Insights into the Holobiont of the Early Branching Metazoan *Vaceletia* sp. and its Biomineralization Strategy

Dissertation
Zur Erlangung des mathematisch-naturwissenschaftlichen Doktorgrades
„Doctor rerum naturalium“
der Georg-August-Universität Göttingen

im Promotionsprogramm Geowissenschaften
der Georg-August University School of Science (GAUSS)

vorgelegt von
Juliane Germer

aus Nordhausen
Göttingen, 2017

Betreuungsausschuss:

Prof. Dr. Daniel J. Jackson, Abteilung Geobiologie, Geowissenschaftliches Zentrum der Georg-August-Universität Göttingen

Prof. Dr. Joachim Reitner, Abteilung Geobiologie, Geowissenschaftliches Zentrum der Georg-August-Universität Göttingen

Mitglieder der Prüfungskommission

Referent:
Prof. Dr. Daniel J. Jackson, Abteilung Geobiologie, Geowissenschaftliches Zentrum der Georg-August-Universität Göttingen

Korreferent:
Prof. Dr. Joachim Reitner, Abteilung Geobiologie, Geowissenschaftliches Zentrum der Georg-August-Universität Göttingen

Prof. Dr. Alexander Schmidt, Abteilung Geobiologie, Geowissenschaftliches Zentrum der Georg-August-Universität Göttingen

Prof. Dr. Gernot Arp, Abteilung Geobiologie, Geowissenschaftliches Zentrum der Georg-August-Universität Göttingen

Dr. Nico Posnien, Abteilung Entwicklungsbiologie, Johann-Friedrich Blumenbach-Institut für Zoologie und Anthropologie, GZMB, Georg-August-Universität Göttingen

Dr. Oliver Voigt, Paläontologie und Geobiologie, Department für Geo- und Umweltwissenschaften, Ludwig-Maximilians-Universität München

Tag der mündlichen Prüfung: 13.09.2017

Versicherung

Hiermit versichere ich an Eides statt, dass die Dissertation mit dem Titel „Insights into the Holobiont of an Early Branching Metazoan *Vaceletia* sp. and its Biomineralization Strategy“ selbständig und ohne unerlaubte Hilfe angefertigt wurde.

Göttingen, den 07.08.2017

Unterschrift:

Contents

Abstract

Sponges are evolutionary and ecologically very successful animals. They are abundant and significant members of benthic communities throughout the world's oceans. Having emerged during the Neoproterozoic era, sponges are among the most ancient of all extant animal lineages. Due to this basal branching position sponges are ideal models to gain insights into the origin and evolution of important metazoan traits. This thesis addresses two important aspects of sponge biology. The first part focuses on host-microbe interactions of the hypercalcifying demosponge *Vaceletia* sp. Like *Vaceletia*, most sponges comprise dense and diverse microbial communities that can constitute up to 50% of the sponge's biomass. The analysis and characterization of *Vaceletia* sp. holo-transcriptome, and the comparison to other sponge transcriptomes and genomes shows that this sponge interacts in various ways with its microbial community. The results of my study imply that the sponge immune system as well as eukaryotic–like proteins from bacteria play an important role in mediating interactions. By studying the underlying molecular mechanisms of lipid pathway components, I showed that short chain fatty acids and mid-chain branched fatty acids are most likely produced by the sponge's bacterial community, whereas long chain fatty acids are most likely synthesized by the sponge itself via elongation and desaturation of short-chain precursors. The second part of this thesis focuses on the biomineralization strategy employed by *Vaceletia* sp. By generating a comprehensive skeletal proteome of this sponge, I identified and characterized 40 proteins that most likely represent the majority of components playing an important role in the mineralization process of *Vaceletia*. The proteome contains components showing similarities to already identified proteins with a known role in biomineralization as well as novel components. The microbial community of *Vaceletia* sp. apparently plays a minimal role in directly contributing proteinaceous compounds to the skeleton formation in this sponge. This thesis represents the first investigation into the molecular mechanisms underlying sponge-microbe interactions and the biomineralization of the early branching *Vaceletia* sp. The studied interactions include innate immunity, eukaryotic-like proteins in bacteria and metabolic interactions. The results of my thesis expand our knowledge of the complex gene repertoire of sponges, show the importance of metabolic interactions between *Vaceletia* and its microbial community, and give insight into the biomineralization strategy of this sponge.

Chapter 1:

Introduction

1.1 Preface to the life of sponges

Sponges (Porifera) are important members of marine benthic communities throughout the world's oceans and occur at all depths [1]. They are often prominent faunal components of temperate and tropical reefs [2] and can even dominate the arctic benthos [3]. Apart from their numerical abundance and biomass dominance, they fulfill a number of important functional roles in these ecosystems by influencing benthic and pelagic processes [1]. Sponges are among the most ancestral animals (metazoans) [4, 5] having evolved during the Neoproterozoic [6]. For understanding the evolution of complex metazoan traits such as neurons, muscles, epithelia, and biomineralization it is essential to resolve the deep metazoan relationships. However, relationships among basal metazoan branches remain controversial, especially regarding the non-bilaterian taxa Porifera, Ctenophora, Cnidaria and Placozoa [7], and recent studies either confirm the view of sponges being the sister group to all other animals [8–10], or reject it, recovering instead ctenophores as sister group [11–13].

Sponges possess a simple bodyplan (Fig. 1) that has remained essentially unchanged during their evolution. However, analysis of sponge transcriptomes and genomes revealed that the genetic repertoire of sponges is more complex than their morphologically simple bodyplan implies [10, 14, 15]. Most of the metazoan genes involved in complex gene pathways are present in sponges, although the function of many genes remains unsolved [15].

Currently, approximately 8,500 sponge species are described and distributed across four classes: calcareous sponges (Calcarea), glass sponges (Hexactinellida), demosponges (Demospongia) and Homoscleromorpha. As sessile filter feeders, sponges pump enormous amounts of water through their bodies (thousands of liters per kilogram of sponge per day) sequestering food particles and nutrients from the water column [1, 16]. Furthermore, many sponges host diverse and abundant microbial communities within their tissue [17].

However, to what extent sponges and their microbial community interact, what functional role microorganisms play, and whether these interactions display a real symbiosis is as yet poorly understood.

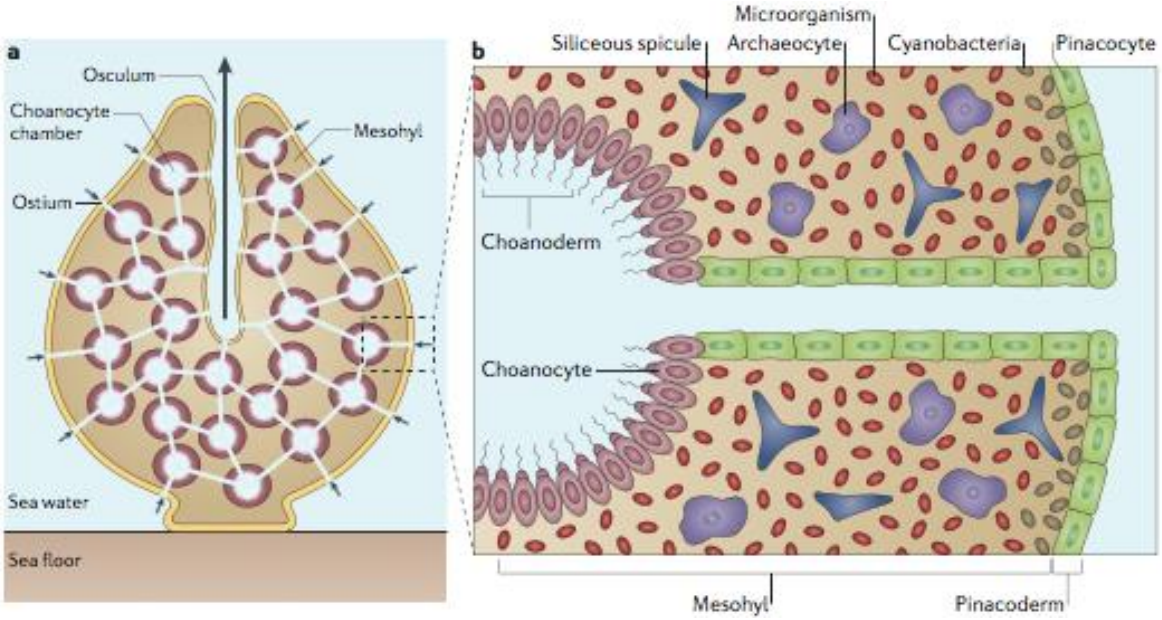

Figure 1. Body plan of a sponge. **(a)** Schematic representation of a typical demosponge. **(b)** An enlargement of the internal structures of a typical demosponge. The sponge body contains pores, canals and chambers forming the aquiferous system. Seawater is inhaled through openings (ostia) in the pinacoderm and exhaled (almost sterile [18, 19]) through the osculum. The water flow is created by the beating of flagellated choanocyte cells that are organized in chambers. Choanocytes also filter nutrients and food particles from the seawater including bacteria [16, 18], unicellular algae [20] and viruses [21] and transfer them to the inner mesohyl of the sponge where they are engulfed by archaeocyte cells. The mesohyl is inhabited by dense and diverse microbial communities in many demosponges. Siliceous spicules provide structural support to most demosponges. Source: Reprinted from [2]. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Microbiology, advanced online publication 30 July 2012 (doi: 10.1038/nrmicro2839

1.2 Living together: the sponge holobiont

1.2.1 Symbiosis

The term symbiosis, originated from the Greek words *syn* "together" and *bíos* "living", was defined about 130 years ago by Anton de Bary as "the living together of unlike organisms" [22]. This original definition encompasses mutualism (both members benefit), commensalism (one member benefits while the other remains unaffected) and parasitism (one member benefits whilst the other is harmed) as symbiotic interactions. Over time the definition of symbiosis has varied from a more restricted use including only persistent

mutualism to a broader one encompassing all types of persistent biological interactions [23]. Today, symbiosis is often used to describe a long-term, often beneficial, intimate association between two or more organisms of different species interacting at the genetic, metabolic or behavioral level. In terms of dependency the relationship between partners can be obligate or facultative. Symbionts can be localized intracellularly as endosymbionts or extracellularly as ectosymbionts. They are acquired by their host either horizontally from the environment, vertically transmitted from parent to offspring, or through a combination of both transmission modes [24].

Prokaryotes are widespread across all environments on the earth and have been around since long before the first occurrence of eukaryotic cells. In the origin and evolution of eukaryotic cells, symbiosis played a crucial role [25]. Since the emergence of eukaryotes, interactions between them and prokaryotic microorganisms were probably abundant and ubiquitous and they most likely played an important role in the evolution of species involved in this symbiosis [26, 27]. Contemporary biology increasingly recognizes that symbiosis "is the rule and not the exception in the animal kingdom", [28] superseding the classical concept of individuals [27, 28], towards a 'holobiont' view – an animal host and all its persistent populations of symbionts. The impact of symbiosis is manifold affecting the organism's development, nutrition, specification, reproduction, immunity and defense against natural enemies. One example of a potential ancient partnership between a host and its microbial community is the sponge holobiont.

