacterial diversity of *H. panicea* have shown that even in the wild its bacterial diversity is highly variable depending on site and time of the year [36–38], indicating that many of the bacterial taxa routinely found in the sponge are transient. Nevertheless, up to 32 % of the bacterial OTUs present within the original sponges was also present after cultivation, showing that some of the sponge-associated bacteria were retained at least over the course of several weeks.

During the co-cultivation experiment and using the sponge cultivation set-up described in method 5 we could re-approve the successful maintenance of the sponge and its dominant symbiont within the sponge body. However, an initial decline in the relative abundance of "*Ca*. H. symbioticus" was measured after 3 weeks before increasing back to a higher relative abundance similar that in its wild counterpart. This temporary decrease in symbiont abundance could indicate that the sponge explants were still adjusting to the new environment. This could be linked, as suggested above, to a reduced potential to feed while for instance rearranging the aquiferous system and regenerating the cut ectosome.

Using the novel co-cultivation method described in this study it was possible to enrich bacteria from a higher number of bacterial classes than using standard plate cultivation. Although isolation studies on *H. panicea* have previously been conducted using standard plating techniques [38,41,42], this is the first time that bacteria from the classes *Spirochaetia*, *Fusobacteriia*, *Delta-* and *Epsilonproteobacteria*, and *Clostridia* were enriched from *H. panicea*. The occurrence of reads associated with "*Ca*. H symbioticus" in the enrichment is intriguing, however it is not excludable that they represent remaining DNA from the inoculation due to the high abundance of the symbiont in the inoculum. However, they could also represent cells of "*Ca*. H. symbioticus" that survived or grew outside of their hosts during co-cultivation. In the latter case further research on the proposed co-cultivation method are highly warranted as it could provide valuable insights into the symbiont's physiology.

The fact that many of the enriched bacterial taxa and isolates were not detected within the original sponge used for inoculation, suggests that both methods also captured very low abundant taxa, possibly those present in seawater attached to the sponge or in its aquiferous system not detected in 16S rRNA gene amplicon sequencing of the sponge body. Though this could also be the effect of PCR bias or primer missmatch [50]. Despite this disparity, most taxa were similar to sequences previously isolated from marine habitats and marine invertebrate excluding contamination as a reason for this difference in detected and cultivated diversity.

Conclusion

Novel approaches are needed to isolate so far uncultivated microorganisms for biotechnological purposes. Here we present a new sponge-microbe co-cultivation device using a diffusion growth chamber that can be deployed in a laboratory setting. Multiple wells allow for high throughput cultivation and re-isolation of microorganisms, while its *ex situ*

deployment enables easy access to the device and close control over environmental parameters. We show that cultivation conditions have an influence on the sponge-associated microbial diversity and thus could impact co-cultivation success. Ensuring suitable conditions to maintain both the sponge and its associated microbes is therefore highly necessary. Whereas we were able to show a higher cultivation success using the co-cultivation method compared to standard isolation techniques, it must be noted that enrichment does not necessarily lead to isolation of microbes in pure culture and this would need to be investigated in future experiments.

Methods

Sponge collection

H. panicea samples for cultivation, co-cultivation and bacterial isolation were collected from the intertidal and subtidal area in Seltjarnarnes (64°09' N 22°00' W) and Eyrarbakki (63°51' N 21°09' W) in South-West Iceland between April 2014 and June 2016. Samples were either cut from their substrate with a sterile scalpel or collected along with the substrate they were fastened to. Sponge individuals and samples were kept in 20 l buckets filled with seawater and were transported to the laboratory or cultivation facility within 1 hour after collection. Sponges were identified as *H. panicea* according to spicule structure and molecular markers as described in Knobloch *et al.* (2019).

Sponge cultivation

For the sponge cultivation experiments *H. panicea* collected from the wild were cut into explants or kept on their original rock substrate. Explants were cut underwater with a sterile scalpel making sure to keep more than 50 % of the explant surface covered with an intact ectosome to improve recovery. In total, five methods for cultivation of *H. panicea* were tested to find a suitable method for *ex situ* sponge maintenance. The first method consisted of a 60 l tank equipped with water aeration and commercial biofiltration mats. Water was circulated through the biofilter and within the tank using external pumps and airlift pumps respectively. Approximately 1 cm^3 sized sponge explants were suspended into the water column from nylon stings with a metal washer attached to the bottom. 400 ml of mixed microalgae cultures at approximately 10^5 cells ml^{-1} consisting of roughly equal concentrations of live *Dunaliella* sp. and *Phaeodactylum tricornutum* were added to the tank twice per week. Debris was removed from the tank twice a week using a syphon and seawater was replaced as necessary. Water loss from evaporation was replaced with deionised water. The second method was identical to the first method apart from that seawater was continuously added to the tank at a rate of approximately 600 l per day. The third method consisted of a tank set-up with a custom built bacterial moving bed biofilter, aeration for water movement and oxygenation, and artificial lighting set to 12 h dark and 12 h light. Water was continuously supplied to replace the system volume of 360 l twice per day. Sponge individuals attached to stones collected in the intertidal area were placed onto a plastic mesh raised above the tank bottom. Additional feed was not supplied to this system. The fourth method consisted of the same specifications as method 3 apart from that a 1000 l tank was added to the system in which *Laminaria* sp. with an intact holdfast attached to a stone were placed. Approximately 1 cm^3 sponge explants were attached to their thallus with nylon string. Artificial lighting set to 12 h dark and 12 h light was placed above the tank

providing a light intensity of 15 klux at the water surface to maintain the seaweed. Water exchange rates in this method were adjusted to approximately four time the water volume per day. In the fifth method effluent water from a public aquarium was supplied to a 15 l tank. Approximately 0.125 cm^3 sized sponge explants were placed individually into wells of a 24-well flat bottom TC-treated cell culture plate (Falcon). Aeration provided water movement and seawater was continuously exchanged at a rate of approximately 750 times the tank volume per day.

In methods 1-4 the water was supplied from a coastal seawater intake of a commercial aquaculture farm and previously sand filtered to remove large particles from the water before entering the farm. The water used in method 5 was from a coastal seawater borehole and was also previously sand filtered before entering the public aquarium tanks. The water temperature in all system remained between 8 and 15 °C degrees throughout the experiments. Attachment of sponge explants was visually observed and by gently touching or tipping the substrate to see if explants were fastened. Sponges where necrosis was detected were removed from the tanks. Sponge explants which had fallen off the substrate or were displaced by other means were also removed from the tank. Pumping activity was measured by placing food grade dye above the osculum and observing the exhalent water current. Ammonia and nitrite concentrations in the tanks were monitored using commercial aquarium test kits from Seachem. The cultivation experiments lasted between 18 and 25 weeks (see Table 1). For microbial diversity analysis three sponge samples from separate individuals were collected before and after each experiment. All sponge samples were stored at -80 °C until proceeding with DNA extraction.

Sponge-bacteria co-cultivation experiment

For the sponge-bacteria co-cultivation experiment a wild sponge was cut into ca. 0.125 cm^3 explants which were distributed individually into the wells of a Bio-Dot SF microfiltration apparatus (BioHit). A bacterial inoculum was made by rinsing 10 g of the same sponge with sterile artificial seawater (450 mM NaCl, 10 mM KCl, 9 mM $CaCl_2$, 14 mM $MgCl_2$, 8 mM $MgSO_4$) to remove bacteria attached to the outer surface and grinding the sponge in sterile calcium-magnesium free seawater (450 mM NaCl, 10 mM KCl). The sponge suspension was centrifuged for 2 minutes at 500 x g to remove the majority of sponge cells and the supernatant was subjected to serial dilutions of 10^2, 10^3, 10^4 and 10^5 in sterile artificial seawater. The bottom container of the microfiltration apparatus was filled with Marine Agar (BD Difco) under sterile conditions by inserting the medium through the effluent tube of the apparatus until the medium reached the top layer of the well separation grid. Then the effluent tube was closed and the media was left to solidify. 3 µl of the 10^5 sponge-bacterial dilution was placed onto the solid medium of each well and covered by 0.2 µm nylon membrane filters (Whatman) followed by seven layers of 25 µm filter paper (Miracloth, Merk-Millipore) to prevent the membrane from clogging. The top half of the microfiltration apparatus containing the sponge explants was briefly removed from the water and tightly screwed onto the bottom half and placed back under the water surface. This co-cultivation device was placed in a 15 l tank and kept under conditions as described for method 5 above. Sponge explants displaced or showing signs of necrosis within the first week were replaced with explants kept in a reserve tank. Explants deceased or displaced after the first week were not replaced. Wells were cleaned regularly to prevent build-up of debris above the filters and membrane separating the explants and the bacterial inoculum.

After ten weeks the device was removed from the tank and transported back to the laboratory on ice. The outside of the device was rinsed thoroughly with sterile water and the top half was removed, being careful not to displace or contaminate the membrane. The bottom of the membrane and surface of the solid media were searched for colony growth which was combined to a single pellet in a microcentrifuge tube and stored at -80 °C until DNA extraction.

Bacterial isolation

For bacterial isolation using the standard plating method the same sponge bacterial suspension as for the co-cultivation experiment was used. Dilutions at 10^4 and 10^5 were plated on Marine Agar and Starch Yeast-Extract Peptone Sea Water Agar (10 g potato starch, 4 g Bacto Yeast Extract, 2 g Bacto Peptone, 15 g Bacto Agar in 1 l 0.2 µm filtered seawater). All media were additionally prepared with *H. panicea* sponge extract, made from a sponge suspension filtered through a 0.2 µm filter and added to the media at 0.5 % (v/v) after autoclaving.

Inoculated plates were incubated at 10 and 22 °C and colonies were picked and restreaked on Marine Agar plates at least three times to assure strains were isolates. Isolated strains were analysed on a Microflex MALDI-TOF mass spectrometer (Bruker) after formic acid extraction. In short, a loopfull of an isolated colony was diluted in 300 µl of ultra-pure water in a microcentrifuge tube. 900 µl of pure ethanol was added and the cell suspension vortexed for 1 min followed by centrifugation at 13.000 x g for 2 min. The supernatant was removed and the pellet was air dried, after which 30 µl of 70 % formic acid was added and mixed with the pellet by pipetting up and down. 30 µl of 100 % acetonitrile was added to the tube and mixed carefully followed by centrifugation at 13.000 x g for 2 min. 1 µl of the supernatant was placed on a stainless steel MALDI target and allowed to air dry. Once dry, the spot was overlaid with freshly prepared HCCA (Bruker) matrix solution and run on the MALDI-TOF using default settings. A main spectra profile dendrogram was created based on m/z spectra comparison and strains clustering into same groups at low distance equivalents were considered as the same species. For each species cluster, one isolate was selected for sequencing of the partial 16S rRNA gene for which a colony of the selected strains was subjected to DNA extraction using the Epicentre DNA Purification Kit (Epicentre) according to the manufacturer's instructions. The partial 16S rRNA gene was amplified using the primer pair F27/R806 (5'-AGAGTTTGATCMTGGCTCAG-3' / 5'-GGACTACVSGGGTATCTAAT-3) and sequenced on a 3730 DNA Analyser (Applied Biosystems, Hitachi). Sequences were trimmed to position 341 and 785 (relative to the *E. coli* 16S rRNA gene) using the bioinformatics software Geneious, to be able to compare against 16S rRNA gene amplicons, and taxonomically classified using the SINTAX algorithm [51] against the SILVA v123 LTP database [52].

Microbial diversity analysis

DNA was extracted from sponge samples, seawater samples and the bacterial pellet retrieved from co-cultivation using the Epicentre DNA Purification Kit according to the manufacturer's instructions, apart from that the sponge samples were first homogenised in laboratory grade Millipore water by passing them at least 10 times through a 90-gauge syringe. A region of the 16S rRNA gene covering the V3-V4 variable region was amplified using the universal prokaryotic primer pair S-D-Bact-0341-b-S-17 (5′-CCTACGGGNGGCWGCAG-3′) / S-D-Bact-0785-a-A-21 (5′-

GACTACHVGGGTATCTAATCC-3′) [53] and sequenced on a Illumina MiSeq desktop sequencer. PCR conditions, library preparation and sequencing were performed as described in Knobloch *et al.* (2019).

Raw sequences were demultiplex and 40 and 60 bps were trimmed off the forward and reverse sequences respectively using the fastx_trimmer command of the FASTX-Toolkit [54]. Trimmed reads were clustered into operational taxonomic units (OTUs) at 97 % sequence identify using the UPARSE pipeline [55]. OTUs were taxonomically classified using the SINTAX algorithm against the SILVA v123 LTP database. The OTU and taxonomy tables were imported and subsequently analysed in the package phyloseq [56] in the statistical software R [57] implemented in RStudio [58]. Plots were created using ggplot2 [59].

Data availability

Raw 16S rRNA gene amplicon reads are deposited in the Sequence Read Archive under BioProject ID PRJNA521872.

Acknowledgements

We thank Ásgeir Guðnason and Þorkell Heiðarsson for providing facilities and help with the cultivation experiments. This work was supported by the EC Marie Curie Actions ITN Project BluePharmTrain [No. 607786].

References

1. Belarbi, E. H., Contreras Gómez, A., Chisti, Y., García Camacho, F. & Molina Grima, E. Producing drugs from marine sponges. *Biotechnol. Adv.* **21,** 585–598 (2003).
2. Thakur, N. L. & Muller, W. E. G. Biotechnological potential of marine sponges. *Curr. Sci.* **86,** 1506–1512 (2004).
3. Laport, M., Santos, O. & Muricy, G. Marine Sponges: Potential Sources of New Antimicrobial Drugs. *Curr. Pharm. Biotechnol.* **10,** 86–105 (2009).
4. Osinga, R., Tramper, J. & Wijffels, R. H. Cultivation of marine sponges. *Mar. Biotechnol.* **1,** 509–532 (1999).
5. Wang, G. Diversity and biotechnological potential of the sponge-associated microbial consortia. *J. Ind. Microbiol. Biotechnol.* **33,** 545–551 (2006).
6. Hentschel, U., Usher, K. M. & Taylor, M. W. Marine sponges as microbial fermenters. *FEMS Microbiol. Ecol.* **55,** 167–177 (2006).
7. Steinert, G. *et al. BluePharmTrain : Biology and Biotechnology of Marine Sponges*. (2018).
8. Webster, N. S. & Thomas, T. The Sponge Hologenome. *MBio* **7,** e00135-16 (2016).
9. Thomas, T. *et al.* Diversity, structure and convergent evolution of the global sponge microbiome. *Nat. Commun.* **7,** 11870 (2016).
10. Amann, R. I., Ludwig, W. & Schleifer, K.-H. Phylogenetic identification and in situ detection of individual microbial cells without cultivation. *Microbiol. Rev.* **59,** 143–169 (1995).
11. Zengler, K. *et al.* Cultivating the uncultured. *Proc. Natl. Acad. Sci. U. S. A.* **99,** 15681–6 (2002).
12. Nichols, D. *et al.* Use of ichip for high-throughput in situ cultivation of "uncultivable" microbial species. *Appl. Environ. Microbiol.* **76,** 2445–50 (2010).