1.2.2 The sponge microbiome

Sponges evolved when bacteria were the dominant organisms throughout the oceans. Today sponges are known for hosting abundant and diverse microbial communities in their tissues [17] and although it is unknown how this association was established and if it already existed in ancient sponges, [2, 17] it is very likely that sponges and microorganisms have been interacting since the first appearance of sponges [2]. The evolutionary success of these animals, the ability to inhabit diverse ecological niches and to survive even under unfavorable conditions is thought to be influenced by their intimate association with their microbial symbionts [17]. These sponge-associated microorganisms can constitute up to 50 % of the sponge's total biomass [29], comprising highly diverse symbiotic bacteria, archaea (mainly thaumarcheota) and unicellular eukaryotes (such as diatoms and dinoflagellates) [2, 17], which contribute to many aspects of the sponges'

physiology and ecology [30]. For these reasons sponges are now described as 'holobionts' [30]. However, some sponge species are notably devoid of microorganisms, resulting in the differentiation of two general categories: "high microbial abundance" (HMA) and "low microbial abundance" (LMA) sponges [31, 32]. Within the tissues of HMA sponges microbial organisms can have densities up to several magnitudes greater than found in the surrounding seawater [17] whereas densities of microbial organism in LMA sponges and the surrounding seawater are roughly equivalent [33].

1.2.3. Community composition and diversity

Cultivation experiments and/or molecular approaches over the last decade have helped to decipher the phylogenetic composition and diversity of microbial communities of many sponges from many different habitats, revealing a striking microbial species richness [2, 17]. Representatives from more than 40 microbial phyla and candidate phyla have been identified. This highlights the complexity of sponge-microorganism associations [2, 17, 34, 35]. The major microbial players within sponge-microbe associations include Proteobacteria (especially the classes Alpha-, Gamma- and Deltaproteobacteria), Chloroflexi, Actinobacter, Acidobacter, Nitrospirae and the candidate phylum Poribacteria [2]. Even though sponges are surrounded by a plethora of microorganisms, and due to their filter-feeding lifestyle ingest these, many sponge–associated microorganisms are specific to their host species [36, 37]. Moreover, most microbial communities within sponge species are remarkably stable across different environmental conditions and geographic locations [17, 35], further highlighting the specificity of their association.

Maintaining the microbial community in sponges over different generations is thought to be achieved through two processes: (i) microbial partners are passed maternally (vertical transmission) from one generation to another or/and (ii) microbes are recruited from the surrounding seawater by the sponges' filter feeding activity (horizontal transmission) [2, 17, 38]. The majority of research has focused on the vertical transmission mode that has been shown to occur in numerous sponge species [39, 40]. Deep sequencing has revealed that many sponge-specific microorganisms may occur in low abundance in seawater [41]. The horizontal transmission mode therefor could provide an explanation for how distantly related species from geographically isolated regions manage to acquire shared microbiomes [41].

1.2.4 The functional role of sponge symbionts

While we do have a considerable understanding on how microbial communities in sponges are composed, very little is known about the functional role of sponge-associated bacteria and physiological interactions between sponges and microbes. Due to the habitat and lifestyle of sponges they display the full range of symbiotic interactions from parasitism to mutualism [17]. It is supposed that they benefit from the diverse metabolic activities of their associated microorganisms. Bacterial symbionts can, for example, provide supplementary nutrition [42, 43], and can remove metabolic waste products such as ammonia, nitrite and nitrate by producing bioactive secondary metabolites [44–46], which in turn, can be used by the sponge as chemical defenses [47]. In return it is thought that bacterial symbionts benefit from the close relationship by having a stable environment, nutrition supply and access to ammonia [2, 17]. However, deciphering the functional roles of specific sponge symbionts is a challenging task because the microbial communities associated with sponges are extremely complex. Moreover, many potential symbionts are not yet cultivable in the laboratory and methods to study complex sponge-microbe interactions are limited.

1.2.5 Features of sponge-microorganism symbiosis

Sponge-associated bacteria can occur intracellular within specialized cells called bacteriocytes, but most of them inhabit the mesohyl, an extracellular matrix comprising most of the sponge's body [17]. Interestingly, the mesohyl not only represents the home of extensive sponge-associated microbial communities, but is also the site where digestion of food bacteria is taking place [17]. Feeding studies have shown that sponges are capable discriminating between food bacteria and their own microbial community [48]. How this discrimination works is still a topic of current research, but first results imply that sponges are either able to recognize their symbionts and/or sponge-associated microorganisms have a mechanism to avoid digestion by the sponge, for example by producing slime capsules [48]. One potential molecular mechanism of sponge-symbiont interactions has been revealed by recent metagenomic analyses: sponge-associated bacteria contain an abundance of eukaryotic-like proteins (ELPs) [42, 49–52]. These ELPs seem to be enriched in sponge symbionts compared to bacteria from the surrounding seawater [53]. The domains in ELPs are thought to mediate protein-protein interactions in many biological processes, including those involved in establishing an intracellular lifestyle,

indicating that symbionts might use ELPs to interact with their host [51]. Also, the surprisingly well-developed innate immune system of sponges [54] is likely to play an important role in distinguishing pathogenic from symbiotic microbes through pattern recognition receptors (PRR) that are able to recognize microbial- or pathogen-associated molecular patterns [10, 55, 56].

Sponge-microbe interactions are incredible complex and many questions still remain: How are the sponge-microbe relationships established and maintained? How does a sponge holobiont manage to control the size of its symbiotic population? And to what extant are these relationships a genuine symbiosis? Sponge-microbe symbioses are likely to be one of the most ancient within the animal kingdom. Thus, studying these symbioses will give valuable insights into our understanding of metazoan-microbiota symbiosis and will help to understand how these relationships can evolve.

1.3 Biomineralization: how to build a skeleton

Sponges occur in many sizes and shapes ranging from millimeter thin encrusting species to giant sponges encompassing a few meters [57]. The skeleton of most sponges is composed of two components: the organic skeleton which consist of collagen filaments or spongin fibers, and the inorganic skeleton composed of either siliceous or calcareous spicules [58]. Exceptions to this general bauplan are for example most members of Keratose sponges that have a well-developed organic-fiber skeleton but no inorganic skeleton [59], or coralline sponges, unique members of the porifera, that are characterized by possessing an uncommon solid calcareous skeleton in addition to a spicular one [60].

The ability of organisms to produce mineral structures like the skeleton of sponges is called biomineralization [61]. It is an extremely widespread phenomenon occurring in members of all five kingdoms (prokaryotes, plants, fungi, protists and animals) [61]. Biologically produced minerals are often different in size and shape from their inorganic counterpart due to an exquisite control over the mineralization process [61]. Organisms have gained the ability to secrete an amazing variety of more than 60 different mineral types with calcium carbonate, calcium phosphate and silica as the principal skeletal mineral classes [61]. Mineralized architectures can range from small-scale structures such as magnetite nanocrystals in bacteria to large-scale structures such as reefs systems built from coral and sponge exoskeletons. The function of biomineralized structures are diverse

including structural support, protection, motion, grinding and cutting, storage and optical, magnetic and gravity sensing [62]. Since its invention biomineralization has shaped the life and the environment on earth [61, 63] by influencing not only biogeochemical cycles at a global scale [64], but also the evolutionary history of the organisms producing biominerals [62].

A milestone in metazoan biomineralization evolution was the Cambrian Explosion (543+ mya), which not only marked the radiation of most animal lineages, but also the first appearance of skeletal parts in many metazoan groups [61]. The widespread occurrence of this ability in the fossil record coincides with a rapid increase in metazoan morphology diversity, suggesting that the evolution of biomineralization was one key factor that supported the Cambrian Explosion [62]. The skeletal parts of these early animals fulfilled functions that can be recognized among extant animals. Within the first 20 million years of the Cambrian Explosion, more than half of the recognized 178 distinct marine animal skeletal architectures known today had appeared [65].

Biomineralization processes giving rise to this chemical and structural diversity are generally divided into two fundamentally different groups, depending on the degree of control organisms have during the mineralization process - biologically induced and biologically controlled mineralization [63, 66].

1.3.1 Biologically induced mineralization

Biologically induced mineralization results from the interaction between biological activity and the environment [67]. Organisms have little control over the mineralization process and the type and habit of the mineral deposited - the precipitated mineral is often considered as a side-product of the organism's metabolic activities [67]. Most macroscopic calcium carbonate structures produced by the bacterial world such as stromatolites, microbialites and microbial mud mounds are products of biologically induced mineralization [68–70]. Additionally, many forms of pathological mineralization such as urinary or pancreatic stones fall within that category [61, 71].

Mineral nucleation and growth often occurs directly at the cell surface or on extracellular polymeric substances (EPS) of microbes [67, 72]. The resulting minerals exhibit morphologies similar to their abiotic, inorganically formed counterparts, which are formed

in the absence of a living system, and are highly dependent on the (local) environmental conditions in which they are formed [63, 67].

1.3.2 Biologically controlled mineralization

Unlike biologically induced mineralization, biologically controlled mineralization is an active process where organisms use specialized pathways to synthesize minerals [61]. These pathways are controlled by a cascade of genes, which in the end determine the nucleation, growth, type, crystallographic axes, microstructure and location of the formed mineral [73]. The end products of this biologically controlled mineralization are morphological diverse and often highly sophisticated skeletal structures imbuing organisms with specialized biological functions [73].

One of the most basic requirements for biologically controlled biomineralization is a delimited space that separates the mineralization site from the ambient medium to provide a local environment that can be supersaturated with inorganic precursor ions (e.g. calcium and bicarbonate to form calcium carbonate) [61]. This compartment, such as vesicles or organic membranes, can be located extra-, inter- or intracellularly, and must be able to control ion flows and the composition of the mother liquor [61]. Another requirement is an organic matrix composed of secreted macromolecules including polysaccharides and/or proteins, which serves as a molecular template and mediates nucleation and growth of the mineral [61].

1.3.3 The skeletal matrix

A shared feature of all biologically controlled biominerals is their close association with proteins, polysaccharides, glycoproteins and other macromolecules forming the skeletal matrix [61]. These organic compounds are occluded in the skeleton during the formation process [61]. The skeletal matrix only makes up between 1 - 4 % of the biomineral [74], but it is this organic/inorganic association that gives the biologically produced material impressive mechanical properties that are still unmatched by engineering technologies [75]. The inorganic minerals provide strength, whereas the organic macromolecules provide the ductility, making biominerals much stronger than their inorganic counterparts [76].

Moreover, it is believed that the skeletal matrix plays an important role in controlling nucleation, crystal growth [77], the morphology of the crystallites, and their spatial organization in well-defined microstructures [74]. Recent high-throughput techniques such as proteomics and transcriptomics have expanded our knowledge on the skeletal matrix by producing partial or complete skeletal matrix protein repertoires of many different metazoans including corals [78], brachiopods [79, 80], mollusks [81, 82] and sea urchins [83, 84], resulting in a striking diversity of skeletal proteins [85]. However, no such transcriptomic or proteomic surveys have been conducted on the skeletal matrix of sponges. Up to now, only a few individual biomineralization associated proteins have been identified in sponges [86–89].