13. Kaeberlein, T., Lewis, K. & Epstein, S. S. Isolating "uncultivable" microorganisms in pure culture in a simulated natural environment. *Science* **296,** 1127–9 (2002).
14. Pham, V. H. T. & Kim, J. Cultivation of unculturable soil bacteria. *Trends Biotechnol.* **30,** 475–484 (2012).
15. Ling, L. L. *et al.* A new antibiotic kills pathogens without detectable resistance. *Nature* **517,** 455–459 (2015).
16. Steinert, G., Whitfield, S., Taylor, M. W., Thoms, C. & Schupp, P. J. Application of Diffusion Growth Chambers for the Cultivation of Marine Sponge-Associated Bacteria. *Mar. Biotechnol.* **16,** 594–603 (2014).
17. Osinga, R., Belarbi, E. H., Grima, E. M., Tramper, J. & Wijffels, R. H. Progress towards a controlled culture of the marine sponge< i> Pseudosuberites andrewsi</i> in a bioreactor same growth feeding on particles or on dissolved matter. *J. Biotechnol.* **100,** 141–146 (2003).
18. Webster, N. S. & Taylor, M. W. Marine sponges and their microbial symbionts: Love and other relationships. *Environ. Microbiol.* **14,** 335–346 (2012).
19. Kumala, L., Riisgård, H. U. & Canfield, D. E. Osculum dynamics and filtration activity studied in small single-osculum explants of the demosponge Halichondria panicea. *Mar Ecol Prog Ser* **572,** 117–128 (2017).
20. Reincke, T. & Barthel, D. Silica uptake kinetics of Halichondria panicea in Kiel Bight. *Mar. Biol.* **129,** 591–593 (1997).
21. Riisgård, H. U. & Larsen, P. S. Filter-feeding in marine macro-invertebrates: Pump characteristics, modelling and energy cost. *Biol. Rev.* **70,** 67–106 (1995).
22. Reiswig, H. M. Particle feeding in natural populations of three marine demosponges. *Biol. Bull.* **141,** 568–591 (1971).
23. Khalaman, V. V., Korchagina, N. M. & Komendantov, A. Y. The impact of waterborne cues from conspecifics and other species on the larvae of Halichondria panacea Pallas, 1766 (Porifera: Demospongiae). *Russ. J. Mar. Biol.* **40,** 36–42 (2014).
24. Gerasimova, E. I. & Ereskovsky, A. V. Reproduction of two species of Halichondria (Demospongiae: Halichondriidae) in the White Sea. *Porifera Res. biodiversity, Innov. Sustain. Livros. Mus. Nac. Rio Janeiro, Proc. 7th Int. Sponge Symp. Livros. Mus. Nac. Rio Janeiro, Proc. 7th Int. Sponge Symp.* (2007).
25. Barthel, D. On the ecophysiology of the sponge Halichondria panicea in Kiel Bight. I. Substrate specificity, growth and reproduction. *Marine Ecology Progress Series* **32,** 291–298 (1986).
26. Barthel, D. On the ecophysiology of the sponge Halichondria panicea in Kiel Bight. 2. Biomass, production, energy budget and integration in environmental processes. *Mar. Ecol. Prog. Ser. Oldend.* **43,** 87–93 (1988).
27. Witte, U., Barthel, D. & Tendal, O. The reproductive cycle of the sponge Halichondria panicea Pallas (1766) and its relationship to temperature and salinity. *J. Exp. Mar. Bio. Ecol.* **183,** 41–52 (1994).
28. Barthel, D. & Wolfrath, B. Tissue sloughing in the sponge Halichondria panicea: a fouling organism prevents being fouled. *Oecologia* **78,** 357–360 (1989).
29. Cimino, G., De Stefano, S. & Minale, L. Paniceins, unusual aromatic sesquiterpenoids linked to a quinol or quinone system from the marine sponge Halichondria panicea. *Tetrahedron* **29,** 2565–2570 (1973).
30. Perovic, S. *et al.* Neuroactive compounds produced by bacteria from the marine sponge Halichondria panicea: Activation of the neuronal NMDA receptor. *Environ. Toxicol. Pharmacol.* **6,** 125–133 (1998).

31. Wicke, C. *et al.* Production and Structure Elucidation of Glycoglycerolipids from a Marine Sponge-Associated Microbacterium Species. (2000). doi:10.1021/NP990313B
32. Lang, S. ., Beil, W. ., Tokuda, H. ., Wicke, C. . & Lurtz, V. . Improved Production of Bioactive Glucosylmannosyl-Glycerolipid by Sponge-Associated Microbacterium Species. *Mar. Biotechnol.* **6,** 152–6 (2004).
33. Toth, G. B. & Lindeborg, M. Water-soluble compounds from the breadcrumb sponge Halichondria panicea deter attachment of the barnacle Balanus improvisus. *Mar. Ecol. Prog. Ser.* **354,** 125–132 (2008).
34. Purushottama, G., Venkateshvaran, K., Pani Prasad, K. & Nalini, P. Bioactivities of extracts from the marine sponge Halichondria panicea. *J. Venom. Anim. Toxins Incl. Trop. Dis.* **15,** 444–459 (2009).
35. Ferreira, M. *et al.* Cytotoxic activity of extracts of marine sponges from NW Spain on a neuroblastoma cell line. *Environ. Toxicol. Pharmacol.* **32,** 430–437 (2011).
36. Althoff, K., Schütt, C., Steffen, R., Batel, R. & Müller, W. E. G. Evidence for a symbiosis between bacteria of the genus *Rhodobacter* and the marine sponge *Halichondria panicea*: Harbor also for putatively toxic bacteria? *Mar. Biol.* **130,** 529–536 (1998).
37. Knobloch, S., Jóhannsson, R. & Marteinsson, V. Bacterial diversity in the marine sponge Halichondria panicea from Icelandic waters and host-specificity of its dominant symbiont "Candidatus Halichondribacter symbioticus". *FEMS Microbiol. Ecol.* **95,** (2019).
38. Wichels, A., Würtz, S., Döpke, H., Schütt, C. & Gerdts, G. Bacterial diversity in the breadcrumb sponge Halichondria panicea (Pallas). *FEMS Microbiol. Ecol.* **56,** 102–118 (2006).
39. Naim, M. A. *et al.* Host-specific microbial communities in three sympatric North Sea sponges. *FEMS Microbiol. Ecol.* **90,** 390–403 (2014).
40. Steinert, G., Rohde, S., Janussen, D., Blaurock, C. & Schupp, P. J. Host-specific assembly of sponge-associated prokaryotes at high taxonomic ranks. *Sci. Rep.* **7,** 2542 (2017).
41. Imhoff, J. F. & Stöhr, R. in *Sponges (Porifera)* 35–57 (Springer, Berlin, Heidelberg, 2003). doi:10.1007/978-3-642-55519-0_2
42. Schneemann, I. *et al.* Comprehensive investigation of marine actinobacteria associated with the sponge halichondria panicea. *Appl. Environ. Microbiol.* **76,** 3702–3714 (2010).
43. Schneemann, I. *et al.* Mayamycin, a cytotoxic polyketide from a streptomyces strain isolated from the marine sponge halichondria panicea. *J. Nat. Prod.* **73,** 1309–1312 (2010).
44. Thomassen, S. & Riisgård, H. U. Growth and energetics of the sponge Halichondria panicea. *Mar. Ecol. Prog. Ser.* **128,** 239–246 (1995).
45. Riisgård, H. U., Kumala, L. & Charitonidou, K. Using the *F/R* -ratio for an evaluation of the ability of the demosponge *Halichondria panicea* to nourish solely on phytoplankton versus free-living bacteria in the sea. *Mar. Biol. Res.* **12,** 907–916 (2016).
46. Hausmann, R., Vitello, M. P., Leitermann, F. & Syldatk, C. Advances in the production of sponge biomass Aplysina aerophoba-A model sponge for ex situ sponge biomass production. *J. Biotechnol.* **124,** 117–127 (2006).
47. Mohamed, N. M., Enticknap, J. J., Lohr, J. E., McIntosh, S. M. & Hill, R. T. Changes in bacterial communities of the marine sponge Mycale laxissima on transfer into

aquaculture. *Appl. Environ. Microbiol.* **74,** 1209–1222 (2008).
48. Isaacs, L. L. T. *et al.* Comparison of the bacterial communities of wild and captive sponge Clathria prolifera from the Chesapeake Bay. *Mar. Biotechnol.* **11,** 758–770 (2009).
49. Webster, N. S. *et al.* Bacterial Community Dynamics in the Marine Sponge Rhopaloeides odorabile Under In Situ and Ex Situ Cultivation. *Mar. Biotechnol.* **13,** 296–304 (2011).
50. Eloe-Fadrosh, E. A., Ivanova, N. N., Woyke, T. & Kyrpides, N. C. Metagenomics uncovers gaps in amplicon-based detection of microbial diversity. *Nat. Microbiol.* **1,** 15032 (2016).
51. Edgar, R. C. SINTAX: a simple non-Bayesian taxonomy classifier for 16S and ITS sequences. *bioRxiv* 074161 (2016). doi:10.1101/074161
52. Yilmaz, P. *et al.* The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res.* **42,** D643–D648 (2014).
53. Klindworth, A. *et al.* Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* **41,** e1–e1 (2013).
54. Gordon, A. & Hannon, G. J. Fastx-toolkit. *FASTQ/A short-reads preprocessing tools (unpublished) http://hannonlab. cshl. edu/fastx_toolkit* **5,** (2010).
55. Edgar, R. C. UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat Meth* **10,** 996–998 (2013).
56. McMurdie, P. J. & Holmes, S. phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. *PLoS One* **8,** e61217 (2013).
57. R Core Team. in *R foundation for Statistical Computing* (2017).
58. RStudio Team. Rstudio: Integrated development environment for r [Computer software manual]. Boston, MA. (2016).
59. Wickham, H. ggplot2: Elegant Graphics for Data Analysis Springer-Verlag. *New York* (2009).

Supplementary Table S1: List of bacterial isolates from *H. panicea* and enriched OTUs from the co-cultivation experiment.

	Assigned class	Best hit (GenBank)	Name (Genbank)	Isolation source	Sequence identity (%)	Rel. abundance *H. panicea* (%)	Rel. abundance in enrichment (%)
***H. panicea* isolates**							
Hp_23	Actinobacteria	MH701860.1	Micrococcus sp. strain CDR-SL-16	Marine habitat	100	0	NA
Hp_26	Actinobacteria	MH681545.1	Rhodococcus sp. strain SYP-A7263	Unknown	100	0	NA
Hp_7	Cytophagia	AB792995.1	Flammeovirga sp. KR07-01	Marine reef	99.1	0	NA
Hp_1	Flavobacteriia	JF827412.1	Uncultured bacterium clone 44F2	Marine invertebrate	100	0.787	NA
Hp_19	Flavobacteriia	AF493686.1	Flavobacteriaceae str. SW334	Coastal seawater	99.1	0.014	NA
Hp_32	Flavobacteriia	KM017080.1	Tenacibaculum sp. HMF2317	Seawater	99.1	0.014	NA
Hp_34	Flavobacteriia	EF491285.2	Uncultured Bacteroidetes bacterium clone S1-24	Marine habitat	99.1	0.014	NA
Hp_35	Flavobacteriia	MH594572.1	Uncultured bacterium clone RU12(10)	Marine algae	100	0	NA

Hp_40	Flavobacteriia	AF493686.1	Flavobacteriaceae str. SW334	Coastal seawater	99.4	0.014	NA
Hp_47	Flavobacteriia	KX245373.2	Flagellimonas sp. strain ECD12	Marine algae	100	0	NA
Hp_51	Flavobacteriia	GQ274057.1	Uncultured Bacteroidetes bacterium clone 24D3	Seawater	97.8	0	NA
Hp_55	Flavobacteriia	JX530989.1	Uncultured Flavobacteriaceae bacterium clone C146300385	Seawater	98.4	0	NA
Hp_31	Bacilli	MG020099.1	Bacillus algicola	Marine invertebrate	100	0	NA
Hp_10	Alphaproteobacteria	MF594171.1	Vibrio sp. strain 2216E-X-15	Steel structure	100	0	NA
Hp_11	Alphaproteobacteria	AY172307.1	Uncultured bacterium clone 38	Unknown	99.7	0.009	NA
Hp_20	Alphaproteobacteria	KP684473.1	Bacterium enrichment culture clone RB2LO08	Halichondria panicea	100	0	NA
Hp_46	Alphaproteobacteria	KX177742.1	Uncultured bacterium clone PW-15G12	Seawater	100	0	NA
Hp_48	Alphaproteobacteria	FR693315.1	Sphingopyxis sp. BB46	Marine invertebrate	100	0.014	NA
Hp_52	Alphaproteobacteria	MF461369.1	Sphingorhabdus sp. strain EL163	Marine invertebrate	99.0	0.014	NA
Hp_54	Alphaproteobacteria	KJ601444.1	Uncultured bacterium clone C34	Coral tissue	100	0.009	NA

Hp_2	Gammaproteobacteria	MG694250.1	Shewanellaceae bacterium strain PS03	Beach sand	100	0	NA
Hp_3	Gammaproteobacteria	MG283319.1	Shewanella halifaxensis strain LPB0184	Sea sand	100	0	NA
Hp_4	Gammaproteobacteria	KP684297.1	Bacterium DA91	Halichondria panicea	100	0	NA
Hp_5	Gammaproteobacteria	JX206721.1	Uncultured bacterium clone TV10-912_C10	Marine sponge	99.1	0	NA
Hp_8	Gammaproteobacteria	MF594182.1	Vibrio coralliilyticus strain 2216E-S-96	Marine coral	100	0	NA
Hp_12	Gammaproteobacteria	JN621578.1	Uncultured bacterium clone G13T5.7_B7	Marine sediment	91.4	0.009	NA
Hp_13	Gammaproteobacteria	MG254532.1	Pseudomonas fluorescens strain 8	River	100	0	NA
Hp_14	Gammaproteobacteria	MH057255.1	Microbulbifer sp. strain LSS-12	Marine habitat	100	0	NA
Hp_16	Gammaproteobacteria	MH512866.1	Vibrio sp. strain 101	Sea water	100	0	NA
Hp_17	Gammaproteobacteria	MG972754.1	Vibrio splendidus strain Vaa2	Marine invertebrate	100	0	NA
Hp_18	Gammaproteobacteria	JX206721.1	Uncultured bacterium clone TV10-912_C10	Marine sponge	99.0	0	NA
Hp_21	Gammaproteobacteria	KJ814573.1	Marine bacterium IIIA016	Marine habitat	100	0	NA