1.3.4. Evolutionary aspects of biomineralization

Sponges are informative animals to decipher the origin and evolution of biomineralization in metazoans. Despite the ongoing discussions of their exact branching position in the metazoan tree [7], sponges diverged early from the rest of the metazoans and are among the first animals that displayed a biologically controlled mode of biomineralization [62]. One of the key questions in the evolution of biomineralization is whether the ability to build a mineralized skeleton is inherited from a common ancestor, or whether it is a result of convergent evolution, and thus different animal lineages evolved their mineralized skeletons independently. It is crucial to include early branching metazoans such as sponges in such investigations. Biomineralization mechanisms common to both sponges and all other animals have most probably been present in the last common ancestor of all metazoans (LCAM). It has been estimated that the ability to secrete a calcium carbonate skeleton originated within the metazoan at least 20 times independently [62]. This estimate however, is based on morphological homology of skeletal parts and a simplistic approach of gain or loss of the biomineralization ability, neglecting the underlying molecular mechanisms that produce skeletons. Recent technical advances have helped overcome this issue by facilitating the generation of entire skeletal proteomes. Comparison of skeleton-forming repertoires of several different organisms reveal an astonishing diversity of skeletal proteins, suggesting that different organisms use very different skeletal repertoires to build their skeletons [85]. On the other hand, many taxa utilize deeply conserved biomineral forming components e.g. carbonic anhydrase and peroxidase [90], hinting at a shared ancestral ‘biomineralization toolkit’. Whether these deeply conserved components

were present in the LCAM or whether they were co-opted to play a role in biomineralization remains unclear.

Another interesting evolutionary aspect of sponge biomineralization is the possible role of the diverse and abundant sponge-associated microorganisms. It has recently been reported that Spherulin, a gene horizontally transferred from a bacterium to the genome of the coralline demosponge *Astrosclera willeyana*, most probably supported the evolution of the biomineralization strategy of that sponge [88]. Spherulin is also present in the genome of the non-calcifying demosponge *Amphimedon queenslandica* but absent from other eukaryotic genomes, showing that this gene was not provided by the LCAM but rather was acquired horizontally by the common ancestor of *A. willeyana* and *A. queenslandica* [88]. In addition to the important role of Spherulin in the biomineralization process of *A. willeyana*, the sponge degrades a portion of its microbial community via the autophagy pathway [91] and uses the organic remains of these microbes to seed crystal growth [92]. Furthermore, the sponge genus *Hemimycale* harbors endosymbiotic calcibacteria that are directly responsible for the precipitation of the calcium carbonate skeleton [93]. These examples show the potential role of microorganisms in the evolution of biomineralization strategies in basal metazoans such as sponges. However, the availability of data for different sponge biomineralization strategies is too small to say whether a role of sponge-associated microorganisms in biomineralization is a common trait or an exception in sponges.

1.3.5 Sponges and their history of biomineralization

Out of the first multicellular eukaryotes, sponges were the pioneers in the development of a "biological controlled" mode of biomineralization [62]. Throughout earth history, calcified (or coralline) sponges were important reef building organisms, beginning with the appearance of the archaeocyatha at the base of the Tommotian age (~ 530 mya) [94]. To date most authorities now agree that the archaeocyatha were an extinct class of sponges with affinities to demosponges [94]. Archaeocyatha are considered as the planet's first metazoan reef-builders, and have been incredibly successful from their first occurrence over the next 10 million years to come; archaeocyatha were globally distributed, ecologically important, diversified in hundreds of recognized species and left evidence in the fossil record of a mode of heavily calcification that is poorly represented among living sponges [95, 96]. As successful as their diversification was, the archaeocyatha were a

short-lived group and 10 to 15 million years after their first appearance these reef-builders disappeared completely from the fossil record [94]. Subsequently, other groups of coralline sponges, sphinctozoans, stromatoporoids and chaetetids contributed greatly in constructing reefs during long periods of earth history [97]. Coralline sponges were thought to be extinct until their rediscovery in the late 1960s [98]. However, they have survived in cryptic niches and their descendants can be found in these habitats in almost all modern coral reefs [29, 97]. Formerly, coralline sponges were grouped into the class of 'Sclerospongia' based on their calcareous skeleton arrangement [99], but were recently identified as being polyphyletic sharing solely a functional character rather than common ancestry [60, 100].

1.4 *Vaceletia* – an ancient hypercalcifying demosponge

Vaceletia, a coralline sponge that is characterized by a hypercalcified aragonitic skeleton, is a monospecific sponge genus with a single described species, *Vaceletia crypta* [101]. It was thought to be extinct until its rediscovery in the late 1970s [101]. Molecular analysis show that systematically *Vaceletia* belongs to the Keratosa (with highest affinities to the order Dictyoceratida). [59, 102, 103]. *Vaceletia* occurs in two different growth forms: solitary and colonial, which most likely represent different species [59]. However, the taxonomic status of *Vaceletia* is not yet resolved [59]. It first appeared in the middle Triassic [29] and has a rich Mesozoic fossil record (reviewed in [104]). Today, *Vaceletia* inhabits cryptic niches such as caves or deeper fore reef areas that are light reduced and have oligotrophic conditions [60, 105]; it has been reported from depths ranging from 10 to 530 m [106].

1.4.1 The microbiome of *Vaceletia*

Vaceletia harbors a dense and diverse microbial community (Fig. 2), which can constitute up to 50 % of the sponge's biomass [29]. The microbial community of the solitary *V. crypta* has been previously studied using 16S rRNA clone libraries and DGGE [107]. This analysis revealed that the microbial community of *V. crypta* is highly diverse and shares features with other sponge derived microbial communities. DGGE cluster analysis of solitary and colonial *Vaceletia* species indicate that the different growth forms harbor distinct microbial communities [107].

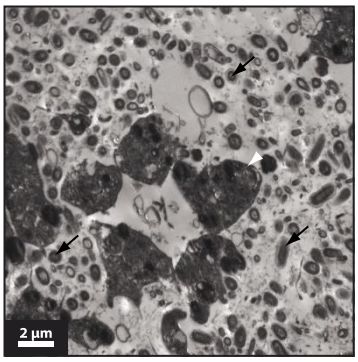

Figure 2. TEM image of *Vaceletia* sp. mesohyl (taken from [108]). The mesohyl is filled with darkly stained sponge cells (white arrow-heads) and abundant and diverse microorganism (black arrows).

1.4.2 Skeleton formation of *Vaceletia* – what is known

Vaceletia is the only living representative that builds its skeleton in a spinctozoan-like way, which is denoted by a chambered structure. This architecture was characteristic for the now extinct sphinctozoan sponges. Superficial morphological similarities to the extinct archaeocyatha led to the hypothesis that *Vaceletia* might represent a modern Archaeocyath [29, 109]. Since molecular data has shown that *Vaceletia* belongs to Dictyoceratida [59, 102, 103], this is rather unlikely. Nevertheless, *Vaceletia* represents an early branching metazoan with a possible ancient mode of biomineralization.

Vaceletia possesses an elaborate skeleton and when viewed from the nm to the cm scales, exhibits exquisite biological control over the formation of its aragonitic $CaCO_3$ skeleton. The overall structure is comprised by chambers terraced one upon another (Fig. 3). Some features of the skeleton formation have been described in detail [29, 110]. Briefly, the sponge grows by first constructing an organic framework, which then is successively mineralized by crystalline aragonite. Hence, the ontogenetically youngest chambers are on top of the animal. In ontogenetically older chambers living tissue is moving upward and chambers are filled with soluble acidic glycoproteins. These chambers are then subsequently mineralized by crystalline aragonite, forming the hypercalcified stalk of the animal (Fig. 3 D-F). The organic matrix consists of proteins and sugars rich in galactose, glucose and fructose, the latter suggesting that bacterial exopolymeric substances (EPS) may be involved in the biocalcification process. Due to the high abundance and diversity of microbes in *Vaceletia* sp., it seems likely that interactions between the sponge and its microbial community play a role in constructing its skeleton. Furthermore, it has recently been shown that sponges employ their microbial communities in the biomineralization process [88, 91–93].

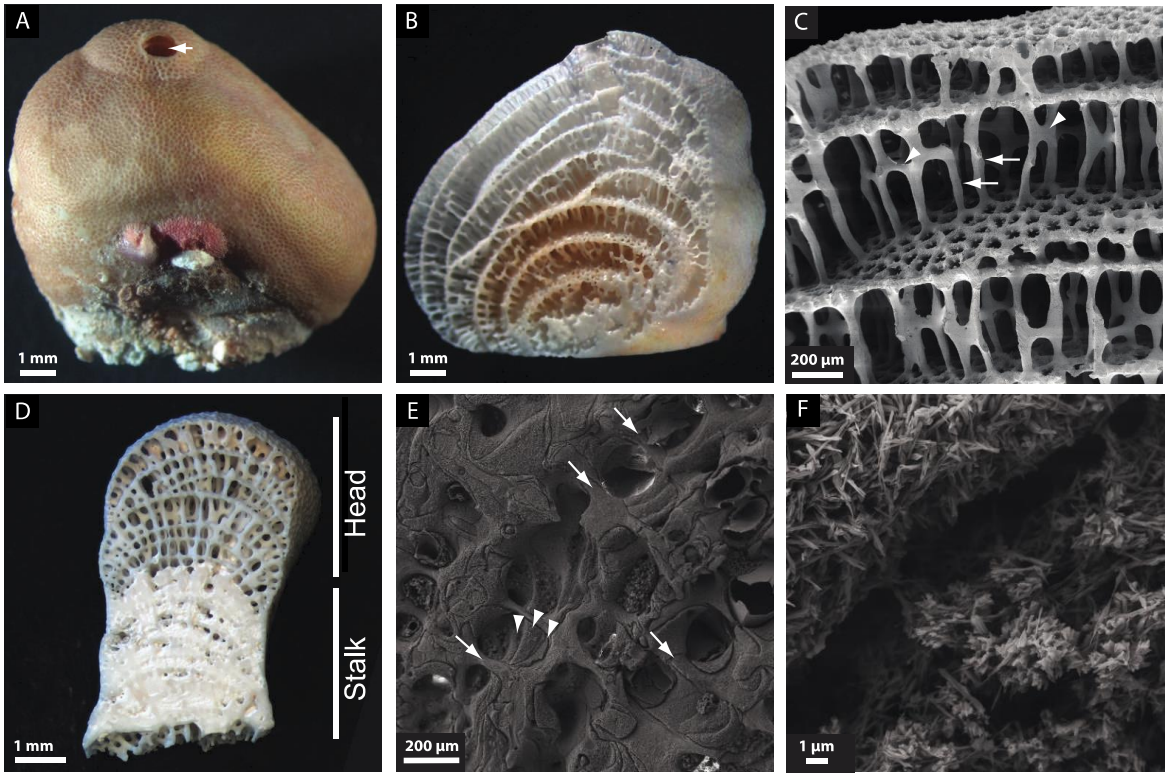

Figure 3. General morphological features of *Vaceletia* sp. and its aragonitic calcium carbonate skeleton (adapted from [108]) **(A)** A lateral view of a fixed animal. The exhalent osculum (arrow) is clearly visible. **(B)** A sagittal section view of an animal after treatment with NaOCl and grinding reveals the structure of the animal. **(C)** An SEM image of the upper part of the skeleton. Pillars (arrow) support the chambers and are reinforced by radial spines (arrowheads). **(D)** A sagittal section after NaOCl treatment and grinding shows the head and hypercalcified stalk region. **(E)** An SEM image of the stalk after etching with EDTA. The pillars of the skeleton are still visible (arrows). Chambers are mineralized in layers (arrowheads). Note that not all chambers are mineralized entirely. **(F)** Both pillars and mineralized chambers are constructed by needles of aragonite.