Hp_22	Gammaproteobacteria	KX453262.1	Pseudoalteromonas citrea strain hHt18-5	Marine invertebrate	100	0	NA
Hp_24	Gammaproteobacteria	MG815851.1	Halomonas subglaciescola strain KU-43	Cheese brine	100	0	NA
Hp_27	Gammaproteobacteria	KT582798.1	Uncultured Enterovibrio sp. clone 285KII	Marine fish	100	0	NA
Hp_28	Gammaproteobacteria	MF594128.1	Pseudoalteromonas sp. strain 2216E-X-10	Unknown	100	0	NA
Hp_29	Gammaproteobacteria	JX436437.1	Shewanella sp. Wash4	Marine sponge	100	0	NA
Hp_30	Gammaproteobacteria	MG833271.1	Shewanella sp. strain 1719I3156	Marine sponge	96.6	0.213	NA
Hp_36	Gammaproteobacteria	KY655372.1	Spongiobacter sp. strain EA271	Marine sponge	95.3	0	NA
Hp_37	Gammaproteobacteria	MG996331.1	Uncultured bacterium clone Plate_K_H05_F	Seawater	100	0	NA
Hp_41	Gammaproteobacteria	MG876115.1	Uncultured bacterium clone 4F_1112_20118_4586	Seawater	100	0	NA
Hp_42	Gammaproteobacteria	MF042743.1	Uncultured Idiomarina sp. clone 381	Groundwater	100	0	NA
Hp_49	Gammaproteobacteria	JX495001.1	Uncultured marine bacterium clone DLGF22	Marine algae	100	0	NA
Hp_53	Gammaproteobacteria	AF466927.1	Uncultured gamma proteobacterium isolate DGGE band D'7	Seawater	100	0	NA

Co-cultivation enrichment OTUs

Otu_8	Bacteroidetes_incertae_sedis	KX172965.1	Uncultured bacterium clone EzlYyy60	Marine sediment	99.3	0	0.007
Otu_15	Bacteroidetes_incertae_sedis	EU050907.1	Uncultured bacterium clone SS1_B_06_18	Marine sediment	100	0.005	9.043
Otu_231	Bacteroidetes_incertae_sedis	FJ716908.1	Uncultured bacterium clone B1_10.1_1	Marine invertebrate	97.5	0	0.082
Otu_668	Bacteroidetes_incertae_sedis	KX172965.1	Uncultured bacterium clone EzlYyy60	Marine sediment	97.8	0	0.004
Otu_671	Bacteroidetes_incertae_sedis	KM203417.1	Uncultured bacterium clone Plate5_4664_0_2	Marine habitat	90.8	0	0.019
Otu_25	Bacteroidia	FJ716944.1	Uncultured bacterium clone A4_10.4_1	Marine sediment	96.2	0	1.547
Otu_275	Bacteroidia	KU533822.1	Bacteroides sp.4SWWS3-28	Marine sediment	99.3	0	0.041
Otu_376	Deltaproteobacteria	KX088584.1	Uncultured bacterium clone MLGsedbac-8	Marine habitat	88.6	0	0.004
Otu_835	Deltaproteobacteria	KC606247.1	Uncultured bacterium clone CarbonSeq013_022210_A07	Groundwater	88.1	0	0.004
Otu_2	Clostridia	KF799151.1	Uncultured bacterium clone Woods-Hole_a5657	Marine invertebrate	100	0.005	32.099
Otu_32	Clostridia	JX391218.1	Uncultured bacterium clone H3078	Marine sediment	100	0	0.245

Otu_49	Clostridia	AJ441231.1	Uncultured firmicute	Marine invertebrate	99.8	0	0.616
Otu_76	Clostridia	LN849485.1	Uncultured bacterium	Biogas reactor	91.1	0	0.627
Otu_82	Clostridia	GU472396.1	Uncultured bacterium clone BBD-Feb09-6BB-70	Coral tissue	96.7	0	0.063
Otu_86	Clostridia	KF799151.1	Uncultured bacterium clone Woods-Hole_a5657	Marine invertebrate	95.8	0	0.434
Otu_104	Clostridia	KR086551.1	Uncultured bacterium clone DH162B34	Marine habitat	100	0	0.319
Otu_116	Clostridia	FJ223458.1	Uncultured bacterium clone 447	Marine sediment	97.5	0	0.156
Otu_162	Clostridia	AB704728.1	Uncultured bacterium M03-2-Bac-F6_27F_1_F06_038	Borehole water	91.1	0	0.200
Otu_193	Clostridia	KF323282.1	Uncultured bacterium clone GXTJ5A301BOFJ0	Marine invertebrate	93.7	0	0.293
Otu_384	Clostridia	MG367102.1	Uncultured Clostridiales Family IV bacterium clone TR_Beef_D2	Marine sediment	99.1	0	0.004
Otu_395	Clostridia	KF179748.1	Uncultured bacterium clone MAY9C13	Coral tissue	96.9	0	0.007
Otu_676	Clostridia	KF179748.1	Uncultured bacterium clone MAY9C13	Coral tissue	92.5	0	0.059
Otu_29	Fusobacteriia	KX956270.1	Uncultured bacterium clone OTU1955	Marine sediment	99.8	0.009	1.688

Otu_1	Alphaproteobacteria	KJ453525.1	Uncultured bacterium clone HP1-2.1	Halichondria panicea	100	73.805	0.030
Otu_4	Alphaproteobacteria	KX550187.1	Uncultured bacterium clone JCC_RecOTU_44	Marine sediment	99.8	0.068	0.004
Otu_45	Alphaproteobacteria	JN412112.1	Uncultured Bacteroidetes bacterium isolate DGGE gel band Chitin292_28	Marine invertebrate	97.0	0	0.638
Otu_122	Alphaproteobacteria	FJ175061.1	Uncultured bacterium clone sbrh_97	Soil	100	0.005	0.004
Otu_146	Alphaproteobacteria	EU458567.1	Uncultured bacterium clone HY2_g05_2	Land animal	92.9	0	0.100
Otu_75	Deltaproteobacteria	FJ202982.1	Uncultured bacterium clone SGUS569	Coral tissue	97.2	0	0.286
Otu_87	Deltaproteobacteria	KC606247.1	Uncultured bacterium clone CarbonSeq013_022210_A07	Groundwater	92.2	0	0.004
Otu_283	Deltaproteobacteria	HQ400926.1	Desulfovibrionaceae bacterium enrichment culture clone MS_OIL_O12	Marine sediment	98.7	0	0.119
Otu_391	Deltaproteobacteria	DQ088266.1	Uncultured bacterium clone cs49	Sediment	91.8	0	0.004
Otu_643	Deltaproteobacteria	HM598559.1	Uncultured bacterium clone Zeebrugge_B07	Brakish sediment	92.3	0	0.078
Otu_651	Deltaproteobacteria	GU061221.1	Uncultured delta proteobacterium clone 5m-23	Seawater	90.1	0	0.100

Otu_941	Deltaproteobacteria	KC607039.1	Uncultured bacterium clone CarbonSeq007_022210_E07	Groundwater	92.0	0	0.122
Otu_984	Deltaproteobacteria	EU236307.1	Uncultured bacterium clone Hg1a2D2	Marine sponge	91.2	0	0.278
Otu_989	Deltaproteobacteria	KC606247.1	Uncultured bacterium clone CarbonSeq013_022210_A07	Groundwater	92.2	0	0.019
Otu_3	Epsilonproteobacteria	JQ862033.1	Uncultured epsilon proteobacterium clone GC234-4-78	Marine sponge	100	0.005	27.698
Otu_6	Epsilonproteobacteria	KY236001.1	Uncultured bacterium clone HC6-92	Marine habitat	100	0.009	5.433
Otu_12	Epsilonproteobacteria	KP183006.1	Uncultured Arcobacter sp. clone 12S_65	Seawater	99.8	0.023	0.282
Otu_19	Epsilonproteobacteria	JF928745.1	Uncultured Arcobacter sp. clone OTU_A1_SP2_101	Marine habitat	100	0.005	2.364
Otu_20	Epsilonproteobacteria	FJ628231.1	Uncultured bacterium clone Nit2Au0637_410	Seawater	99.8	0.005	0.004
Otu_27	Epsilonproteobacteria	JX170296.1	Uncultured bacterium clone AJ-U-CD-188	Marine invertebrate	100	0.014	3.047
Otu_30	Epsilonproteobacteria	HQ203919.1	Uncultured bacterium clone SW-Apr-23	Seawater	100	0.005	0.152
Otu_31	Epsilonproteobacteria	LC133156.1	Uncultured Arcobacter sp.	Marine invertebrate	99.8	0.827	0.200
Otu_39	Epsilonproteobacteria	FJ202693.1	Uncultured bacterium clone SGUS1281	Coral tissue	97.4	0	0.946

Otu_72	Epsilonproteobacteria	KY969129.1	Uncultured Arcobacter sp. clone D4	Marine habitat	99.3	0.149	0.148
Otu_413	Epsilonproteobacteria	EF029030.1	Uncultured bacterium clone root_Ell_39	Seagrass	100	0	0.404
Otu_591	Epsilonproteobacteria	KP994506.1	Uncultured Arcobacter sp. clone Pichicolo-G05	Seawater	91.6	0	0.004
Otu_827	Epsilonproteobacteria	AY768985.1	Uncultured bacterium clone FE2TopBac38	Marine habitat	97.6	0	0.004
Otu_875	Epsilonproteobacteria	FJ628346.1	Uncultured bacterium clone Nit5Au0628_769	Seawater	90.6	0	0.007
Otu_17	Gammaproteobacteria	KT341319.1	Uncultured bacterium clone a7-47	Oil field	94.1	0	6.086
Otu_41	Gammaproteobacteria	DQ357824.1	Moritella sp. SW43 clone 148.1	Seawater	98.9	0	1.180
Otu_132	Gammaproteobacteria	FJ618889.1	Uncultured bacterium clone BP5	Marine invertebrate	97.2	0	0.237
Otu_145	Gammaproteobacteria	KF325298.1	Uncultured bacterium clone G250WV301AHQ3C	Marine invertebrate	95.1	0	0.393
Otu_157	Gammaproteobacteria	KT341319.1	Uncultured bacterium clone a7-47	Oil field	93.2	0	0.397
Otu_173	Gammaproteobacteria	EU617814.1	Uncultured bacterium clone B13S-47	Marine sediment	100	0	0.219
Otu_208	Gammaproteobacteria	KP952941.1	Uncultured bacterium clone I3Q1XXJ01AMV7I	Marine invertebrate	96.9	0	0.442

Otu_991	Gammaproteobacteria	KF179764.1	Uncultured bacterium clone MAY9C35	Coral tissue	98.3	0.005	0.141
Otu_94	Spirochaetia	KC261847.1	Spirochaeta sp. K2	Marine sediment	97	0	0.590
Otu_135	Spirochaetia	KC261848.1	Spirochaeta sp. K1	Marine sediment	89.4	0	0.271

Paper IV

Pelagibaculum spongiae gen. nov., sp. nov., isolated from a marine sponge in South-West Iceland

Stephen Knobloch, Aurélien Daussin, Ragnar Jóhannsson, Viggó Þór Marteinsson

Accepted for publication in *International Journal of Systematic and Evolutionary Microbiology*

Author contributions:

Conceived and designed the study: SK, RJ, VM
Performed sampling: SK, AD
Performed laboratory work: SK, AD
Analysed data: SK, AD
Wrote the manuscript: SK
Reviewed and edited the manuscript: SK, AD, RJ, VM
Correspondence (publication process): SK

Abstract

A Gram-stain-negative, motile, mesophilic, aerobic, rod-shaped bacterium, designated Hp12^T, was isolated from a marine sponge in the intertidal zone off the coast of Seltjarnarnes (64° 16' N 22° 00' W), Iceland. Strain Hp12^T grew optimally at 20 - 22 °C, at pH 7 - 8 and in the presence of 1 - 2 % (w/v) NaCl. Phylogenetic analysis based on 16S rRNA gene sequences placed strain Hp12^T in the class *Gammaproteobacteria*, related to members of the genus *Alcanivorax* in the order *Oceanospirillales* with 90.3 – 88.5 % sequence similarity. The strain had a draft genome size of 4.99 Mbp with a DNA G + C content of 43.0 mol%. Cellular fatty acids were dominated by the $C_{16:1}\omega7c$, $C_{18:1}\omega7c$, and $C_{16:0}$. The predominant polar lipids were Phosphatidylglycerol and Phosphatidylethanolamine. The major respiratory lipoquinones were ubiquinone Q8 and menaquinone MK8. From the taxonomic information and phenotypic properties obtained in this study, it is proposed that strain Hp12^T be placed into a novel genus and species named *Pelagibaculum spongiae* gen. nov., sp. nov. The type strain of *Pelagibaculum spongiae* is Hp12^T (= DSM 104963^T = CECT 9367^T).

Results

Marine sponges often host diverse microbial communities, many of which produce bioactive compounds of biotechnological interest [1]. A screening for novel bacteria from marine sponges in Iceland was conducted in September 2016. One isolate designated Hp12^T, which showed 16S rRNA gene sequence similarity of less than 91.0 % to other members of the class *Gammaproteobacteria* is described in this study. Using high-throughput amplicon sequencing of the partial 16S rRNA gene, it was shown that Hp12^T was present at less than 0.01 % of the relative abundance in a sample of the sponge *Halichondria panicea* used for inoculation, but was not detectable in other samples of *H. panicea* or in seawater samples collected close to the sampling site across different seasons. A search against the Sequence Read Archive [2] (accessed Jan. 2018) using IMNGS [3] matched the 16S rRNA gene sequence of Hp12^T to sequences in 56 marine metagenomic datasets with 99% or higher sequence similarity covering at least 200 bp and with an average relative abundance of 0.005% in the bacterial communities. This indicates that Hp12^T is likely a low-abundant, free-living seawater bacterium and not a member of the sponge-specific microbial community. Phylogenetic analysis based on the full 16S rRNA gene showed that Hp12^T was distinct from other characterised members of the class *Gammaproteobacteria*. The aim of the present work was to determine the taxonomic position of strain Hp12^T using a polyphasic approach.

A marine sponge identified as *Halichondria panicea* was collected from the intertidal area in Seltjarnarnes, Iceland (64° 16' N 22° 00' W), rinsed several times in sterile seawater to remove transient bacteria attached to the sponge surface and subsequently used for bacterial isolation. Strain Hp12^T was obtained by standard dilution plating technique on Marine Agar 2216 (MA, BD Difco) at 20 °C. Hp12^T was routinely cultivated at 20 °C on MA or in MB. For comparative studies, reference strain *Alcanivorax borkumensis* DSM 11573^T was obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ) and grown on MA or in MB.

Cell morphology was observed by light microscopy (BX51, Olympus) and field emission scanning electron microscopy (Supra 25, Zeiss) of gold-coasted, freeze-dried cells on 0.2 µm GTTP membranes (Millipore). The Gram reaction was performed using the BD BBL Gram Stain Kit (BD) according to the manufacturer's instructions. Growth at 4, 7, 10, 17, 20, 25, 28, 30, 35, 37, 40, 45 and 50 °C was observed on MA to determine the temperature range for growth. Optimal growth temperature, pH range, NaCl tolerance and antibiotic susceptibility were performed on a Bioscreen C (Oy Growth Curves). The optimal growth temperature was determined in Marine Broth 2216 (MB, BD Difco) in 1 °C increments. The pH range for growth was determined in MB adjusted to pH 6.0 – 10.0 (at intervals of 0.5 pH units). Growth at different NaCl concentrations, from 0 to 22 % (w/v) in 1 % increments, was measured by adjusting the appropriate amount of NaCl in MB prepared according to the manufacturer´s formula except that NaCl was excluded. Growth under anaerobic conditions was determined after incubation for 22 days in a GasPak (BD) on MA and MA supplemented with nitrate (2 mM), nitrite (2 mM), fumarate (20 mM), sulphite (5 mM), thiosulfate (10 mM) and sulfate (10 mM). Catalase and oxidase activity were measured using catalase reagent droppers (BD Difco) and DrySlide reagent slides (BD Difco) respectively. Hydrolysis of casein, Tween 20 and starch were investigated on MA at 20°C using the substrate concentrations according to Cowan and Steel (2004). Susceptibility to antibiotics was investigated in MB containing the following (concentrations 10, 30 and 100 µg ml^{-1}): ampicillin, apramycin, chloramphenicol, enteromycin, kanamycin, nalidixic acid, penicillin,

polymyxin B, spectinomycin, streptomycin, tetracycline, vancomycin and zeocin. Additionally, API ZYM and API 20NE systems (bioMérieux) were used to test for enzyme activities and other physiological and biochemical traits according to the manufacturer's instructions, apart from that cells were suspended in artificial seawater (450 mM NaCl, 10 mM KCl, 9 mM $CaCl_2$, 14 mM $MgCl_2$, 8 mM $MgSO_4$), or inoculated in artificial seawater supplemented media. Both tests were run in triplicate.

Cellular biomass for chemotaxonomic analysis was obtained from cultures of Hp12^T and the reference strain grown in MB at 20 °C until the late exponential phase. Chromosomal DNA was extracted using the MasterPure DNA purification Kit (Epicentre) according to the manufacturer's instructions for total DNA extraction. Genomic libraries were constructed using the Nextera XT DNA Library Prep Kit (Illumina) and sequenced on an Illumina MiSeq platform using paired-end 300bp chemistry. Raw reads were quality trimmed and assembled using SPAdes version 3.9.1 [5] in paired-end mode with default settings resulting in a draft genome with the length of 4.99 Mbp and an average genome coverage of 197X. The full length 16S rRNA gene sequence was predicted using RNAmmer [6]. The DNA G + C content was determined from the draft genome assembly using the sequence analysis software Geneious (Biomatters). Average Nucleotide Identity (ANI) analysis was computed on EZBioCloud [7] and digital DNA-DNA hybridisation (dDDH) was calculated by the Genome-to-Genome Distance Calculator 2.1 on http://ggdc.dsmz.de/ [8]. Analysis of cellular fatty acids and polar lipids for all strains, and respiratory lipoquinones for Hp12^T were carried out from freeze dried cells by the Identification Service of the DSMZ, Braunschweig, Germany.

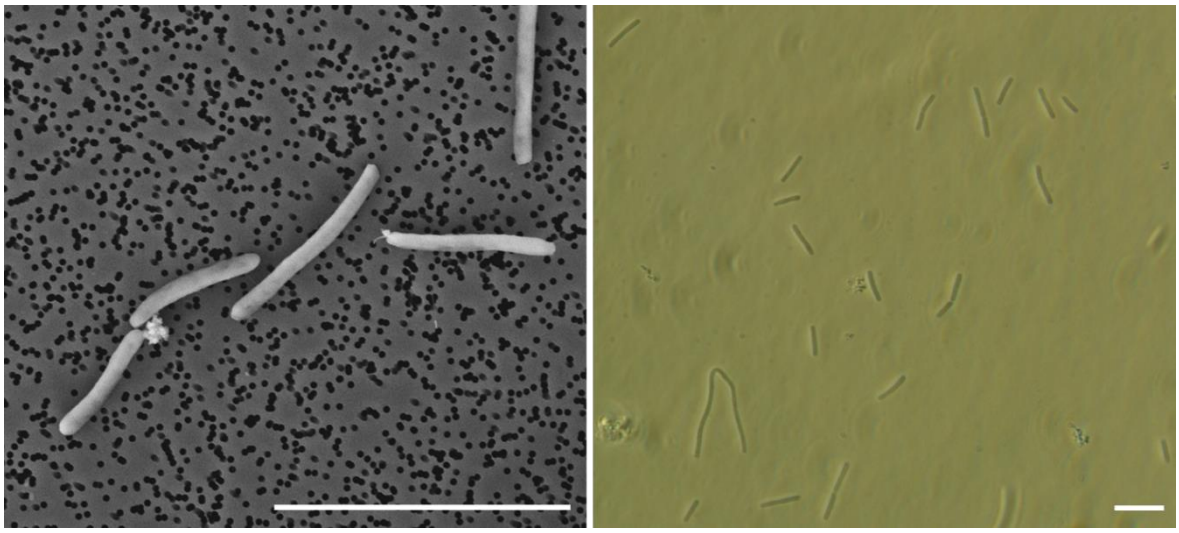


Figure 1*: Scanning electron microscopy (left) and light microscopy image (right) of Hp12^T during exponential growth phase. Bar = 10 µm*

Cells of strain Hp12^T were Gram-Stain-negative, non-spore-forming, motile, rod-shaped, 1.0 – 1.2 µm wide, 4 - 35 µm long and with an average length of 8 µm (Figure 1). Strain Hp12^T was aerobic, oxidase-positive and catalase-negative. Colonies grown on MA plates for 3 days at 20 °C were 1-2 mm in diameter, circular, raised and tan colour. Strain Hp12^T was able to grow at 4 – 28 °C but not at 30 °C. Growth occurred at NaCl concentration of 1 – 3 % and pH 6.5 – 9.5 but not above or below these limits. Optimal growth conditions for strain Hp12^T were at 20 – 22 °C, a NaCl concentration of 1 – 2 % and pH 7 - 8. The test for hydrolysis of casein, Tween 20 and starch were negative. Positive enzyme activity detected

on the API ZYM system were alkaline phosphatase, esterase, esterase lipase, leucine arylamidase, acid phosphatase and naphthol-AS-BI-phosphohydrolase. Positive reactions on the API 20NE test were the reduction of nitrate to nitrite, hydrolysis of esculin and hydrolysis of gelatin. Strain Hp12^T was resistant to apramycin (10 µg ml^{-1}), spectinomycin (10 and 30 µg ml^{-1}) and tetracycline (10 and 30 µg ml^{-1}), but susceptible to all other tested antibiotics and concentration. The DNA G + C content of strain Hp12^T was 43.0 mol% according to the draft genome sequence. The dominant respiratory lipoquinones in strain Hp12^T were both ubiquinone Q8 (98 %) and menaquinone MK8 (100 %). Other lipoquinones detected in lower amounts were Q7 (2 %). The cellular fatty acid profiles of strain Hp12^T and reference strain DSM 11573^T are shown in Table 1. The major fatty acids detected in strain Hp12^T (> 10.0 %) were $C_{16:1}\omega7c$ (55.3 %), $C_{18:1}\omega7c$ (12.4 %) and $C_{16:0}$ (10.8 %). The polar lipid profile of strain Hp12^T contained phosphatidylglycerol and phosphatidylethanolamine as the major polar lipids and diphosphatidylglycerol, two uncharacterized aminolipids, and three uncharacterized lipids as the minor components (Figure 3). Physiological, morphological and biochemical characteristics of strain Hp12^T, DSM 11573^T and other type strains of related taxa are presented in Table 3.

Table 1*: Cellular fatty acid profile of strain Hp12^T and reference strain DSM11573^T. Values are presented as total fatty acid content (%). Fatty acids present below 0.5 % are indicated as trace amounts (tr.).*

Fatty acid	***Pelagibaculum spongiae* Hp12^T**	***Alcanivorax borkumensis* DSM 11573^T**
$C_{10:0}$	tr.	4.0
$C_{10:0}$ 3-OH	1.9	2.9
$C_{12:0}$	tr.	4.7
$C_{12:0}$ 2-OH	-	0.8
$C_{12:0}$ 3-OH	3.6	9.0
$C_{12:1}$ 3-OH	2.2	tr.
$C_{14:0}$	1.4	1.5
$C_{15:0}$	-	tr.
$C_{16:0}$	10.8	24.1
$C_{18:0}$	3.2	0.5
$C_{16:1}$ ω7c	55.3	22.2
$C_{16:1}$ ω9c	1.2	-
$C_{17:1}$ ω8c	0.8	tr.
$C_{17:0}$	tr.	tr.
$C_{18:1}$ ω9c	3.1	1.1
$C_{18:1}$ ω7c	12.4	28.1
$C_{20:1}$ ω7c	0.9	-

The full length 16S rRNA gene sequence (1530 nt) of strain Hp12^T was compared to available sequences in the GenBank database using the BLAST program [9, 10] and also through identity analysis performed on EzBioCloud [11], showing that strain Hp12^T fell within the class of *Gammaproteobacteria*. The closest neighbours in the EzTaxon database with full length sequences were *Marinobacter litoralis* SW-45^T (1322/1426, 90.42 %), *Marinobacter zhejiangensis* CGMCC1.7061^T (1320/1426, 90.29 %) and *Alcanivorax*

borkumensis (1318/1461, 90.21 %). The 16S rRNA sequences of Hp12^T was aligned in SINA version 1.2.11 [12] against the global SSU SILVA alignment and merged with the LTPs132 database [13] using the ARB software [14]. Tree reconstruction was performed with 565 related sequences using the neighbour joining (ARB), maximum parsimony (DNAPars, [15]) and maximum-likelihood (PhyML, [16]) methods. Tree topology was further tested with 30%, 40% and 50% positional conservatory filters. The final tree was calculated with 565 sequences based on 1234 valid columns (50% positional conservatory filter) with PhyML (HKY85 evolutionary model) and 100 bootstraps. Despite highest 16S rRNA gene sequence similarity to *Marinobacter litoralis* and *Marinobacter zhejiangensis*, strain Hp12^T did not occupy a phylogenetically close position to the genus *Marinobacter* in the order *Alteromonadales* but instead clustered with members of the order *Oceanospirillales* (families *Oceanospirillaceae*, *Halomonadaceae*, *Alcanivoracaceae* and *Kangiellaceae*) (Figure 2). Based on all phylogenetic trees, strain Hp12^T occupied the basal position in a monophyletic group with only members of the genus *Alcanivorax* (Figure 2). Nevertheless, low 16S rRNA gene sequence similarity (<91 %) and a clear differentiation of ANI values to members of the *Alcanivorax* (Table 2) suggests that strain Hp12^T represents a member of a new genus in the family *Alcanivoracaceae*.

***Table 2**: 16S rRNA sequence similarity (%) and ANI value (%) between strain Hp12^T and members of the genus* Alcanivorax. *Genomes (RefSeq no.) used for ANI calculation: 1, NZ_QDDL00000000.1; 2, NC_008260.1; 3, NC_018691.1; 4, NZ_AMRJ00000000.1; 5, NZ_ARXV00000000.1; 6, NZ_ARXU00000000.1; 7, NZ_CP004387.1; 8, NZ_CP012331.1.*

Strain	**1**	2	3	4	5	6	7	8	
1. *Pelagibaculum spongiae* Hp12^T		**65.75**	**65.55**	**65.69**	**65.75**	**65.96**	**65.85**	**65.30**	**ANI (%)**
2. *A. borkumensis* DSM 11573^T	**90.37**		72.28	75.20	75.51	77.22	70.61	72.16	
3. *A. dieselolei* B5^T	**90.14**	94.44		74.66	73.17	74.14	72.91	93.22	
4. *A. hongdengensis* A-11-3^T	**89.79**	96.74	94.95		76.66	78.16	72.73	74.68	
5. *A. nanhaiticus* 19-m-6^T	**89.40**	96.74	95.48	97.27		77.43	71.63	73.28	
6. *A. jadensis* T9^T	**89.44**	97.93	94.60	96.85	97.52		71.99	74.14	
7. *A. pacificus* W11-5^T	**88.53**	92.90	93.77	92.97	94.36	92.90		72.67	
8. *A. xenomutans* JC109^T	**88.95**	94.02	99.64	94.52	95.16	94.16	94.36		

16S rRNA sequence similarity (%)

Hp12^T can be phenotypically differentiated from members of the genus *Alcanivorax* and members of other closely related families in the order *Oceanospirillales* based on cell size, maximum salinity tolerance, temperature growth range, catalase activity and DNA G + C content (Table 3). The concentrations of major fatty acids ($C_{16:1}\omega7c$, $C_{18:1}\omega7c$ and $C_{16:0}$)

found in Hp12^T were clearly differentiated from the major fatty acid concentrations in other related species (Table 1 and Table 3).