1.5 Aim of this thesis

Despite the previous work described earlier, nothing is known about the highly coordinated molecular basis of the biocalcification process of this phylogenetically informative taxon. In fact, very little is known about biomineralization and the underlying molecular mechanisms in sponges in general. The main aim of this thesis was to identify and characterize the molecular mechanisms of biocalcification employed by the as yet to be described colonial-branching *Vaceletia* sp. species. Due to the high abundance and diversity of *Vaceletia*'s microbial community it seems reasonable that interactions between the sponge and its microbes may also play some role in fabricating its calcified skeleton. This thesis aims to find the first answers for the following key questions:

- How does *Vaceletia* sp. interact with its microbial community?
- Which proteins are involved in biocalcification process of *Vaceletia* sp.?
- Does the microbial community of *Vaceletia* sp. contribute to the biocalcification process, and if so how?
- Which genetic repertoire was present in the last common ancestor of all Metazoa that possibly contributed to the evolution of the ability to biocalcify?
- Where are the skeleton-forming genes expressed and what is their potential function?

As a first step to address these questions an Illumina transcriptome dataset from *Vaceletia* sp. was generated. *Vaceletia* sp. harbors an abundant and diverse microbial community [107], however, so far little is known about the function of these potential symbionts and how the sponge manages its microbial community. To expand our knowledge on sponge-microbe interactions the holo-transcriptome of *Vaceletia* sp. was analyzed and characterized and compared to other sponge transcriptomic and genomic data in **Chapter 2**. An *in silico* approach was used to characterize the underlying molecular mechanisms of different lipid pathway components and to study potential metabolic interactions between the sponge and its microbial community.

In the next step, proteins from purified skeletal elements of *Vaceletia* sp. were extracted. This proteomic data was combined with the transcriptomic dataset to generate a skeletal proteome of the head and stalk region of this coralline demosponge. This proteome represents the first comprehensive biomineralization dataset from a sponge and provides information about the components *Vaceletia* sp. employs in constructing its skeleton (**Chapter 3**). It also gives a first clue to a potential role of bacteria in the biomineralization process. Skeletogenic proteome surveys will expand our knowledge of the protein repertoires animal uses in the biomineralization process and will help us to understand how this ability may have evolved.

With the list of biocalcification gene candidates generated in Chapter 3 it is now possible to further characterize these genes and visualize their spatial expression profiles by using the technique of *in situ* hybridization (**Chapter 4**). Gene candidates expressed in sponge or bacterial cells that are intimately associated with the sites of active biomineralization are

likely to be involved in this process. Knowing the location of these genes helps in inferring their potential function.

References

1. Bell, J. J. (2008) **The functional roles of marine sponges.** *Estuar*. *Coast*. *Shelf*. *Sci*., 79(3), 341-353, doi:10.1016/j.ecss.2008.05.002.

2. Hentschel, U., Piel, J., Degnan, S. M., Taylor, M. W. (2012) **Genomic insights into the marine sponge microbiome.** *Nat*. *Rev*. *Microbiol*., 10(9), 641-654, doi:10.1038/nrmicro2839.

3. McClintock, J. B., Amsler, C. D., Baker, B. J., Van Soest, R. W. M. (2005) **Ecology of Antarctic Marine Sponges: An Overview.** *Integr*. *Comp*. *Biol*., 45(2), 359-368, doi:d10.1093/icb/45.2.359.

4. Dohrmann, M., Wörheide, G. (2013) **Novel Scenarios of Early Animal Evolution - Is It Time to Rewrite Textbooks?** *Integr*. *Comp*. *Biol*., 53(3), 503-511, doi:10.1093/icb/ict008.

5. Wörheide, G., Dohrmann, M., Erpenbeck, D., Larroux, C., Maldonado, M., Voigt, O. et al. (2012) **Deep phylogeny and evolution of sponges (phylum Porifera).** *Adv*. *Mar*. *Biol*., 61, 1-78, doi:10.1016/B978-0-12-387787-1.00007-6.

6. Sperling, E. A., Robinson, J. M., Pisani, D., Peterson, K. J. (2010) **Where's the glass? Biomarkers, molecular clocks, and microRNAs suggest a 200-Myr missing Precambrian fossil record of siliceous sponge spicules.** *Geobiology*, 8(1), 24-36, doi:10.1111/j.1472-4669.2009.00225.x.

7. Nosenko, T., Schreiber, F., Adamska, M., Adamski, M., Eitel, M., Hammel, J. et al. (2013) **Deep metazoan phylogeny: when different genes tell different stories.** *Mol*. *Phylogenet*. *Evol*., 67(1), 223-233, doi:10.1016/j.ympev.2013.01.010.

8. Philippe, H., Derelle, R., Lopez, P., Pick, K., Borchiellini, C., Boury-Esnault, N. et al. (2009) **Phylogenomics Revives Traditional Views on Deep Animal Relationships.** *Curr*. *Biol*., 19(8), 706-712, doi:10.1016/j.cub.2009.02.052.

9. Pick, K. S., Philippe, H., Schreiber, F., Erpenbeck, D., Jackson, D. J., Wrede, P. et al. (2010) **Improved phylogenomic taxon sampling noticeably affects nonbilaterian relationships.** *Mol*. *Biol*. *Evol*., 27(9), 1983-1987, doi:10.1093/molbev/msq089.

10. Srivastava, M., Simakov, O., Chapman, J., Fahey, B., Gauthier, M. E. A., Mitros, T. et al. (2010) **The *Amphimedon queenslandica* genome and the evolution of animal complexity.** *Nature*, 466(7307), 720-726, doi:10.1038/nature09201.

11. Sperling, E. A., Peterson, K. J., Pisani, D. (2009) **Phylogenetic-Signal Dissection of Nuclear Housekeeping Genes Supports the Paraphyly of Sponges and the Monophyly of Eumetazoa.** *Mol*. *Biol*. *Evol*., 26(10), 2261-2274, doi:10.1093/molbev/msp148.

12. Ryan, J. F., Pang, K., Schnitzler, C. E., Nguyen, A. D., Moreland, R. T., Simmons, D. K. et al. (2013) **The Genome of the Ctenophore *Mnemiopsis leidyi* and Its Implications**

for Cell Type Evolution. *Science*, 342(6164), 1242592-1242592, doi:10.1126/science.1242592.

13. Moroz, L. L., Kocot, K. M., Citarella, M. R., Dosung, S., Norekian, T. P., Povolotskaya, I. S. et al. (2014) **The ctenophore genome and the evolutionary origins of neural systems.** *Nature*, 510(7503), 109-114, doi:10.1038/nature13400.

14. Nichols, S. A., Roberts, B. W., Richter, D. J., Fairclough, S. R., King, N. (2012) **Origin of metazoan cadherin diversity and the antiquity of the classical cadherin/β-catenin complex.** *PNAS*, (109), 13046-13051, doi:10.1073/pnas.1120685109.

15. Riesgo, A., Farrar, N., Windsor, P. J., Giribet, G., Leys, S. P. (2014) **The analysis of eight transcriptomes from all poriferan classes reveals surprising genetic complexity in sponges.** *Mol. Biol. Evol.*, 31(5), 1102-1120, doi:10.1093/molbev/msu057.

16. Reiswig, H. M. (1975) **Bacteria as food for temperate-water marine sponges.** *Can. J. Zool.*, 53(5), 582-589, doi:10.1139/z75-072.

17. Taylor, M. W., Radax, R., Steger, D., Wagner, M. (2007) **Sponge-Associated microorganisms: evolution, ecology, and biotechnological potential.** *Microbiol. Mol. Biol. Rev.*, 71(2), 295-347, doi:10.1128/MMBR.00040-06.

18. Reiswig, H. M. (1971) **Particle Feeding in Natural Populations of Three Marine Demosponges.** *Biol. Bull.*, 141(3), 591568-591591, doi:10.2307/1540270.

19. Wehrl, M., Steinert, M., Hentschel, U. (2007) **Bacterial Uptake by the Marine Sponge *Aplysina aerophoba*.** *Microb. Ecol.*, 53(2), 355-365, doi:10.1007/s00248-006-9090-4.

20. Pile, A. J., Patterson, M. R., Witman, J. D. (1996) **In situ grazing on plankton.** *Mar. Ecol. Prog. Ser.* 141, 95-102.

21. Hadas, E., Marie, D., Shpigel, M., Ilan, M. (2006) **Virus predation by sponges is a new nutrient-flow pathway in coral reef food webs.** *Limnol. Oceanogr.*, 51(3), 1548-1550, doi:10.4319/lo.2006.51.3.1548.

22. de Bary, A. (1879) **Die Erscheinung der Symbiose: Vortrag gehalten auf der Versammlung Deutscher Naturforscher und Aerzte zu Cassel.** Trübner;

23. Martin, B. D., Schwab, E. (2012) **Current Usage of Symbiosis and Associated Terminology.** *Int. J. Biol.*, 5(1), doi:10.5539/ijb.v5n1p32.

24. Bright, M., Bulgheresi, S. (2010) **A complex journey: transmission of microbial symbionts.** *Nat. Rev. Microbiol.*, 8(3), 218-230, doi:10.1038/nrmicro2262.

25. Sagan, L. (1967) **On the origin of mitosing cells.** *J. Theor. Biol.*, 14(3), 225-IN6, doi:10.1016/0022-5193(67)90079-3.

26. Moya, A., Peretó, J., Gil, R., Latorre, A. (2008) **Learning how to live together: genomic insights into prokaryote-animal symbioses.** *Nat*. *Rev*. *Genet*., 9(3), 218-229, doi:10.1038/nrg2319.

27. 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*. *USA*, 110(9), 3229-3236, doi:10.1073/pnas.1218525110.

28. Gilbert, S. F. (2014) **Symbiosis as the way of eukaryotic life: The dependent co-origination of the body.** *J*. *Biosciences*., 39(2), 201-209, doi:10.1007/s12038-013-9343-6.