***Table 3**: Characteristics that differentiate Hp12^T from type species and type strains of related genera. Strains: 1, Hp12^T (data from this study); 2, Alcanivorax borkumensis DSM 11573^T (data from this study); 3, A. dieselolei B5^T [18]; 4, A. xenomutans JC109^T [17]; 5, A. marinus R8-12^T [19]; 6, Kangiella koreensis SW-125^T [20]; 7, Halomonas elongata 1H9^T [21–23]. ND, no data available; NA, not applicable; +, positive; -, negative; DPG, diphosphatidylglycerol; PG, phosphatidylglycerol; PE, phosphatidylethanolamine; dDDH values are estimated based on Formula 2 of GGDC 2.1.*

Characteristic	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Cell shape	rod	rod	rod	rod	rod	rod	rod
Cell length [µm]	4.0-35	1.5-2.5	0.8-2.0	1.8	1.0-1.2	1.5-4.5	2.8-5.2
Max. salinity for growth (% w/v)	3	12	15	20	15	12	20
Optimal salinity range for growth (% w/v)	1-2	3-10	3-7.5	2-5	3	2-3	3-8
Temperature range for growth (°C)	4-27	4-35	15-45	25-40	10-42	4-43	4-45
Optimal growth temperature (°C)	20-22	25-30	28	30	28	30-37	30-35
Catalase	-	+	+	+	+	+	+
DNA G+C content (mol%)	43.0	54.7	62.1	54.5	66.1	44	60.5
Major fatty acids (% of total and > 10 %)	C16:1ω7c (55.3), C18:1ω7c (12.4), C16:0 (10.8)	C18:1ω7c (28.1), C16:0 (24.1), C16:1ω7c (22.2)	C16:0 (32.1), C18:1ω7c (22.41), C19:0cyclo ω8c (14.3), C16:1ω7c (11.3)	C16:0 (31.0), C19:0cyclo ω8c (18.0), C12:0 (10.2)	C16:0 (31.8), C18:1ω7c (20.3), C19:0ω8c cyclo (15.8)	iso-C15:0 (57.6), iso-C11:0 3OH (10.5)	C18:1ω7c (56.9), C16:0 (22.4)
ANI (%) to Hp12^T	NA	65.75	65.55	65.30	ND	66.63	65.5
dDDH (%) to Hp12^T	NA	34.3	34.3	37.2	ND	27.2	32.8

A distinct phylogenetic relationship, low 16S rRNA gene sequence similarity to related species and several differential phenotypic characteristics demonstrated that strain Hp12^T belongs to a new species in a new genus of the class *Gammaproteobacteria*, for which the name *Pelagibaculum spongiae* is proposed.

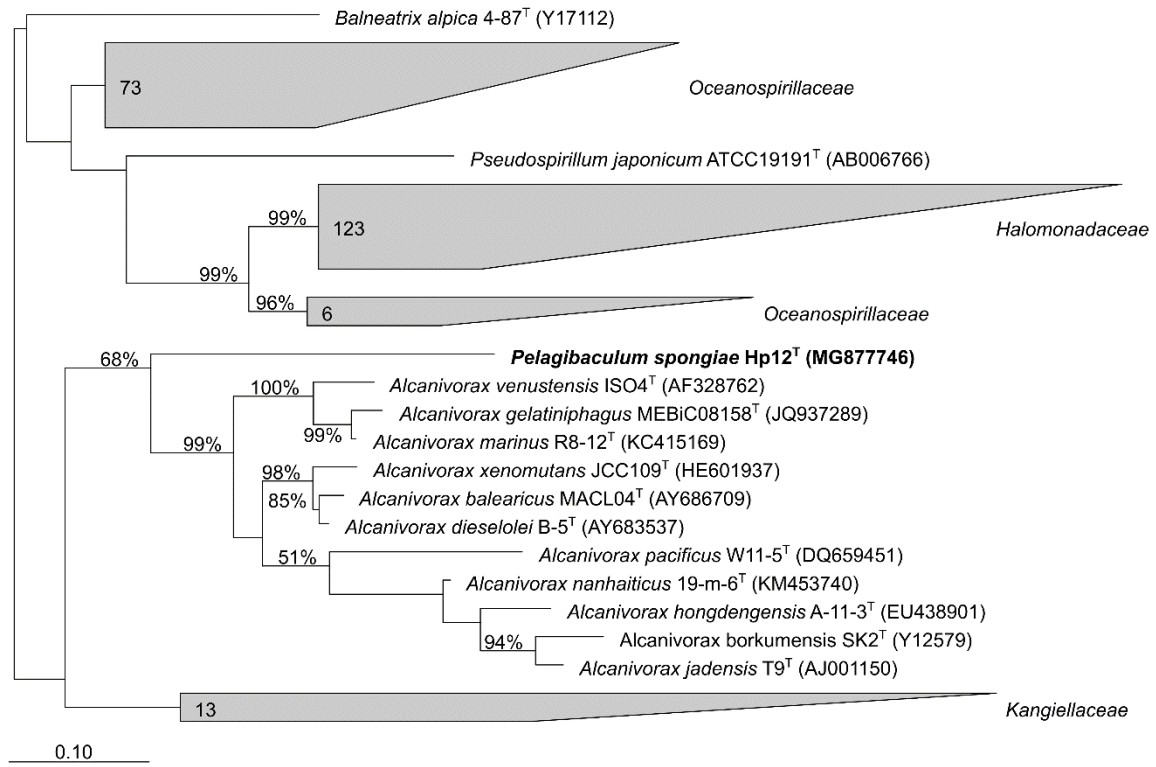


Figure 2*: Maximum-likelihood phylogenetic tree based on 565 16S rRNA gene sequences showing the relationship of strain Hp12^T to related members in the phylum Proteobacteria. For better clarity a subset of sequences are shown and type strains of the same taxonomic rank are grouped, with the number of sequences per group shown next to the group node. When not grouped, family names are given next to the species name. Phylogenetic analysis was performed using ARB [14] after sequence alignment using the SINA alignment service against the global SILVA SSU alignment [12]. Bootstrap values (>50%) are given as percentages at the branching points and are based on 100 resamplings. Bar: 0.10 substitutions per nucleotide position.*

Description of *Pelagibaculum* gen. nov.

Pelagibaculum (Pe.la.gi.ba'cu.lum. L. n. *pelagus* the sea; L. neut. n. *baculum* rod, stick; N.L. neut. n. *Pelagibaculum* rod-shaped bacterium from the sea). Cells are Gram-stain-negative, aerobic, rod-shaped, catalase-negative and oxidase positive. The predominant cellular fatty acids are $C_{16:1}\omega7c$, $C_{18:1}\omega7c$, and $C_{16:0}$. The predominant respiratory lipoquinones are Q8 and MK8. The major polar lipids are Phosphatidylglycerol and Phosphatidylethanolamine. The DNA G + C content of the type species is 43.0 mol%, determined from the draft genome sequence. Phylogenetically, the genus is affiliated with the class *Gammaproteobacteria* and related to the genus *Alcanivorax.* The type and only species is *Pelagibaculum spongiae*.

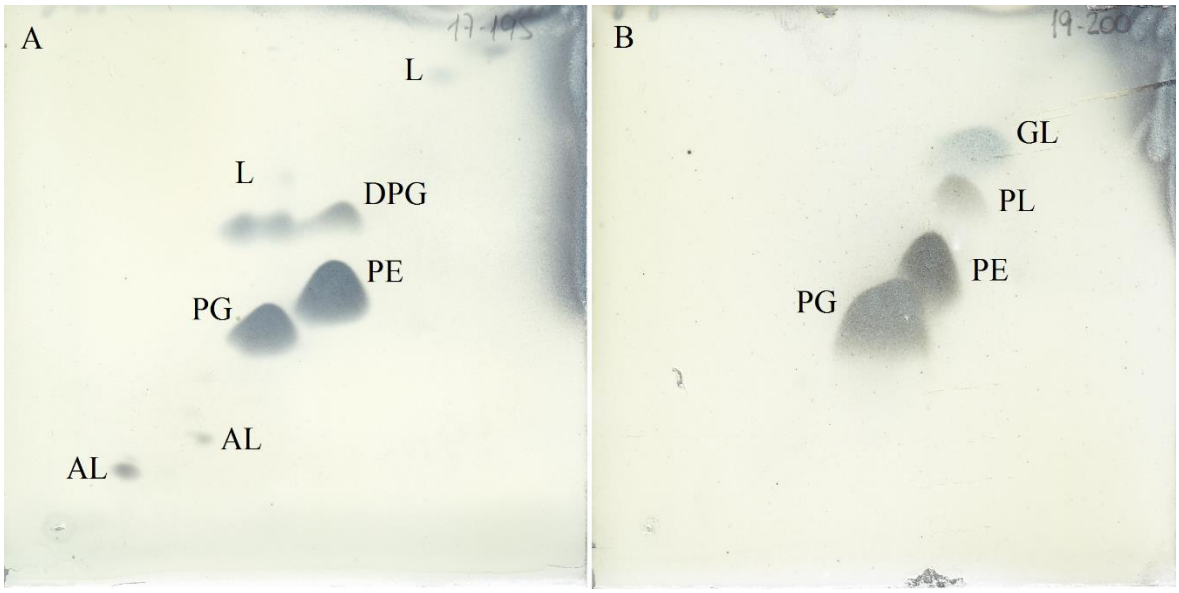


Figure 3*: TLC of polar lipid analysis for Hp12^T (A) and* Alcanivorax borkumensis *DSM 11573^T (B).* PG, phosphatidylglycerol; PE, phosphatidylethanolamine; DPG, diphosphatidylglycerol; GL, glycolipid; PL, phospholipid AL, aminolipid; L, uncharacterized lipid.

Description of *Pelagibaculum spongiae* sp. nov.

Pelagibaculum spongiae (spon'gi.ae. L. gen. n. *spongiae* of a sponge, pertaining to the isolation source of the type strain). In addition to the properties given in the genus description *Pelagibaculum spongiae* exhibits following characteristics: Cells are 1.0 – 1.2 µm wide and 4 - 35 µm long (typically 8 µm). Colonies are 1-2 mm in diameter, circular, raised and tan colour after incubation on MA at 20 °C for 3 days. Growth occurs between 4 and 28 °C with an optimal growth between 20 - 22 °C. Optimal pH for growth is between pH 7 - 8. Growth occurs at pH 6.5 and 9.5 but not at pH 6.0 or 10.0, and at NaCl concentrations from 1 - 3 % with an optimal growth rate at 1 - 2 % NaCl. No growth under anaerobic conditions on MA or MA supplemented with nitrate, nitrite, fumarate, sulphite, thiosulphate or sulphate. Neither casein, Tween 20 nor starch are hydrolysed. In the API ZYM tests, alkaline phosphatase, esterase, esterase lipase, leucine arylamidase, acid phosphatase and naphthol-AS-BI-phosphohydrolase are positive. Lipase, valine arylamidase, cystine arylamidase, trypsin, α-chymotrypsin, α-galactosidase, β-galactosidase, β-glucuronidase, α-glucosidase, β-glucosidase, N-acetyl-β-glucosaminidase, α-mannosidase and α-fucosidase are negative. In API 20NE strips, nitrate to nitrite reduction, hydrolysis of esculin and hydrolysis of gelatin are positive. Indole production, D-glucose fermentation, arginine dihydrolase, urease, β-galactosidase, and the assimilation of D-glucose, L-arabinose, D-mannose, D-mannitol, N-acetyl-glucosamine, D-maltose, potassium gluconate, D-maltose, potassium gluconate, capric acid, adipic acid, malic acid, trisodium citrate and phenylacetic acid are negative. The major respiratory lipoquinones are Q8 and MK8 with minor amounts of Q7. The polar lipid profile of strain Hp12^T contained phosphatidylglycerol and phosphatidylethanolamine as the major polar lipids and diphosphatidylglycerol, two uncharacterized aminolipids, and three uncharacterized lipids as the minor components.

The type strain is Hp12^T (=DSM 104963^T = CECT 9367^T) and was isolated from a marine sponge collected in an intertidal area in Seltjarnarnes, Iceland. The 16S rRNA sequence of strain Hp12^T is deposited under DDBJ/ENA/GenBank accession number MG877746.

Funding information

This work was supported by the EC Marie Curie Actions ITN Project No. 607786 BluePharmTrain.

Acknowledgements

We would like to acknowledge editorial support and thank Birgir Jóhannesson for help with the scanning electron microscopy and Bernhard Schink for support with the nomenclature.

References

1. **Taylor MW, Radax R, Steger D, Wagner M**. Sponge-Associated Microorganisms: Evolution, Ecology, and Biotechnological Potential. *Microbiol Mol Biol Rev* 2007;71:295–347.
2. **Leinonen R, Sugawara H, Shumway M**. The Sequence Read Archive. *Nucleic Acids Res* 2011;39:D19–D21.
3. **Lagkouvardos I, Joseph D, Kapfhammer M, Giritli S, Horn M, *et al.*** IMNGS: A comprehensive open resource of processed 16S rRNA microbial profiles for ecology and diversity studies. *Sci Rep* 2016;6:33721.
4. **Cowan ST, Steel KJ**. *Cowan and Steel's manual for the identification of medical bacteria*. Cambridge university press; 2004.
5. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, *et al.*** SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J Comput Biol* 2012;19:455–477.
6. **Lagesen K, Hallin P, Rødland EA, Stærfeldt H-H, Rognes T, *et al.*** RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 2007;35:3100–3108.
7. **Yoon S-H, Ha S, Lim J, Kwon S, Chun J**. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 2017;110:1281–1286.
8. **Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M**. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14:60.
9. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ**. Basic local alignment search tool. *J Mol Biol* 1990;215:403–410.
10. **Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Ostell J, *et al.*** GenBank. *Nucleic Acids Res* 2018;46:D41–D47.
11. **Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y, *et al.*** Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* 2017;67:1613–1617.
12. **Pruesse E, Peplies J, Glöckner FO**. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 2012;28:1823–1829.
13. **Yilmaz P, Parfrey LW, Yarza P, Gerken J, Pruesse E, *et al.*** The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic Acids Res* 2014;42:D643–D648.
14. **Ludwig W, Strunk O, Westram R, Richter L, Meier H, *et al.*** ARB: a software environment for sequence data. *Nucleic Acids Res* 2004;32:1363–1371.
15. **Felsenstein J**. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. *http://www.evolution.gs.washington.edu/phylip.html*.

http://ci.nii.ac.jp/naid/10027221536/en/ (2004, accessed 23 January 2019).
16. **Guindon S, Gascuel O**. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 2003;52:696–704.
17. **Rahul K, Sasikala C, Tushar L, Debadrita R, Ramana C V.** Alcanivorax xenomutans sp. nov., a hydrocarbonoclastic bacterium isolated from a shrimp cultivation pond. *Int J Syst Evol Microbiol* 2014;64:3553–3558.
18. **Liu C, Shao Z**. Alcanivorax dieselolei sp. nov., a novel alkane-degrading bacterium isolated from sea water and deep-sea sediment. *Int J Syst Evol Microbiol* 2005;55:1181–1186.
19. **Lai Q, Wang J, Gu L, Zheng T, Shao Z**. Alcanivorax marinus sp. nov., isolated from deep-sea water. *Int J Syst Evol Microbiol* 2013;63:4428–4432.
20. **Yoon J-H, Oh T-K, Park Y-H**. Kangiella koreensis gen. nov., sp. nov. and Kangiella aquimarina sp. nov., isolated from a tidal flat of the Yellow Sea in Korea. *Int J Syst Evol Microbiol* 2004;54:1829–1835.
21. **Vreeland RH, Litchfield CD, Martin EL, Elliot E**. Halomonas elongata, a New Genus and Species of Extremely Salt-Tolerant Bacteria. *Int J Syst Bacteriol* 1980;30:485–495.
22. **Giordano A, Vella FM, Romano I, Gambacorta A**. Structural elucidation of a novel phosphoglycolipid isolated from six species of Halomonas. *J Lipid Res* 2007;48:1825–31.
23. **Miao C, Jia F, Wan Y, Zhang W, Lin M, *et al.*** Halomonas huangheensis sp. nov., a moderately halophilic bacterium isolated from a saline-alkali soil. *Int J Syst Evol Microbiol* 2014;64:915–920.

Paper V

Tenacibaculum islandicus sp. nov., isolated from a marine sponge contains secondary metabolite gene clusters

Stephen Knobloch, Ragnar Jóhannsson, Viggó Þór Marteinsson

Manuscript

Author contributions:

Conceived and designed the study: SK, RJ, VM
Performed sampling: SK
Performed laboratory work: SK
Analysed data: SK
Wrote the manuscript: SK
Reviewed and edited the manuscript: SK, RJ, VM

Abstract

A rod-shaped bacterium, designated Hp32^T, was isolated from a marine sponge in the intertidal zone off the coast of Seltjarnarnes (64° 16' N 22° 00' W), Iceland. Analysis of the full length 16S rRNA gene sequence revealed an affiliation with the genus *Tenacibaculum*. The sequence similarities between Hp32^T and type strains of the genus *Tenacibaculum* were between 93.92 % and 97.15 % with the highest similarity to *Tenacibaculum haliotis* strain RA3-2^T. Cells were gram-stain-negative, motile, oxidase- and catalase-negative, mesophilic and aerobic. Optimal growth of strain Hp32^T was at 25 °C, at pH 8 and in the presence of 2 % (w/v) NaCl. The isolate was able to hydrolyse starch, but not gelatin, and was able to grow on succinate as the only carbon source. Strain Hp32^T had a draft genome size of 4.69 Mbp with a DNA G + C content of 29.3 mol%. From the taxonomic information and phenotypic properties obtained in this study, it is proposed that strain Hp32^T be placed into the genus *Tenacibaculum* with the species name *Tenacibaculum islandicus* sp. nov. The type strain of *Tenacibaculum islandicus* is Hp32^T (= ISCAR6971^T).

Marine sponges are one of the most prolific sources of novel marine natural product discoveries [1]. Sponges host a large diversity of microorganisms which are often the producers of compounds with biotechnological and pharmaceutical potential [2]. A screening for novel bacteria from marine sponges was conducted in September 2016. One strain, designated Hp32^T, isolated from the marine sponge *Halichondria panicea* collected in the intertidal area of Seltjarnarnes, Iceland, showed highest 16S rRNA gene sequence similarity to members of the genus *Tenacibaculum* (< 97.2 %), but branched separately from the members of this genus in a phylogenetic tree analysis. The genus *Tenacibaculum* was first described by Suzuki *et al*. (2001) [3] and currently contains 28 validly published species (http://www.bacterio.net/tenacibaculum.html). The type species is *Tenacibaculum maritimum*, originally published as *Flexibacter maritimus* [4]. Type species of the genus were isolated from seawater or marine organisms, are rod-shaped and have a GC content of 29.8 to 35.3 %. The objective of the present study was to determine the taxonomic position of strain Hp32^T using a polyphasic approach.

A sponge, identified as *Halichondria panicea*, was collected from the intertidal area in Seltjarnarnes, Iceland (64° 16' N 22° 00' W). The surface of the sponge was rinsed multiple times with sterile seawater to remove bacteria attached to the sponge surface, ground in sterile seawater and subsequently used for bacterial isolation. Strain Hp32^T was obtained on Marine Agar 2216 (MA, BD Difco) by standard dilution plating technique and was routinely cultivated at 20°C on MA or in MB.

Cell morphology was observed on a BX51 light microscopy (Olympus). Gram reaction, oxidase- and catalase-reaction, optimal growth temperature, pH range, NaCl tolerance and antibiotic susceptibility were performed on a Bioscreen C (Oy Growth Curves) as previously described [5]. Hydrolysis of gelatin, casein, Tween 20 and starch were investigated on MA at 20 °C using the substrate concentrations according to Cowan and Steel [6]. Susceptibility to antibiotics was investigated as previously described [5] against ampicillin, apramycin, chloramphenicol, enteromycin, kanamycin, nalidixic acid, penicillin, polymyxin B, spectinomycin, streptomycin, tetracycline, vancomycin and zeocin. The API ZYM and 20NE system (bioMérieux) was used to test for enzyme activities and other physiological properties according to the manufacturer's instructions, with the exception of cells being suspended in artificial seawater (450 mM NaCl, 10 mM KCl, 9 mM $CaCl_2$, 14 mM $MgCl_2$, 8 mM $MgSO_4$). Growth on single carbon sources was tested against 26 compounds dissolved as 1% (w/v) in basal medium (w/v: 2.5% NaCl, 0.2% KCl, 0.002% $MgSO_4$, 0.1% KNO_3, 0.1% $(NH_4)_2HPO_4$, 0.05% KH_2PO_4). Antimicrobial activity of aqueous and organic extracts of Hp32^T was tested against the strains *Escherichia coli* ATCC 8739, *Staphylococcus aureus* ATCC 6538 and *Pseudomonas aeruginosa* ATCC 9027. In short, Hp32^T was grown in MB at 22 °C until the late exponential phase and sonicated in a Sonifier 250 (Branson Ultrasonics) for 3 min with the output control set to 4. The suspension was then extracted twice using water-saturated butanol with 10 % (v/v) methanol. The aqueous phase was freeze dried and the organic phase concentrated on a rotary evaporator at 40°C and dissolved in Dimethyl sulfoxide (DMSO) to obtain a final concentration of 25 mg ml^{-1}. Both extracts were added to Brain-Heart-Infusion media at concentrations of 0, 0.16, 0.33, 0,66 and 0.98 % (v/v) and growth inhibition against the test strains was evaluated on a Bioscreen C.

Chromosomal DNA was extracted from a culture of Hp32^T grown in MB at 20 °C until the late exponential phase using the MasterPure DNA purification Kit (Epicentre) according to the manufacturer's instructions for total DNA purification. Genomic libraries were constructed with the Nextera XT DNA Library Prep Kit (Illumina) according to the

manufacturer's instructions and sequenced on an Illumina MiSeq platform using 2 x 300 paired-end v3 chemistry. Raw sequence reads were quality trimmed with Trimmomatic [7] and assembled in SPAdes version 3.9.1 [8] in paired-end mode with default settings. The resulting draft genome had a size of 4.69 Mbp, DNA G + C content of 29.3 % and an average genome coverage of 150X. The full length 16S rRNA gene sequence of strain Hp32^T was predicted using the sotware RNAmmer [9]. Digital DNA-DNA hybridisation (dDDH) was calculated by the Genome-to-Genome Distance Calculator 2.1 on http://ggdc.dsmz.de/ [10] and Average Nucleotide Identity (ANI) analysis was computed on EZBioCloud [11]. Presence of secondary metabolite biosynthetic gene clusters was tested through the web-service of antiSMASH [12].

Cells of strain Hp32^T were Gram-stain-negative, motile, straight rods (0.5 - 1 µm wide and 2 – 4 µm long). Strain Hp32^T was oxidase- and catalase-negative. After 3 days at 20 °C, colonies grown on MA plates were 2-4 mm in diameter, circular with uneven edge, raised and yellow in colour. Strain Hp32^T was able to grow at 4 – 28 °C but not at 30 °C. Growth occurred at NaCl concentration of 1 – 3 % and pH 6.0 – 9.0 but not above or below these limits. The test for hydrolysis of starch was positive, but for casein, gelatin and Tween 20 it was negative. Positive enzyme activity detected on the API ZYM system were alkaline phosphatase, esterase, esterase lipase, lipase and valine arylamidase. All reactions on the API 20NE test were negative. Strain Hp32^T was susceptible to chloramphenicol (≥ 10 µg ml^{-1}), tetracycline (≥ 10 µg ml^{-1}), penicillin (≥ 10 µg ml^{-1}), streptomycin (≥ 30 µg ml^{-1}), nalidixic acid (100 µg ml^{-1}), spectinomycin (100 µg ml^{-1}) and ampicillin (100 µg ml^{-1}), but resistant to all other tested antibiotics and concentrations. Analysis of secondary metabolite biosynthetic gene clusters on AntiSMASH showed that strain Hp32^T contained Siderophone, Terpene, T3pks, Nrps and Lantipeptide-Ladderane producing gene clusters. Antimicrobial activity against *Escherichia coli* ATCC 8739, *Staphylococcus aureus* ATCC 6538 and *Pseudomonas aeruginosa* ATCC 9027 was not detected. Physiological and morphological characteristics of strain Hp32^T and other type strains of the genus *Tenacibaculum* are presented in Table 2.

The full length 16S rRNA gene sequence (1508 nt) of strain Hp32^T was compared to available sequences through identity analysis performed on EzBioCloud [13], revealing that strain Hp32^T was closely related to members of the genus *Tenacibaculum*. The closest neighbours in the EzTaxon database with full length sequences were *Tenacibaculum haliotis* RA3-2^T (1399/1440, 97.15 %), *Tenacibaculum ovolyticum* JDTF-31^T (1397/1440, 97.01 %) and *Tenacibaculum insulae* (1394/1442, 96.67 %). The full length 16S rRNA gene sequence of strain Hp32^T and all type strains of the genus *Tenacibaculum* were aligned to the global SSU SILVA alignment with SINA [14] and imported into the ARB software [15]. Related sequences were selected and neighbour joining [15], maximum parsimony [16] and maximum-likelihood [17] trees constructed. Tree topology was tested with 30, 40 and 50 % positional conservatory filters and validated with 100 bootstraps. Based on all phylogenetic trees, strain Hp32^T fell within a monophyletic clade consisting of *Tenacibaculum aesturariivivum*, *T. dicentrarchi* and *T. ovolyticum* with bootstrap support above 50% (Figure 1).

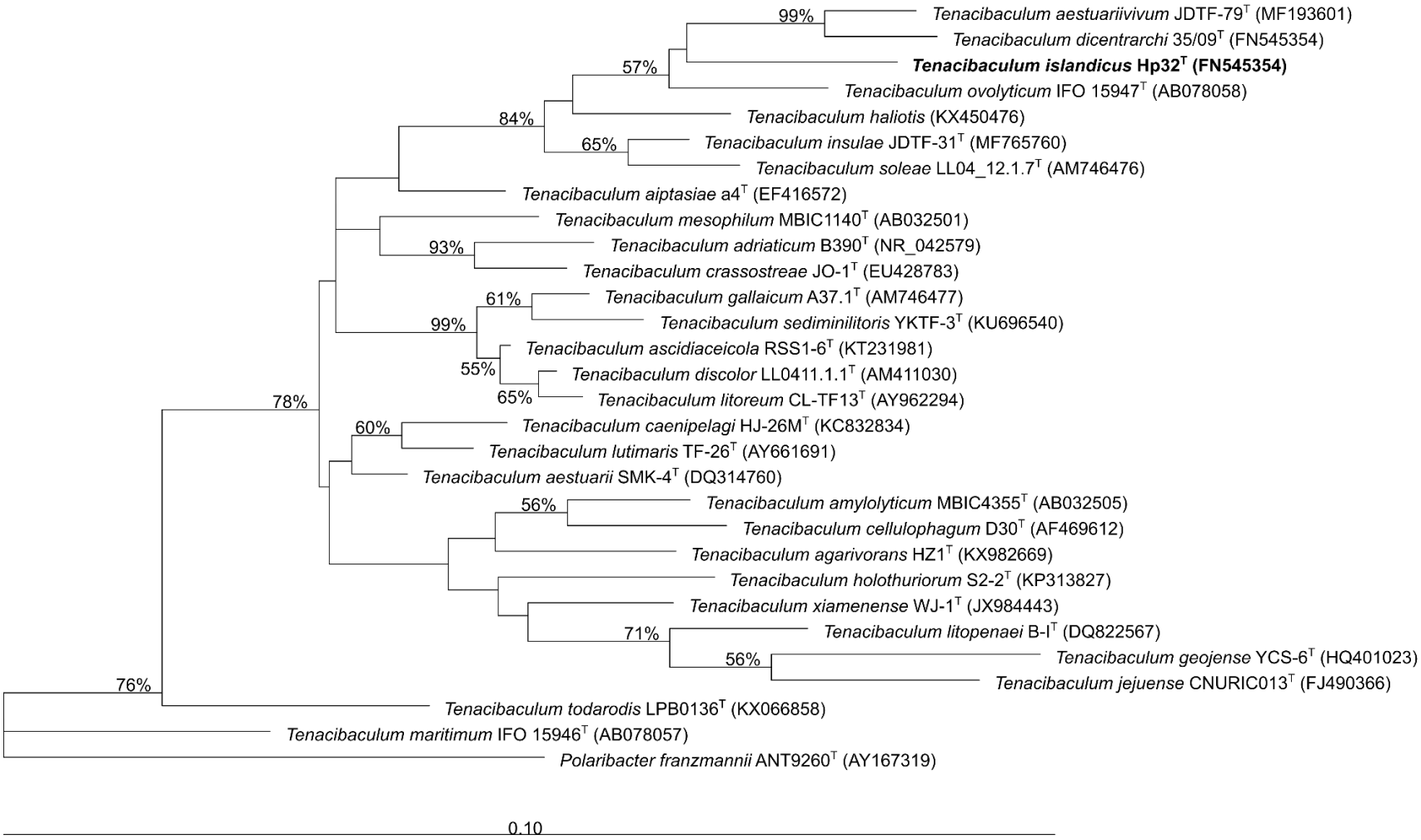


Figure 1: *Maximum-likelihood phylogenetic tree showing the relationship of strain Hp32^T to related members in the genus* Tenacibaculum*.*
Bootstrap values (>50%) are given as percentages at the branching points and are based on 100 resamplings. Bar: 0.10 substitutions per nucleotide position.