29. Reitner, J., Wörheide, G., Lange, R., Thiel, V. (1997) **Biomineralization of calcified skeletons in three Pacific coralline demosponges - an approach to the evolution of basal skeletons.** *Cour*. *Forsch-Inst*. *Senckenberg* 201, 371-383.

30. Webster, N. S., Thomas, T. (2016) **The Sponge Hologenome.** *mBio*, 7(2), e00135-16, doi:10.1128/mBio.00135-16.

31. Reiswig, H. M. (1981) **Partial carbon and energy budgets of the bacteriosponge *Verohgia fistularis* (Porifera: Demospongiae) in Barbados.** *Mar*. *Ecol*. 2(4), 273-293.

32. Hentschel, U., Fieseler, L., Wehrl, M., Gernert, C., Steinert, M., Hacker, J. et al. (2003) **Microbial Diversity of Marine Sponges.** In: Sponges (Porifera), edited by Müller, W. E. G., Berlin, Heidelberg: Springer Berlin Heidelberg; 59-88.

33. Hentschel, U., Usher, K. M., Taylor, M. W. (2006) **Marine sponges as microbial fermenters.** *FEMS Microbiol*. *Ecol*., 55(2), 167-177, doi:10.1111/j.1574-6941.2005.00046.x.

34. Schmitt, S., Tsai, P., Bell, J., Fromont, J., Ilan, M., Lindquist, N. et al. (2011) **Assessing the complex sponge microbiota: core, variable and species-specific bacterial communities in marine sponges.** *ISME J*., 6(3), 564-576, doi:10.1038/ismej.2011.116.

35. Reveillaud, J., Maignien, L. I. S., Eren, A. M., Huber, J. A., Apprill, A., Sogin, M. L. et al. (2014) **Host-specificity among abundant and rare taxa in the sponge microbiome.** *ISME J*., 8(6), 1198-1209, doi:10.1038/ismej.2013.227.

36. Hentschel, U., Hopke, J., Horn, M., Friedrich, A. B., Wagner, M., Hacker, J. et al. (2002) **Molecular evidence for a uniform microbial community in sponges from different oceans.** *Appl*. *Environ*. *Microbiol*., 68(9), 4431-4440, doi:10.1128/AEM.68.9.4431-4440.2002.

37. Simister, R. L., Deines, P., Botté, E. S., Webster, N. S., Taylor, M. W. (2012) **Sponge-specific clusters revisited: a comprehensive phylogeny of sponge-associated microorganisms.** *Eviron*. *Microbiol*., 14(2), 517-524, doi:10.1111/j.1462-2920.2011.02664.x.

38. Webster, N. S., Taylor, M. W. (2011) **Marine sponges and their microbial symbionts: love and other relationships.** *Eviron*. *Microbiol*., 14(2), 335-346, doi:10.1111/j.1462-2920.2011.02460.x.

39. Sharp, K. H., Eam, B., Faulkner, D. J., Haygood, M. G. (2007) **Vertical transmission of diverse microbes in the tropical sponge *Corticium* sp.** *Appl*. *Environ*. *Microbiol*., 73(2), 622-629, doi:10.1128/AEM.01493-06.

40. Schmitt, S., Weisz, J. B., Lindquist, N., Hentschel, U. (2007) **Vertical Transmission of a Phylogenetically Complex Microbial Consortium in the Viviparous Sponge *Ircinia felix*.** *Appl*. *Environ*. *Microbiol*., 73(7), 2067-2078, doi:10.1128/AEM.01944-06.

41. 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.** *Eviron*. *Microbiol*., 12(8), 2070-2082, doi:10.1111/j.1462-2920.2009.02065.x.

42. 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(12), 1557-1567, doi:10.1038/ismej.2010.74.

43. Fiore, C. L., Labrie, M., Jarett, J. K., Lesser, M. P. (2015) **Transcriptional activity of the giant barrel sponge, *Xestospongia muta* Holobiont: molecular evidence for metabolic interchange.** *Front*. *Microbiol*., 6, 364, doi:10.3389/fmicb.2015.00364.

44. Piel, J., Hui, D., Wen, G., Butzke, D., Platzer, M., Fusetani, N. et al. (2004) **Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge *Theonella swinhoei*.** *Proc*. *Natl*. *Acad*. *Sci*. *USA*, 101(46), 16222-16227, doi:10.1073/pnas.0405976101.

45. Hochmuth, T., Niederkrüger, H., Gernert, C., Siegl, A., Taudien, S., Platzer, M. et al. (2010) **Linking chemical and microbial diversity in marine sponges: possible role for Poribacteria as producers of methyl-branched fatty acids.** *ChemBioChem*, 11(18), 2572-2578, doi:10.1002/cbic.201000510.

46. Wilson, M. C., Mori, T., Rückert, C., Uria, A. R., Helf, M. J., Takada, K. et al. (2014) **An environmental bacterial taxon with a large and distinct metabolic repertoire.** *Nature*, 506(7486), 58-62, doi:10.1038/nature12959.

47. Pawlik, J. R. (2011) **The chemical ecology of sponges on caribbean reefs: natural products shape natural systems.** *BioScience*, 61(11), 888-898, doi:10.1525/bio.2011.61.11.8.

48. Wilkinson, C. R., Garrone, R., Vacelet, J. (1984) **Marine sponges discriminate between food bacteria and bacterial symbionts: electron microscope radioautography and *in situ* evidence.** *Proc*. *R*. *Soc*. *Lond*. *B*. *Biol*. *Sci*., 220(1221), 528-519, doi:10.1098/rspb.1984.0018.

49. Siegl, A., Kamke, J., Hochmuth, T., Piel, J. O. R., Richter, M., Liang, C. et al. (2010) **Single-cell genomics reveals the lifestyle of *Poribacteria*, a candidate phylum symbiotically associated with marine sponges.** *ISME J*., 5(1), 61-70, doi:10.1038/ismej.2010.95.

50. Liu, M., Fan, L., Zhong, L., Kjelleberg, S., Thomas, T. (2012) **Metaproteogenomic analysis of a community of sponge symbionts.** *ISME J*., 6(8), 1515-1525, doi:10.1038/ismej.2012.1.

51. Nguyen, M. T., Liu, M., Thomas, T. (2014) **Ankyrin-repeat proteins from sponge symbionts modulate amoebal phagocytosis.** *Mol Ecol*, 23(6), 1635-1645, doi:10.1111/mec.12384.

52. 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(1), e87353, doi:10.1371/journal.pone.0087353.

53. Fan, L., Reynolds, D., Liu, M., Stark, M., Kjelleberg, S., Webster, N. S. et al. (2012) **Functional equivalence and evolutionary convergence in complex communities of microbial sponge symbionts.** *Proc. Natl. Acad. Sci. USA*, 109(27), E1878-E1887, doi:10.1073/pnas.1203287109.

54. Müller, W., E. G., Müller, I. M. (2003) **Origin of the metazoan immune system: identification of the molecules and their functions in sponges.** *Integr. Comp. Biol*., 43(2), 292-281, doi:10.1093/icb/43.2.281.

55. Gauthier, M. E. A., Du Pasquier, L., Degnan, B. M. (2010) **The genome of the sponge *Amphimedon queenslandica* provides new perspectives into the origin of Toll-like and interleukin 1 receptor pathways.** *Evol. Dev*., 12(5), 519-533, doi:10.1111/j.1525-142X.2010.00436.x.

56. Yuen, B., Bayes, J. M., Degnan, S. M. (2014) **The characterization of sponge NLRs provides insight into the origin and evolution of this innate immune gene family in animals.** *Mol. Biol. Evol*., 31(1), 106-120, doi:10.1093/molbev/mst174.

57. Hooper, J. N. A., Van Soest, R. W. M. (2002) **Systema Porifera. A Guide to the Classification of Sponges.** In: *Systema Porifera*, edited by Hooper, J. N. A., Van Soest, R. W. M., Willenz, P., Boston, MA: Springer US; 1-7.

58. Bergquist, P. R. (1979) **Sponges.** Univ of California Pr; 1-268.

59. Wörheide, G. (2008) **A hypercalcified sponge with soft relatives: *Vaceletia* is a keratose demosponge.** *Mol. Phylogenet. Evol*., 47(1), 433-438, doi:10.1016/j.ympev.2008.01.021.

60. Reitner, J. (1992) **Coralline Spongien: der Versuch einer phylogenetisch-taxonomischen Analyse.** *Berliner Geowisschenschaftliche Abhandlungen* (Reihe E) 1, 1-352.

61. Lowenstam, H. A., Weiner, S. (1989) **On Biomineralization.** Oxford University Press; 1-324.

62. Knoll, A. H. (2003) **Biomineralization and evolutionary history.** *Rev*. *Mineral*. *Geochem*., 54(1), 329-356, doi:10.2113/0540329.

63. Lowenstam, H. (1981) **Minerals formed by organisms.** *Science*, 211(4487), 1126-1131, doi:10.1126/science.7008198.

64. Van Cappellen, P. (2003) **Biomineralization and Global Biogeochemical Cycles.** *Rev*. *Mineral*. *Geochem*., 54(1), 357-381, doi:10.2113/0540357.

65. Thomas, R. D., Shearman, R. M., Stewart, G. W. (2000) **Evolutionary exploitation of design options by the first animals with hard skeletons.** *Science*, 288(5469), 1239-1242, doi:10.1126/science.288.5469.1239.

66. Mann, S. (1983) **Mineralization in biological systems.** In: *Inorganic Elements in Biochemistry*, edited by Connett, P. H., Folmann, H., Lammers, M., Mann, S., Odom, J. D., Wetterhahn, K. E., Berlin, Heidelberg: Springer Berlin Heidelberg; 125-174.

67. Weiner, S., Dove, P. M. (2003) **An Overview of Biomineralization Processes and the Problem of the Vital Effect.** *Rev*. *Mineral*. *Geochem*., 54(1), 1-29, doi:10.2113/0540001.

68. Reitner, J., Gautret, P., Marin, F., Neuweiler, F. (1995) **Automicrites in modern marine microbialite. Formation model via organic matrices (Lizard Island, Great Barrier Reef, Australia).** *Bulletin de l'Institut Océanographique (Monaco)* Numéro Spécial 14, 237 - 264.

69. Gautret, P., Camoin, G., Golubic, S., Sprachta, S. (2004) **Biochemical control of calcium carbonate precipitation in modern lagoonal microbialites, Tikehau Atoll, French Polynesia.** *J*. *Sediment*. *Res*., 74(4), 462-478, doi:10.1306/012304740462.

70. Monty, C. L. V., Bosence, D. W. J., Bridges, P. H., Pratt, B. R. (2009) **Carbonate mud-mounds: Their origin and evolution (Special Publication 23 of the IAS).** John Wiley & Sons;

71. Kontoyannis, C. G., Vagenas, N. V. (2000) **Calcium carbonate phase analysis using XRD and FT-Raman spectroscopy.** *Analyst*, 125(2), 251-255, doi:10.1039/a908609i.