Table 1*: Characteristics that differentiate Hp32^T from type strains of related genera. Strains: 1, Hp32^T (data from this study); 2,* T. ovolyticum *DSM 18103^T [3, 18]; 3,* T. dicentrarchi *AY7486TD^T [19]; 4,* T. soleae *UCD-KL19^T [20]; 5,* T. agarivorans *HZ1^T [21]. ND, no data available; +, positive; -, negative; * does not grow in media supplemented with NaCl only.*

Characteristic	**1**	**2**	**3**	**4**	**5**
Cell shape	rod	rod	rod	rod	rod
Cell length [µm]	2-4	2-5	2-40	2-25	1-4
Colony shape	Uneven edge	Regular edge	Uneven edge	Uneven edge	Irregular edges
Colony color	Yellow	Pale yellow	Pale yellow	Yellow	Beige
Salinity range for growth (% w/v)	1-3	1-3	*	*	2-3
Temperature range for growth (°C)	4-28	4-25	4-30	14-30	15–37
Optimal growth temperature (°C)	25	ND	22-25	22-25	28
Oxidase	-	+	+	+	-
Catalase	-	+	+	+	+
Growth with 1% peptone	-	+	ND	ND	ND
Degradation of starch	+	-	-	-	-
Degradation of gelatin	-	+	+	+	+
Nitrate reduction	-	+	-	+	+
DNA G+C content (mol%)	29.3	31-32	30.3	29.8	31.8

Hp32^T can be phenotypically differentiated from other type strains of the genus *Tenacibaculum* based on cell size, oxidase and catalase activity, degradation of starch and gelatin, nitrate reduction and DNA G + C content (Table 1), as well as based on ANI values and dDDH calculations (Table 2). A distinct phylogenetic relationship, low 16S rRNA gene sequence similarity to related type strains and several differential phenotypic characteristics demonstrated that strain Hp32^T belongs to a new species in the genus *Tenacibaculum*, for which the name *Tenacibaculum islandicus* is proposed.

Description of *Tenacibaculum islandicus* sp. nov.

Tenacibaculum islandicus (is.lan′di.cus. N.L. masc. adj. *islandicus* from Iceland, pertaining to the isolation source of the type strain). Cells are 0.5 - 1 µm wide and 2 – 4 µm long. After incubation on MA at 20°C for three days colonies are 2-4 mm in diameter, circular with uneven edge, raised and yellow in colour. Growth occurred at pH 6 – 9, NaCl concentrations from 1 – 3% and at a temperature of 4 – 28°C. Starch was hydrolysed, but not gelatine, casein or Tween 20. Antibiotic susceptibility to chloramphenicol, tetracycline, penicillin, streptomycin, nalidixic acid, spectinomycin and ampicillin. In the API ZYM tests, alkaline phosphatase, esterase, esterase lipase, lipase and valine arylamidase are positive. In API 20NE strip all tests are negative.

The type strain is Hp32^T (=ISCAR6971^T) and was isolated from a marine sponge collected in an intertidal area in Seltjarnarnes, Iceland. The 16S rRNA sequence of strain Hp32^T is deposited under DDBJ/ENA/GenBank accession number MK633875.

***Table 2**: dDDH estimate (%) and ANI (%) between strain Hp32^T and members of the genus* Tenacibaculum*. Genomes (GenBank/RefSeq no.) used for calculations: 1, SIHP00000000; 2, NZ_AUMF00000000.1; 3, NZ_CP013671.1; 4, NZ_MAKX00000000.1; 5, NZ_MSMP00000000.1.*

Strain	**1**	2	3	4	5	
1. *T. islandicus* Hp32^T		**76.04**	**75.67**	**76.40**	**74.33**	**ANI (%)**
2. *T. ovolyticum* DSM 18103^T	**20.30**		80.49	82.86	74.16	
3. *T. dicentrarchi* AY7486TD^T	**20.50**	24.30		80.26	73.97	
4. *T. soleae* UCD-KL19^T	**20.30**	26.70	24.00		74.51	
5. *T. agarivorans* HZ1^T	**19.40**	19.20	19.90	19.60		
	dDDH estimate (%)					

Funding information

This work was supported by the EC Marie Curie Actions ITN Project No. 607786 BluePharmTrain.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

1. **Sacristán-Soriano O, Becerro MA, Becerro MA**. Publication impact in sponge chemical and microbial ecology. *Sci Mar* 2016;80:555.

2. **Piel J**. Metabolites from symbiotic bacteria. *Nat Prod Rep* 2009;26:338–362.

3. **Suzuki M, Nakagawa Y, Harayama S, Yamamoto S**. Phylogenetic analysis and taxonomic study of marine Cytophaga-like bacteria: proposal for Tenacibaculum

gen. nov. with Tenacibaculum maritimum comb. nov. and Tenacibaculum ovolyticum comb. nov., and description of Tenacibaculum mesophilum sp. nov. and Tenacibaculum amylolyticum sp. nov. *Int J Syst Evol Microbiol* 2001;51:1639–1652.

4. **Wakabayashi H, Hikida M, Masumura K**. Flexibacter maritimus sp. nov., a Pathogen of Marine Fishes. *Int J Syst Bacteriol* 1986;36:396–398.

5. **Knobloch S, Jóhannsson R, Marteinsson V**. Pelagibaculum spongiae gen. nov., sp. nov., isolated from a marine sponge in South-West Iceland. *Int J Syst Evol Microbiol*;(accepted).

6. **Cowan ST, Steel KJ**. *Cowan and Steel's manual for the identification of medical bacteria*. Cambridge university press; 2004.

7. **Bolger AM, Lohse M, Usadel B**. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30:2114–2120.

8. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, *et al.*** SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J Comput Biol* 2012;19:455–477.

9. **Lagesen K, Hallin P, Rødland EA, Stærfeldt H-H, Rognes T, *et al.*** RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 2007;35:3100–3108.

10. **Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M**. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14:60.

11. **Yoon S-H, Ha S, Lim J, Kwon S, Chun J**. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 2017;110:1281–1286.

12. **Blin K, Wolf T, Chevrette MG, Lu X, Schwalen CJ, *et al.*** antiSMASH 4.0—improvements in chemistry prediction and gene cluster boundary identification. *Nucleic Acids Res* 2017;45:W36–W41.

13. **Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y, *et al.*** Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* 2017;67:1613–1617.

14. **Pruesse E, Peplies J, Glöckner FO**. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 2012;28:1823–1829.

15. **Ludwig W, Strunk O, Westram R, Richter L, Meier H, *et al.*** ARB: a software environment for sequence data. *Nucleic Acids Res* 2004;32:1363–1371.

16. **Felsenstein J**. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. *http://www.evolution.gs.washington.edu/phylip.html*. http://ci.nii.ac.jp/naid/10027221536/en/ (2004, accessed 23 January 2019).

17. **Guindon S, Gascuel O**. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 2003;52:696–704.

18. **Hansen GH, Bergh O, Michaelsen J, Knappskog D**. Flexibacter ovolyticus sp. nov., a Pathogen of Eggs and Larvae of Atlantic Halibut, Hippoglossus hippoglossus L. *Int J Syst Bacteriol* 1992;42:451–458.

19. **Pineiro-Vidal M, Gijon D, Zarza C, Santos Y**. Tenacibaculum dicentrarchi sp. nov., a marine bacterium of the family Flavobacteriaceae isolated from European sea bass. *Int J Syst Evol Microbiol* 2012;62:425–429.

20. **Pineiro-Vidal M, Carballas CG, Gomez-Barreiro O, Riaza A, Santos Y**. Tenacibaculum soleae sp. nov., isolated from diseased sole (Solea senegalensis Kaup). *Int J Syst Evol Microbiol* 2008;58:881–885.

21. **Xu Z-X, Yu P, Mu D-S, Liu Y, Du Z-J**. Tenacibaculum agarivorans sp. nov., an agar-degrading bacterium isolated from marine alga Porphyra yezoensis Ueda. *Int J Syst Evol Microbiol* 2017;67:5139–5143.

Paper VI

VI

Endozoicomonas halichondricola sp. nov., isolated from the marine sponge *Halichondria panicea*

Stephen Knobloch, Ragnar Jóhannsson, Viggó Þór Marteinsson

Manuscript

Author contributions:

Conceived and designed the study: SK, RJ, VM
Performed sampling: SK
Performed laboratory work: SK
Analysed data: SK
Wrote the manuscript: SK
Reviewed and edited the manuscript: SK, RJ, VM

Abstract

A Gram-negative, mesophilic, rod-shaped bacterium, designated Hp36^T, was isolated from the marine sponge *Halichondria panicea* collected off the coast of Seltjarnarnes (64° 16' N 22° 00' W), Iceland. 16S rRNA gene sequence analysis showed that strain Hp36^T clustered closely with *Endozoicomonas numazuensis* strain DSM25634^T (96.79% sequence similarity) and *Endozoicomonas arenosclerae* Ab112^T (96.72% sequence similarity). The novel strain required NaCl for growth and grew optimally at 22°C and in the presence of 2% NaCl. Strain Hp36^T had a draft genome size of 5.49 Mbp, a DNA G + C content of 44.9 mol% and was distinguishable from other related type strains based on Average Nucleotide Identity values and digital DNA-DNA hybridisation. It is evident from the results of taxonomic and phenotypic analysis that strain Hp36^T represents a novel species in the genus *Endozoicomonas*, for which the name *Endozoicomonas halichondricola* sp. nov. is proposed. The type strain of *Endozoicomonas halichondricola* is Hp36^T (= ISCAR6976^T).

Marine sponges can host a large diversity of microorganisms, in some cases accounting for over 40% of the sponge body [1]. Whereas sponge-associated microbes can possess commensal symbiotic functions [2], the role of bacterial symbiont in sponges is not always clear [3]. Members of the genus *Endozoicomonas* are often found in the cultivable microbial community of sponges [4, 5] and other marine animals [6, 7]. The genus *Endozoicomonas* consists of nine validly published species, all of which were isolated from benthic marine animals, including corals [8–10], a bivalve [11], a nudibranch [12], an ascidian [13] and sponges [14, 15]. Genome analysis has shown that members of the genus *Endozoicomonas* participate in putative commensal activity with their animal hosts through host-associated protein and carbohydrate transport and cycling [16]. One novel strain, designated Hp36^T, was isolated from the marine sponge *Halichondria panicea* during a screening for novel bacteria from sponges in Iceland conducted in September 2016. Based on the partial 16S rRNA gene sequence, strain Hp36^T was detected in over 50% of *H. panicea* samples in a sponge microbiome dataset collected between 2014 and 2016 in Icelandic waters [17]. 16S rRNA gene sequence analysis showed that strain Hp36^T exhibited highest similarity to validly published members of the genus *Endozoicomonas* (96.79% to 93.72% sequence similarity). The aim of the present study was to determine the taxonomic position of strain Hp36^T using a polyphasic approach.

A specimen of *Halichondria panicea* was collected from a tidal pool in Seltjarnarnes, Iceland (64° 16' N 22° 00' W), rinsed on the outside with sterile seawater and used for bacterial isolation. Strain Hp36^T was obtained by standard dilution plating technique on Starch Yeast-Extract Peptone Sea Water Agar (10 g potato starch, 4 g Bacto Yeast Extract, 2 g Bacto Peptone, 15 g Bacto Agar in 1 l 0.2 µm filtered seawater) at 10 °C. Routine cultivation of Hp36^T was performed on MA or in MB at 20 °C. Cell morphology was observed on an Olympus BX51 light microscope. Optimal growth temperature, pH value and NaCl concentrations was evaluated in MB on a Bioscreen C (Oy Growth Curves) as previously described [18]. The Gram reaction, catalase and oxidase activity was conducted using the BD BBL Gram Stain Kit (BD), catalase reagent droppers (BD Difco) and DrySlide reagent slides (BD Difco) respectively according to the manufacturer's instructions. Anaerobic growth was tested on MA in a GasPak (BD) for 14 days at 20°C. Hydrolysis of casein, gelatine, Tween 20 and starch was investigated on MA at 20°C using the substrate concentrations according to Cowan and Steel (2004) [19]. Enzyme activities and other physiological or biochemical traits were tested on API ZYM and API 20NE systems (bioMérieux) according to the manufacturer's instructions, with the suspension media adapted to artificial seawater (450 mM NaCl, 10 mM KCl, 9 mM $CaCl_2$, 14 mM $MgCl_2$, 8 mM $MgSO_4$), or artificial seawater supplemented media.

Chromosomal DNA of strain Hp36^T was extracted from colonies grown on MA using the MasterPure DNA purification Kit (Epicentre) according to the manufacturer's instructions for total DNA purification. Whole genome shotgun sequencing was performed on an Illumina MiSeq desktop sequencer using the Nextera XT DNA Library Prep Kit (Illumina) and paired-end 300bp v3 sequencing chemistry. Quality filtering of raw reads was performed in Trimmomatic [20]. Trimmed reads were assembled in SPAdes version 3.9.1 [21] using paired-end mode and default settings. The resulting draft genome of Hp36^T had a length of 5.49 Mbp and an average genome coverage of 120X. The full length 16S rRNA gene sequence was predicted using the tool RNAmmer [22] and DNA G + C content was determined in the bioinformatics software Geneious (Biomatters). Average Nucleotide Identity (ANI) analysis was computed through the OrthoANI-usearch tool [23]. Digital

DNA-DNA hybridisation (dDDH) was estimated by the Genome-to-Genome Distance Calculator 2.1 on http://ggdc.dsmz.de/ [24].