72. Dupraz, C., Reid, R. P., Braissant, O., Decho, A. W., Norman, R. S., Visscher, P. T. (2009) **Processes of carbonate precipitation in modern microbial mats.** *Earth Sci*. *Rev*., 96(3), 141-162, doi:10.1016/j.earscirev.2008.10.005.

73. Dove, P. M. (2010) **The Rise of Skeletal Biominerals.** *Elements*, 6(1), 37-42, doi:10.2113/gselements.6.1.37.

74. Marin, F., Le Roy, N., Marie, B., Ramos-Silva, P., Bundeleva, I., Guichard, N. et al. (2014) **Metazoan calcium carbonate biomineralizations: macroevolutionary trends - challenges for the coming decade.** *B. Soc. Geol. Fr.*, 185(4), 217, doi:10.2113/gssgfbull.185.4.217.

75. Meyers, M. A., Chen, P.-Y., Lin, A. Y.-M., Seki, Y. (2008) **Biological materials: Structure and mechanical properties.** *Prog. Mater. Sci.*, 53(1), 1-206, doi:10.1016/j.pmatsci.2007.05.002.

76. Addadi, L., Weiner, S. (2014) **Biomineralization: mineral formation by organisms.** *Phys. Scr.*, 89(9), 098003, doi:10.1088/0031-8949/89/9/098003.

77. Weiner, S. (1984) **Organization of Organic Matrix Components in Mineralized Tissues.** *American Zoologist*, 24(4), 945, doi:10.1093/icb/24.4.945.

78. Drake, J. L., Mass, T., Haramaty, L., Zelzion, E., Bhattacharya, D., Falkowski, P. G. (2013) **Proteomic analysis of skeletal organic matrix from the stony coral *Stylophora pistillata*.** *Proc Natl Acad Sci U S A*, 110(10), 3788-3793, doi:10.1073/pnas.1301419110.

79. Immel, F., Gaspard, D., Marie, A., Guichard, N., Cusack, M., Marin, F. (2015) **Shell proteome of rhynchonelliform brachiopods.** *J. Struct. Biol.*, 190(3), 360-366, doi:10.1016/j.jsb.2015.04.001.

80. Jackson, D. J., Mann, K., Häussermann, V., Schilhabel, M., Lüter, C., Griesshaber, E. et al. (2015) **The *Magellania venosa* biomineralizing proteome: a window into brachiopod shell evolution.** *Genome Biol. Evol.*, doi:10.1093/gbe/evv074.

81. Jackson, D. J., Mcdougall, C., Green, K., Simpson, F., Wörheide, G., Degnan, B. M. (2006) **A rapidly evolving secretome builds and patterns a sea shell.** *BMC Biol*, 4, 40, doi:10.1186/1741-7007-4-40.

82. Marie, B., Marie, A., Jackson, D. J., Dubost, L., Degnan, B. M., Milet, C. et al. (2010) **Proteomic analysis of the organic matrix of the abalone *Haliotis asinina* calcified shell.** *Proteome Sci.*, 8, 54, doi:10.1186/1477-5956-8-54.

83. Mann, K., Poustka, A. J., Mann, M. (2008) **The sea urchin (*Strongylocentrotus purpuratus*) test and spine proteomes.** *Proteome Sci.*, 6, 22, doi:10.1186/1477-5956-6-22.

84. Mann, K., Wilt, F. H., Poustka, A. J. (2010) **Proteomic analysis of sea urchin (*Strongylocentrotus purpuratus*) spicule matrix.** *Proteome Sci.*, 8, 33, doi:10.1186/1477-5956-8-33.

85. Marin, F., Bundeleva, I., Takeuchi, T., Immel, F., Medakovic, D. (2016) **Organic matrices in metazoan calcium carbonate skeletons: Composition, functions, evolution.** *J*. *Struct*. *Biol*., 196(2), 98-106, doi:10.1016/j.jsb.2016.04.006.

86. Shimizu, K., Cha, J., Stucky, G. D., Morse, D. E. (1998) **Silicatein α: Cathepsin L-like protein in sponge biosilica.** *Proc*. *Natl*. *Acad*. *Sci*. *USA* 95(11), 6234-6238.

87. Jackson, D. J., Macis, L., Reitner, J., Degnan, B. M., Wörheide, G. (2007) **Sponge paleogenomics reveals an ancient role for carbonic anhydrase in skeletogenesis.** *Science*, 316(5833), 1893-1895, doi:10.1126/science.1141560.

88. Jackson, D. J., Macis, L., Reitner, J., Wörheide, G. (2011) **A horizontal gene transfer supported the evolution of an early metazoan biomineralization strategy.** *BMC Evol*. *Biol*., 11, 238, doi:10.1186/1471-2148-11-238.

89. Voigt, O., Adamski, M., Sluzek, K., Adamska, M. (2014) **Calcareous sponge genomes reveal complex evolution of α-carbonic anhydrases and two key biomineralization enzymes.** *BMC Evol*. *Biol*., 14, 230, doi:10.1186/s12862-014-0230-z.

90. Jackson, D. J., Degnan, B. M. (2016) **The importance of evo-devo to an integrated understanding of molluscan biomineralisation.** *J*. *Struct*. *Biol*., 196(2), 67-74, doi:10.1016/j.jsb.2016.01.005.

91. Jackson, D. J., Wörheide, G. (2014) **Symbiophagy and biomineralization in the "living fossil" *Astrosclera willeyana*.** *Autophagy*, 10(3), 408-415, doi:10.4161/auto.27319.

92. Jackson, D. J., Thiel, V., Wörheide, G. (2010) **An evolutionary fast-track to biocalcification.** *Geobiology*, 8(3), 191-196, doi:10.1111/j.1472-4669.2010.00236.x.

93. Uriz, M. J., Agell, G., Blanquer, A., Turon, X., Casamayor, E. O. (2012) **Endosymbiotic calcifying bacteria: a new cue to the origin of calcification in metazoa?** *Evolution*, 66(10), 2993-2999, doi:10.1111/j.1558-5646.2012.01676.x.

94. Debrenne, F., Zhuravlev, A. Y., Kruse, P. D. (2002) **Class Archaeocyatha Bornemann, 1884.** In: *Systema Porifera*, edited by Hooper, J. N. A., Van Soest, R. W. M., Willenz, P., Boston, MA: Springer US; 1539-1699.

95. Debrenne, F., Reitner, J. (2001) **Sponges, cnidarians, and ctenophores.** In: *The ecology of the Cambrian radiation*, edited by Zhuravlev, A. Y., Riding, R., New York: Columbia University Press; 301-325.

96. Kerner, A., Debrenne, F., Vignes-Lebbe, R. (2011) **Cambrian archaeocyathan metazoans: revision of morphological characters and standardization of genus descriptions to establish an online identification tool.** *Zookeys*, (150), 381-395, doi:10.3897/zookeys.150.1566.

97. Wood, R. (1990) **Reef-building sponges.** *Am*. *Sci*. 78(3), 224-235.

98. Hartman, W. D. (1969) **New genera and species of coralline sponges (Porifera) from Jamaica.** New Haven, Conn.: Peabody Museum of Natural History; 1-39.

99. Hartman, W. D., Goreau, T. F. (1970) **Jamaican coralline sponges: their morphology, ecology and fossil relatives.** *Symp. Zool. Soc. Lond.* 25, 205-243.

100. Vacelet, J. (1985) **Coralline sponges and the evolution of Porifera.** In: *The origins and relationships of lower invertebrates*, edited by Morris, S. C., Georg, J. D., Gibson, R., Platt, H. M., Clarendon Press, Oxford; 1-13.

101. Vacelet, J. (1977) **Une nouvelle relique du Secondaire: un représentant actuel des Eponges fossiles Sphinctozoaires.** *Comptes Rendus De L'Academie Des Sciences Paris* (série D) 285, 509-511.

102. Erpenbeck, D., Voigt, O., Wörheide, G., Lavrov, D. V. (2009) **The mitochondrial genomes of sponges provide evidence for multiple invasions by Repetitive Hairpin-forming Elements (RHE).** *BMC Genomics*, 10, 591, doi:10.1186/1471-2164-10-591.

103. Morrow, C., Cárdenas, P. (2015) **Proposal for a revised classification of the Demospongiae (Porifera).** *Front. Zool.*, 12(1), 7, doi:10.1186/s12983-015-0099-8.

104. Reitner, J. (1992) **Coralline Spongien: der Versuch einer phylogenetisch-taxonomischen Analyse.** Selbstverlag Fachbereich Geowissenschaften, FU Berlin;

105. Wörheide, G., Reitner, J. (1996) **"Living fossil" sphinctozoan coralline sponge colonies in shallow water caves of the Osprey Reef (Coral Sea) and the Astrolabe Reefs (Fiji Islands).** In: *Göttinger Arbeiten zur Geologie und Palaeontologie*, edited by Reitner J, Neuweiler F, F, G., Göttingen: 145-148.

106. Vacelet, J. (2002) **Recent 'Sphinctozoa', Order Verticillitida, Family Verticillitidae Steinmann, 1882.** In: *Systema Porifera: A Guide to the Classification of Sponges*, edited by Hooper, J. N. A., Van Soest, R. W. M., New York: Springer; 1097-1098.

107. Karlińska-Batres, K., Wörheide, G. (2013) **Microbial diversity in the coralline sponge *Vaceletia crypta*.** *Antonie Van Leeuwenhoek*, 103(5), 1041-1056, doi:10.1007/s10482-013-9884-6.

108. Germer, J., Mann, K., Wörheide, G., Jackson, D. J. (2015) **The skeleton forming proteome of an early branching metazoan: a molecular survey of the biomineralization components employed by the coralline sponge *Vaceletia* sp.** *PLoS One*, 10(11), e0140100, doi:10.1371/journal.pone.0140100.

109. Reitner, J., Wörheide, G. (2002) **Non-Lithistid Fossil Demospongiae - Origins of their Palaeobiodiversity and Highlights in History of Preservation.** In: *Systema Porifera: A Guide to the Classification of sponges*, edited by Hooper, John, van Soest, R. W. M., New York: Springer; 52-68.

110. Reitner, J., Wörheide, G., Lange, R., Schumann-Kindel, G. (2001) **Coralline demosponges, a geobiological portrait.** *Bull*. *Tohoku Univ*. *Museum* 1, 229-235.

Chapter 2:

The Holo-Transcriptome of a Calcified Early Branching Metazoan

Juliane Germer, Nicolas Cerveau, Daniel John Jackson

correspondence: djackso@uni-goettingen.de

Frontiers in Marine Science (2017), 4

DOI: 10.3389/fmars.2017.00081, Open Access

2.1 Abstract

Symbiotic interactions are widespread throughout the animal kingdom and are increasingly recognized as an important trait that can shape the evolution of a species. Sponges are widely understood to be the earliest branching clade of metazoans and often contain dense, diverse yet specific microbial communities which can constitute up to 50% of their biomass. These bacterial communities fulfil diverse functions influencing the sponge's physiology and ecology, and may have greatly contributed to the evolutionary success of the Porifera. Here we have analyzed and characterized the holo-transcriptome of the hypercalcifying demosponge *Vaceletia* sp. and compare it to other sponge transcriptomic and genomic data. *Vaceletia* sp. harbours a diverse and abundant microbial community; by identifying the underlying molecular mechanism of a variety of lipid pathway components we show that the sponge seems to rely on the supply of short chain fatty acids by its bacterial community. Comparisons to other sponges reveal that this dependency may be more pronounced in sponges with an abundant microbial community. Furthermore, the presence of bacterial polyketide synthase genes suggests bacteria are the producers of *Vaceletia*'s abundant mid-chain branched fatty acids, whereas demospongic acids may be produced by the sponge host via elongation and desaturation of short-chain precursors. We show that the sponge and its microbial community have the molecular tools to interact through different mechanisms including the sponge's immune system, and the presence of eukaryotic-like proteins in bacteria. These results expand our knowledge of the complex gene repertoire of sponges and show the importance of metabolic interactions between sponges and their endobiotic microbial communities.

Keywords: sponge, microbes, symbiosis, transcriptome, immunity, metabolism, fatty acid, signalling pathway

2.2 Introduction

Sponges (Porifera) are widely thought to be the earliest branching clade of metazoans [1, 2], with a simple body plan that has remained essentially unchanged for hundreds of millions of years [3]. They have a widespread distribution and are important members of benthic communities. Due to their position on the metazoan tree of life sponges can be used to test hypotheses regarding the evolution of a variety of important traits such as biologically controlled biomineralization [4–7], signaling pathways and gene regulation [8, 9], gastrulation and mesoderm formation [10], and genome evolution.

Sponges are known for hosting dense and diverse microbial communities in their tissues [11] and although it is unclear when this association was established ([11] and references therein) it could be reasonably assumed that sponges and microorganisms have been interacting since the emergence of sponges [12]. Sponge-associated microorganisms have been reported to constitute up to 50% of the sponge's total biomass at densities several orders of magnitude higher than the microbial density in the surrounding seawater [11, 13]. For these reasons sponges are no longer considered to be individual organisms, but as 'holobionts' comprising the sponge host and the sum total of its endobiotic microbial community [14]. Research over the last decade has characterized the phylogenetic composition and diversity of the microbial communities of many shallow water sponges from many different habitats [11]. Despite the fact that sponges are surrounded by, and ingest, a plethora of microorganisms due to their filter-feeding lifestyle, many sponge–associated microorganisms are endemic to sponges [15, 16]. Knowledge regarding the functional roles of sponge-associated microorganisms is also growing. Bacterial symbionts can, for example, provide supplementary nutrition [17, 18], and can remove metabolic waste products such as ammonia, nitrite and nitrate by producing bioactive secondary metabolites [19–21] which can in turn be used by the sponge as chemical defences [22]. How these sponge-microbe relationships are established and maintained, and to what extant these relationships are a genuine symbiosis, remains unclear.

Most sponge-associated microorganisms reside in the sponge mesohyl, an extracellular matrix comprising most of the sponge's body, which is also the site of digestion [11]. This implies that sponge-associated microorganisms must either have a mechanism to avoid digestion by the sponge, for example by producing slime capsules [23], and/or sponges are able to recognize their symbionts. The surprisingly well-developed innate immune system

of sponges [24] is likely to play an important role in distinguishing pathogenic from symbiotic microbe.

Vaceletia is a monospecific, hypercalcified sponge genus with a single described species, *Vaceletia crypta* [25], belonging to the Dictyoceratida within the Demospongiae [26–28]. *Vaceletia* first appears in the Middle Triassic [13] and has been referred to as a "living fossil". Today, *Vaceletia* inhabits cryptic niches, such as caves or deeper fore reef areas that are light reduced and have oligotrophic conditions [29, 30]. *Vaceletia* harbours a dense and diverse microbial community that can make up to 50% of the sponges' biomass [13]. The bacterial community composition of *V. crypta* has been previously studied using 16S rRNA clone libraries and DGGE [31]. This microbial community was highly diverse and shared features with other sponge derived microbial communities. *Vaceletia* occurs in different growth forms (solitary vs. colonial), which are likely to represent different species, however their taxonomy is not yet fully resolved [26]. DGGE cluster analyses indicate distinct microbial communities exist in the solitary and colonial species [31]. In this study we characterize the transcriptome of the yet to be described colonial-branching *Vaceletia* sp. species. We previously generated this sequence resource to assist a proteomic survey of *Vaceletia*'s calcified skeleton [4]. Due to the high abundance and diversity of microbes in this sponge, we reason that interactions between the sponge and its microbial community may also play some role in fabricating its calcified skeleton. This reasoning also follows from our previous work that demonstrated *Astrosclera willeyana* (another hyperclacified demosponge) degrades a proportion of its microbial community via the autophagy pathway [32], and then initiates calcification on the organic remains of these microbes [6]. Here we characterize the transcriptome of an individual adult *Vaceletia* sp. and search for evidence of a variety of sponge-microbe interactions including mechanisms relating to innate immunity, eukaryotic-like proteins in bacteria, and metabolic interactions between the sponge host and its associated bacteria. We also identify the presence of a variety of signaling pathway components and transcription factors that further support the notion that sponges possess a complex gene regulatory repertoire more than adequate to both specify sponge the cell types that must interact with and manage the microbiome, and to coordinate the deposition of an intricate and elaborate calcified skeleton.

2.3 Materials and Methods

2.3.1 Sample Collection and RNA Extraction

Specimens of *Vaceletia* sp. were collected in 2009 during the Deep DownUnder Expedition (http://www.deepdownunder.de) at Osprey- and Bougainville Reefs (Coral Sea, Australia) in depths from 5 to 24 m by SCUBA diving. Sponges do not fall under the German animal protection act §8. This work was therefore exempt from the University of Göttingen Ethics Committee. Samples for RNA extraction were preserved in RNAlater and stored at -20°C. Before processing, samples were carefully inspected under a stereomicroscope for contaminating organisms which were carefully removed. Total RNA derived from the head of one individual was extracted using the miRNeasy Kit (Qiagen) according to the manufacturer's instructions. RNA was checked by gel electrophoresis on a 1.5 % TAE agarose gel and by spectrometry using a Nanodrop spectrometer.

2.3.2 Library Preparation, RNA Sequencing and Sequence Assembly

Library preparation and sequencing was carried out at the Transcriptome and Genome Analysis Laboratory, University Medical Center Göttingen, Germany. A TrueSeq RNA Sample Kit (Illumina, Cat. N°RS-122-2002) was used to generate an mRNA library using 500 ng of total RNA. During TrueSeq Illumina library construction rRNA was depleted (oligo dT beads enriched poly-adenylated mRNAs). While prokaryotes do not extensively adenylate RNA as eukaryotes do, and in theory such poly-A selection should exclude all prokaryotic RNA, it is apparent that generating cDNA libraries with these methods does capture a diverse population of microbial sequences, for example [18] and [33]. However it should be kept in mind that this methodology will be biased against capturing prokaryotic sequences. Accurate quantitation of cDNA libraries was performed using the QuantiFluor™ dsDNA System (Promega) and the size range of the final cDNA libraries was determined using a Bioanalyzer 2100 from Agilent (280 bp). Sequencing was conducted on the Illumina HiSeq 2000 platform to generate 100 bp paired-end reads. Read qualities were assessed using FastQC [34], and Trimmomatic [35] was used to remove low quality reads and adapter sequences. Reads were assembled *de novo* using our recently developed assembly pipeline [36]. Briefly, we employed three assembly packages with unique assembly strategies: Trinity V2.0.3 [37], CLC Genomics Workbench [38] and IDBA-tran V1.1.1 [39]. The CLC assembler was run using a word size of 20 base pairs (bp), a bubble size of 50 bp, with reads mapped back to the transcriptome using default parameters. IDBA_tran was run with kmer values ranging from 20 to 100 bp with a step

size of 10 bp. Trinity assemblies were run with default parameters and a k-mer value of 25. The coding transcriptomes of these three assemblies were merged into a concatenated assembly, and redundancy was removed. The raw reads are deposited to NCBI under SRA accession number SRR4423080.

2.3.3 Transcriptome Annotation and Characterization

The assembled *Vaceletia* sp. transcriptome was annotated using BLASTx searches (with an e-value cut-off of 1e-5) against the NCBI non-redundant database, using Trinotate [37], BLASTx and BLASTp searches (with an e-value cut-off of 1e-5) against the Swissprot/Uniprot database, and against Pfam [40] using Hmmer [41]. SignalP 4.0 [42] and TMHMM [43] were used to predict signal peptides and transmembrane regions, respectively.

Transcripts were assigned a putative taxonomic origin using the lowest common ancestor algorithm implemented in MEGAN5 [44] with the following parameters: MinScore = 50, MaxExpected = 0.01, TopPercent 10 and MinSupport = 20. In order to separate and compare contigs with bacteria and metazoan origin, contigs identified as such were extracted from MEGAN and used in individual BLASTx searches. The results were visualized in MEGAN, and a comparison between the bacterial and metazoan contigs was constructed using the 'compare' option with normalized counts. Functional annotation was performed using the KEGG database implemented in MEGAN. Sequences that were assigned by MEGAN to a bacterial origin were extracted and further characterized using BLAST searches as described above. Potential eukaryotic-like proteins were identified via keyword search of the BLASTx output file using a customized Perl script. Matching contigs were further evaluated using hmmer against the Pfam 28.0 database (with a cut-off value of 1e-5) for the presence of eukaryotic-like protein domains (ankyrin repeat (ANK): PF00023, tetratricopeptide repeat (TPR): PF00515.24, PF13374.2, PF13414.2, PF13424.2, PF13428.2, PF13429.2, PF13432.2, PF13431.2, PF13512.2, PF14559.2, PF07719.13, PF09976.5, PF07720.8, PF07721.10, PF13174.2, PF13176.2, PF13181.2, PF13371.2, SEL1: PF08238, leucine rich repeat (LRR): PF00560.29, PF07723.9, PF07725.8, PF12799.3, PF13306.2, PF13516.2, PF13855.2, PF14580.2, fibronectin: PF00041, cadherin: PF00028, NHL (named after NCL-1, HT2A and Lin-41): PF01436).

Pattern Recognition Receptors (PRRs) were identified from the *Vaceletia* sp. BLASTx output file using a customized Perl script. The script scanned the description lines of the BLASTx output file for key words and collected all positive contigs. These contigs were further evaluated for the presence of the PRR domains using hmmer against the Pfam 28.0 database with a cut-off value of 1e-5.