Table 1: *Characteristics that differentiate Hp36^T from closely related type strains of the genus* Endozoicomonas. *Strains: 1, Hp36^T (data from this study); 2,* Endozoicomonas numazuensis *HC50^T [14]; 3,* Endozoicomonas arenosclerae *Ab112^T [15]; 4,* Endozoicomonas montiporae *CL-33^T [8]; and 5,* Endozoicomonas ascidiicola *AVMART05^T [13]. ND, no data available; +, positive; -, negative; w, weakly positive.*

Characteristics	**1**	**2**	**3**	**4**	**5**
Isolation source	Sponge	Sponge	Sponge	Coral	Ascidian
Cell shape	rod	rod	ND	rod	rod
Cell length [µm]	4 – 15	3.0 – 10.0	ND	1.0 – 3.0	1.2 – 11.3
Colony colour	Pale yellow	pale creamy white	cream coloured	Beige	Beige
Motility	+	-	+	+	+
Max. salinity for growth (% w/v)	3	5	5	3	5
Temperature range for growth (°C)	4 - 30	15 - 37	12 – 35	15 – 35	5 – 27
Anaerobic fermentation	-	+	-	-	+
Catalase	-	+	ND	+	w
Oxidase	+	+	ND	+	+
DNA G+C content (mol%)	44.9	48.3 – 48.7	47.6	50.0	46.7

Cells of strain Hp36^T were 1.5 – 2 µm wide, 4 – 15 µm long and rod-shaped. Cells were Gram-stain negative, motile, oxidase-positive and catalase-negative. Grown on MA for 3 days at 20°C, colonies were pale yellow, 2-4 mm in diameter, raised and circular. Strain Hp36^T grew at 4 and 30 °C but not below or above these temperatures. Hp36^T required NaCl for growth but did not grow above a NaCl concentration of 3 %. Optimal grow was observed at 22 °C, pH 8 and the presence of 2% NaCl. Hp36^T did not hydrolysis starch, casein, Tween 20 or gelatine. Positive activity detected by the API ZYM and 20NE systems were nitrate reduction, aesculin hydrolysis, alkaline phosphatase, C4 esterase, C8 esterase, acid phosphatase, naphthol-AS-BI-phosphohydrolase. The DNA G + C content of the Hp36^T draft genome was 44.9 %. Physiological and morphological characteristics of strain Hp36^T and other type strains of the genus *Endozoicomonas* can be found in Table 1.

The full length 16S rRNA gene sequence of strain Hp36^T consisting of 1531 nucleotides and showed closest sequence similarity to the validly published type strains *Endozoicomonas numazuensis* (1418/1465, 96.79 %), *Endozoicomonas arenosclerae* (1414/1462, 96.72 %), *Endozoicomonas montiporae* (1399/1465, 95.49 %) and *Endozoicomonas ascidiicola* (1398/1465, 95.43 %) on the EzTaxon database [25]. Neighbour joining [26], maximum

parsimony [27] and maximum-likelihood [28] phylogenetic trees were built in ARB [26] using the 16S rRNA sequences of Hp36^T and type strains of related species aligned against the global SSU SILVA alignment in SINA version 1.2.11 [29]. Based on all phylogenetic trees, strain Hp36^T formed a monophyletic clade with the type strains of *Endozoicomonas numazuensis* and *Endozoicomonas arenosclerae* with bootstrap values above 90 % (maximum-likelihood tree shown in Figure 1).

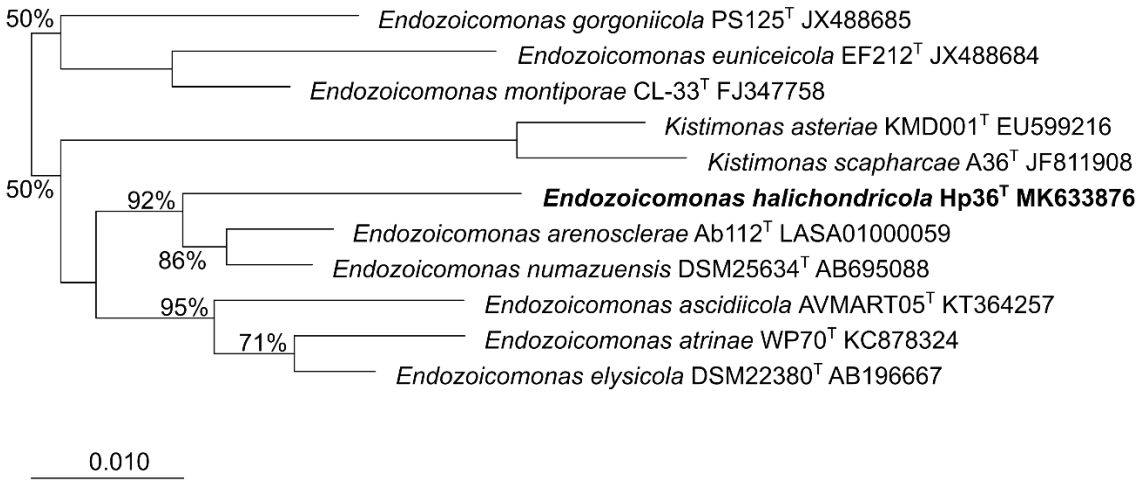


Figure 1: *Maximum-likelihood phylogenetic tree based on 16S rRNA gene sequences of strain Hp36^T and related type strains. Bootstrap values above 50 % are given as percentages at the branching points and are based on 100 resamplings. Bar: 0.010 substitutions per nucleotide position.*

Hp36^T can be phenotypically differentiated from other members of the genus *Endozoicomonas* based on cell size, maximum salinity tolerance, temperature growth range, catalase activity, anaerobic fermentation and DNA G + C content (Table 1). In addition, 16S rRNA gene sequence similarity below 97 % and a differentiation of ANI values and low dDDH estimates to other members of the genus *Endozoicomonas* (Table 2) clearly shows that strain Hp36^T represents a novel species in the genus *Endozoicomonas*, for which the name *Endozoicomonas halichondricola* is proposed.

Description of *Endozoicomonas halichondricola* sp. nov.

Endozoicomonas halichondricola (ha.li.chon.dri.co'la N.L. n. *Halichondria* name of a zoological genus; L. suff. *cola* dweller; N.L. n. *halichondricola Halichondria* dweller). Cells are 1.5 – 2 µm wide and 4 – 15 µm long. Colonies are pale yellow, 2-4 mm in diameter, raised and circular after incubation on MA for 3 days at 20 °C. Growth occurs between 4 and 30 °C with an optimal growth 22 °C. Optimal pH for growth is 8. Requires presence of NaCl, grows optimally in 2% NaCl, but does not grow above 3 % NaCl. No growth under anaerobic conditions. Neither casein, gelatin, Tween 20 nor starch are hydrolysed. Positive activity in the API ZYM and 20NE tests are nitrate reduction, aesculin hydrolysis, alkaline phosphatase, C4 esterase, C8 esterase, acid phosphatase, naphthol-AS-BI-phosphohydrolase.

The type strain is Hp36^T (= ISCAR6976^T) and was isolated from the marine sponge *Halichondria panicea* in Seltjarnarnes, Iceland. The 16S rRNA sequence of strain Hp36^T is deposited under DDBJ/ENA/GenBank accession number MK633876.

Table 2: *ANI value (%) and dDDH estimate (%) between strain Hp36T and availably genomes of type strains in the genus* Endozoicomonas. *Genomes (RefSeq no.) used for calculations: 1, SIHQ00000000 (this study); 2, NZ_JOKH00000000.1; 3, GCF_001562015.1; 4, NZ_JOJP00000000.1; 5, NZ_CP013251.1; 6, NZ_PJPV00000000.1; 7, NZ_LUKQ00000000.2; 8, GCF_001646945.1.*

Strain	**1**	2	3	4	5	6	7	8
1. *E. halichondricola* Hp36^T		**72.57**	**72.66**	**70.20**	**71.17**	**70.32**	**70.36**	**70.61**
2. *E. numazuensis* DSM 25634^T	**22.30**		79.62	70.89	72.18	71.13	71.03	71.04
3. *E. arenosclerae* Ab112^T	**22.10**	22.90		70.76	72.15	71.22	70.83	71.01
4. *E. elysicola* DSM 22380^T	**22.20**	24.60	24.10		70.93	80.28	83.98	74.44
5. *E. montiporae* CL-33^T	**22.80**	26.10	23.20	24.20		71.46	70.93	71.15
6. *E. acroporae* Acr-14^T	**20.60**	22.50	22.60	24.30	24.50		85.33	74.60
7. *E. atrinae* WP70^T	**20.50**	22.30	21.80	28.50	22.50	30.80		74.52
8. *E. ascidiicola* AVMART05^T	**22.40**	24.30	22.40	22.50	24.00	22.20	22.00	
	dDDH estimate (%)							

ANI (%)

Funding information

This work was supported by the EC Marie Curie Actions ITN Project No. 607786 BluePharmTrain.

Conflicts of interest

The authors declare that there are no conflicts of interest.

References

1. **Wilkinson CR**. Immunological Evidence for the Precambrian Origin of Bacterial Symbioses in Marine Sponges. *Proceedings of the Royal Society of London. Series B, Biological Sciences* 1984;220:509–518.
2. **Webster NS, Thomas T**. The Sponge Hologenome. *MBio* 2016;7:e00135-16.
3. **Taylor MW, Radax R, Steger D, Wagner M**. Sponge-Associated Microorganisms: Evolution, Ecology, and Biotechnological Potential. *Microbiol Mol Biol Rev* 2007;71:295–347.
4. **Neave MJ, Michell CT, Apprill A, Voolstra CR**. Whole-genome sequences of three symbiotic endozoicomonas strains. *Genome Announc* 2014;2:e00802-14.
5. **Mohamed NM, Colman AS, Tal Y, Hill RT**. Diversity and expression of nitrogen fixation genes in bacterial symbionts of marine sponges. *Environ Microbiol* 2008;10:2910–2921.

6. **Goffredi SK, Johnson SB, Vrijenhoek RC**. Genetic diversity and potential function of microbial symbionts associated with newly discovered species of Osedax polychaete worms. *Appl Environ Microbiol* 2007;73:2314–23.

7. **Du Z, Zhang W, Xia H, Lü G, Chen G**. Isolation and diversity analysis of heterotrophic bacteria associated with sea anemones. *Acta Oceanol Sin* 2010;29:62–69.

8. **Yang CS, Chen MH, Arun AB, Chen CA, Wang JT, *et al.*** Endozoicomonas montiporae sp. nov., isolated from the encrusting pore coral Montipora aequituberculata. *Int J Syst Evol Microbiol* 2010;60:1158–1162.

9. **Sheu S-Y, Lin K-R, Hsu M, Sheu D-S, Tang S-L, *et al.*** Endozoicomonas acroporae sp. nov., isolated from Acropora coral. *Int J Syst Evol Microbiol* 2017;67:3791–3797.

10. **Pike RE, Haltli B, Kerr RG**. Description of Endozoicomonas euniceicola sp. nov. and Endozoicomonas gorgoniicola sp. nov., bacteria isolated from the octocorals Eunicea fusca and Plexaura sp., and an emended description of the genus Endozoicomonas. *Int J Syst Evol Microbiol* 2013;63:4294–4302.

11. **Hyun D-W, Shin N-R, Kim M-S, Oh SJ, Kim PS, *et al.*** Endozoicomonas atrinae sp. nov., isolated from the intestine of a comb pen shell Atrina pectinata. *Int J Syst Evol Microbiol* 2014;64:2312–2318.

12. **Kurahashi M, Yokota A**. Endozoicomonas elysicola gen. nov., sp. nov., a γ-proteobacterium isolated from the sea slug Elysia ornata. *Syst Appl Microbiol* 2007;30:202–206.

13. **Schreiber L, Kjeldsen KU, Obst M, Funch P, Schramm A**. Description of Endozoicomonas ascidiicola sp. nov., isolated from Scandinavian ascidians. *Syst Appl Microbiol* 2016;39:313–318.

14. **Nishijima M, Adachi K, Katsuta A, Shizuri Y, Yamasato K**. Endozoicomonas numazuensis sp. nov., a gammaproteobacterium isolated from marine sponges, and emended description of the genus Endozoicomonas Kurahashi and Yokota 2007. *Int J Syst Evol Microbiol* 2013;63:709–714.

15. **Appolinario LR, Tschoeke DA, Rua CPJ, Venas T, Campeão ME, *et al.*** Description of Endozoicomonas arenosclerae sp. nov. using a genomic taxonomy approach. *Antonie Van Leeuwenhoek* 2016;109:431–438.

16. **Neave MJ, Apprill A, Ferrier-Pagès C, Voolstra CR**. Diversity and function of prevalent symbiotic marine bacteria in the genus Endozoicomonas. *Appl Microbiol Biotechnol* 2016;100:8315–8324.

17. **Knobloch S, Jóhannsson R, Marteinsson V**. Bacterial diversity in the marine sponge Halichondria panicea from Icelandic waters and host-specificity of its dominant symbiont "Candidatus Halichondribacter symbioticus". *FEMS Microbiol Ecol*;95. Epub ahead of print 10 November 2019. DOI: 10.1093/femsec/fiy220.

18. **Knobloch S, Jóhannsson R, Marteinsson V**. Pelagibaculum spongiae gen. nov., sp. nov., isolated from a marine sponge in South-West Iceland. *Int J Syst Evol Microbiol*;(accepted).

19. **Cowan ST, Steel KJ**. *Cowan and Steel's manual for the identification of medical bacteria*. Cambridge university press; 2004.

20. **Bolger AM, Lohse M, Usadel B**. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 2014;30:2114–2120.

21. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, *et al.*** SPAdes: A New

Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. *J Comput Biol* 2012;19:455–477.

22. **Lagesen K, Hallin P, Rødland EA, Stærfeldt H-H, Rognes T, *et al.*** RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res* 2007;35:3100–3108.

23. **Yoon S-H, Ha S, Lim J, Kwon S, Chun J**. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek* 2017;110:1281–1286.

24. **Meier-Kolthoff JP, Auch AF, Klenk H-P, Göker M**. Genome sequence-based species delimitation with confidence intervals and improved distance functions. *BMC Bioinformatics* 2013;14:60.

25. **Yoon S-H, Ha S-M, Kwon S, Lim J, Kim Y, *et al.*** Introducing EzBioCloud: a taxonomically united database of 16S rRNA gene sequences and whole-genome assemblies. *Int J Syst Evol Microbiol* 2017;67:1613–1617.

26. **Ludwig W, Strunk O, Westram R, Richter L, Meier H, *et al.*** ARB: a software environment for sequence data. *Nucleic Acids Res* 2004;32:1363–1371.

27. **Felsenstein J**. PHYLIP (Phylogeny Inference Package) version 3.6. Distributed by the author. *http://www.evolution.gs.washington.edu/phylip.html*. http://ci.nii.ac.jp/naid/10027221536/en/ (2004, accessed 23 January 2019).

28. **Guindon S, Gascuel O**. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Syst Biol* 2003;52:696–704.

29. **Pruesse E, Peplies J, Glöckner FO**. SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* 2012;28:1823–1829.