Genes involved in short-chain fatty acid biosynthesis were identified via the blast based KEGG pathway map coverage implemented in MEGAN. To get an overview of the type I fatty acid synthase (FAS), type II FAS and polyketide synthase (PKS) diversity the transcriptomes of *Vaceleti*a sp., *Amphimedon queenslandica*, *Stylissa cateri*, *Petrosia ficiformis*, *Spongilla lacustris*, *Pseudospongosorites suberitoides*, *Xestospongia testudinaris*, *Sycon coactum* and *Corticum candelabrum* were screened for contigs containing a ketosynthase (KS) domain using hmmer against the KS HMM profiles PF00109 and PF02801 with a cut-off value of 1e-5. Positive matches were extracted from the respective transcriptomes. For phylogenetic analyses protein reference sequences were downloaded from NCBI genebank and the NaPDoS database [45]. Searching animal derived type I FAS sequences against porifera sequences (taxid: 6040) stored in the NCBI nr database did not yield any significant hits. Since only KS domains were used in the construction of a phylogenetic tree, the NaPDoS server [45] was used to extract the KS domain from all sequences. Many *Vaceletia*-derived sequences only span a small region of the KS domain and were discarded. Full and partial KS amino acid sequences (at least 160 amino acids) were aligned using clustalO [46]. Gblocks [47] was used to identify conserved sites. Phylogenetic analysis were conducted using MrBayes v 3.2.6 [48] with the following parameters: lset rates = gamma, preset aamodelpr = mixed, mcmcp, nruns = 4, ngen = 10,000,000, relburn = yes, burninfrac = 0.25, printfreq = 1000, samplefreq = 100, nchains = 4, saveprlens = yes and RAxML on the RAxML blackbox webserver [49] using the combined bootstrapping and maximum likelihood search algorithm with the WGA model and 100 bootstraps.

Bacterial PKS genes were identified by searching a list of sponge-specific PKS sequences retrieved from Genbank against the bacterial derived contigs of the *Vaceletia* sp. transcriptome. Each predicted *Vaceletia* sp. sequence was then searched against the nr database using BLASTx. Domain structure was evaluated with hmmer against the Pfam 28.0 database with a cut-off value of 1e-5.

Genes involved in fatty acid elongation and desaturation were identified from the *Vaceletia* sp. BLASTx output file using a customized Perl script. The script scanned the description lines of the BLASTx output file for key words and collected all positive contigs. Additionally, genes were identified by searching a list of fatty acid elongation and desaturation sequences retrieved from Genebank against *Vaceletia* sp. transcriptome. Positive and matching contigs were further evaluated using hmmer against the Pfam 28.0 database with a cut-off value of 1e-5.

Sterol-24/28-methyltransferase (SMT) genes were identified by searching a list of sponge-specific SMT sequences retrieved from Genbank against *Vaceletia* sp. transcriptome. Each predicted *Vaceletia* sp. sequence was then searched against the NCBI-nr database using BLASTx. Positive matches were aligned to the SMT protein alignment provided by Gold et. al [50] using clustalO. Phylogenetic analyses were conducted using MrBayes v 3.2.6 [48] with the following parameters: lset rates = gamma, preset aamodelpr = mixed, mcmcp, nruns = 4, ngen = 10,000,000, relburn = yes, burninfrac = 0.25, printfreq = 1000, samplefreq = 100, nchains = 4, saveprlens = yes and RAxML on the RAxML blackbox webserver [49] using the combined bootstrapping and maximum likelihood search algorithm with the GTA model and 100 bootstraps.

Protein sequences from a collection of recently reported sponge transcriptomes [33] were used in local BLAST searches against the *Vaceletia* sp. transcriptome to identify conserved signaling proteins. Matching contigs were further evaluated using hmmer against the Pfam 28.0 database with a cut-off value of 1e-5.

Homeobox containing genes were identified by searching the homeobox domain (PF00046) and homeobox KN domain (PF05920) against the *Vaceletia* sp. transcriptome using hmmer with a cut-off value of 1e-5. Contigs containing these domains were extracted and further characterized using the Homeobox database [51, 52].

2.3.4 Comparison of the *Vaceletia* sp. Transcriptome to Other Sponge Transcriptomes and Transcriptome Completeness

In order to place our results within a broader poriferan context, we included a range of other datasets. Riesgo reported eight transcriptomes derived from representatives of all poriferan classes [33] using the same sequencing strategy as we employed here (Illumina

TruSeq poly-A mRNA library preparation). These were BLAST searched against the nr database, loaded into MEGAN and annotated against KEGG as described in section 2.3. Raw Illumina reads for *Amphimedon queenslandica* (SRR1511621), *Xestospongia testudinaris* (SRR1738066/68) and *Stylissa cateri* (SR1738063) were assembled as described for the *Vaceletia* data, and also BLAST searched against the nr database, loaded into MEGAN and annotated against KEGG as described above. In addition, to evaluate the completeness of all the transcriptomes we investigated, we performed a BUSCO analysis [53] (Supplementary Table S1).

2.3.5 Lipid Extraction and GC/MS Analysis

Lipids were extracted from crushed, freeze-dried heads of *Vaceletia* sp. by ultrasonication with DCM/methanol (1:3, v/v), DCM and n-hexane, respectively. The total organic extract was derivatized using trimethylchlorsilane (TMCS)/methanol (1:9; v:v; 1.5 h at 80°C), yielding fatty acid methyl esters (FAME). Separation was achieved by column chromatography. The extract was dried on silica gel 60, loaded onto the column and eluted successively with n-hexane/DCM (5:1, v/v), DCM, DCM/acetone (9:1, v/v) and DCM/methanol (1:1, v/v) to yield hydrocarbons, fatty acid, alcohols and polar residue, respectively. Gas chromatography-mass spectrometry (GC/MS) analyses were carried out using a Varian CP-3800 gas chromatograph coupled to a Varian 1200L mass spectrometer. The system was equipped with a silica column (Phenomenex Zebron ZB-5MS, 30 m, 0.25 µm film thickness, inner diameter 0.32 mm). Samples were injected on column. The carrier gas was helium. The temperature programme was 80°C (3 min) to 310 °C at 4 °C min per minute (held 25 min). Mass spectra were produced using an electron energy of 70 eV, a mass range of m/z 50-650 and a scan time of 0.25 s in full scan mode.

2.4 Results and Discussion

2.4.1 Transcriptome Characterization

Our sampling and sequencing strategy (sequence using the Illumina HiSeq2000 platform from a single *Vaceletia* sp. individual) was designed to minimize allelic variation in order to assist bioinformatic assembly of the raw data. While the bioinformatic assembly of such data derived from nonconventional organisms is always challenging [36], this experimental approach should aid in generating an assembly as close to the biological truth as possible. However we do acknowledge that biological and/or temporal replication may reveal additional patterns not detected here. Quality filtering of raw reads resulted in 34,592,975

paired end reads that were assembled as described by Cerveau and Jackson [36] (see Methods section). The assembly generated 91,443 contigs with an average length of 933 bp and an N50 value of 1,307 (Table 1). Assembled contigs were compared to the NCBI non-redundant database and to the Swissprot/Uniprot database using BLASTx (with an e-value of 1e-5). In total 76,207 contigs (83.3%) and 70,734 (65.7%) were annotated, respectively. Using TransDecoder 107,655 coding sequences (cds) were identified and used in BLASTp searches against the Swissprot/Uniprot database which resulted in 65,722 (61.0%) hits. Further, conserved Pfam domains were assigned to 69,083 (64.2%) of the translated queries. In total 24,893 coding sequences (23,1%) could not be annotated by any of these sequence similarity searches (Table 2). This is a relatively high proportion in comparison with other sponge transcriptomes generated using similar methods [33].

TABLE 1 | *Vaceletia*. sp. transcriptome assembly metrics

Number of transcripts	91,443
Maximum transcript length (bp)	12,237
Minimum transcript length (bp)	300
Mean transcript length (bp)	933
Median transcript length (bp)	643
Total length (bp)	82,274,648
N50 (bp)	1,307

TABLE 2 | Results of sequence similarity searches

Input	**Method**	**Number of annotated transcripts**
all transcripts	BLASTx against NCBI nr prot	76,207 (83.3 %)
all transcripts	Transdecoder – coding sequence (cds) finder	107,655
cds	BLASTx against Swissprot	70,734 (65.7 %)
cds	BLASTp against Swissprot	65,722 (61.0 %)
cds	Pfam	69,083 (64.2 %)

2.4.2 Taxonomic Distribution of *Vaceletia* sp. Contigs

To obtain information regarding the taxonomic origin of the assembled contigs, BLASTx output files were used as input for MEGAN [44]. MEGAN uses a Lowest Common Ancestor algorithm to assign protein-coding (mRNA) genes to a taxon, which can then be visualized in a phylogenetic tree. A MEGAN output of *Vaceletia* sp. contigs is shown in Fig. 1. 367,018 *Vaceletia* sp. contigs were assigned to bacteria (40%), 29,947 contigs to Metazoa (33%) and 18,221 contigs (20%) could not be taxonomically assigned.

Among the *Vaceletia* sp. sequences that shared significant similarity with a metazoan sequence, a high proportion (6,377 contigs, 21%) were assigned to sponges (mostly *Amphimedon*). However, the majority shared the highest similarity either with a specific bilaterian sequence (11,674 contigs, 39%), or with an unspecific sequence and were therefore assigned to the metazoan node (11,302 contigs, 38%) (Fig. 1B). A recent characterization of eight sponge transcriptomes gave similar results [33]. Although the sequence similarity hit rate for our *Vaceletia* transcriptome is higher, the significant proportion of contigs assigned to the "unspecific" category highlights the lack of curated database entries for sponges.

A striking number of contigs in our dataset (40%) were taxonomically assigned to Bacteria. This is despite the fact that a polyA RNA selection was performed during TruSeq library construction (see Materials and Methods above). Some sponges are known to harbour a remarkable diversity of microorganisms. In *Vaceletia crypta* the total biomass derived from bacteria has been reported to be as high as 50% [13]. Recent molecular studies using 16S rRNA characterized the microbial communities of sponges from many marine habitats. This work paints a picture of a handful of dominant sponge-associated bacterial phyla, including Proteobacteria (especially the classes Alpha-, Gamma- and Deltaproteobacteria), Chloroflexi, Actinobacteria, Acidobacteria, Nitrospirae and the candidate phylum Poribacteria [12]. All of these sponge-associated Bacteria phyla are present in our *Vaceletia* sp. transcriptome dataset (Fig. 1C). As much as 24% of the bacterial contigs were assigned to Proteobacteria, followed by 13% to Chloroflexi and 9% to the candidate phylum Poribacteria. Traditionally, 16S rRNA is used for both taxonomic characterization of microbial communities, as was previously performed for *V. crypta*, the solitary sister species to *Vaceletia* sp. [31]. Taxonomic assignment as performed by MEGAN employs protein-coding genes. Interestingly, in comparison to previously reported sponge transcriptomes [33, 54] the *Vaceletia* sp. transcriptome contains by far the highest proportion of bacterially derived contigs (40%) in comparison to high microbial abundance (HMA) sponges such as *X. testudinaris* (1.4%), *C. nucula* (7.5%) and *C. candelabrum* (13.7%). It must be kept in mind that comparisons between studies such as this are likely to suffer from some systematic error associated with library preparation and downstream data handling. However broadly similar library preparation (ribosomally depleted mRNA and/or polyA selected RNA inherent to the TruSeq library preparation

protocol), sequencing (Illumina platform) and sequence data handling methods were employed by all of these studies.

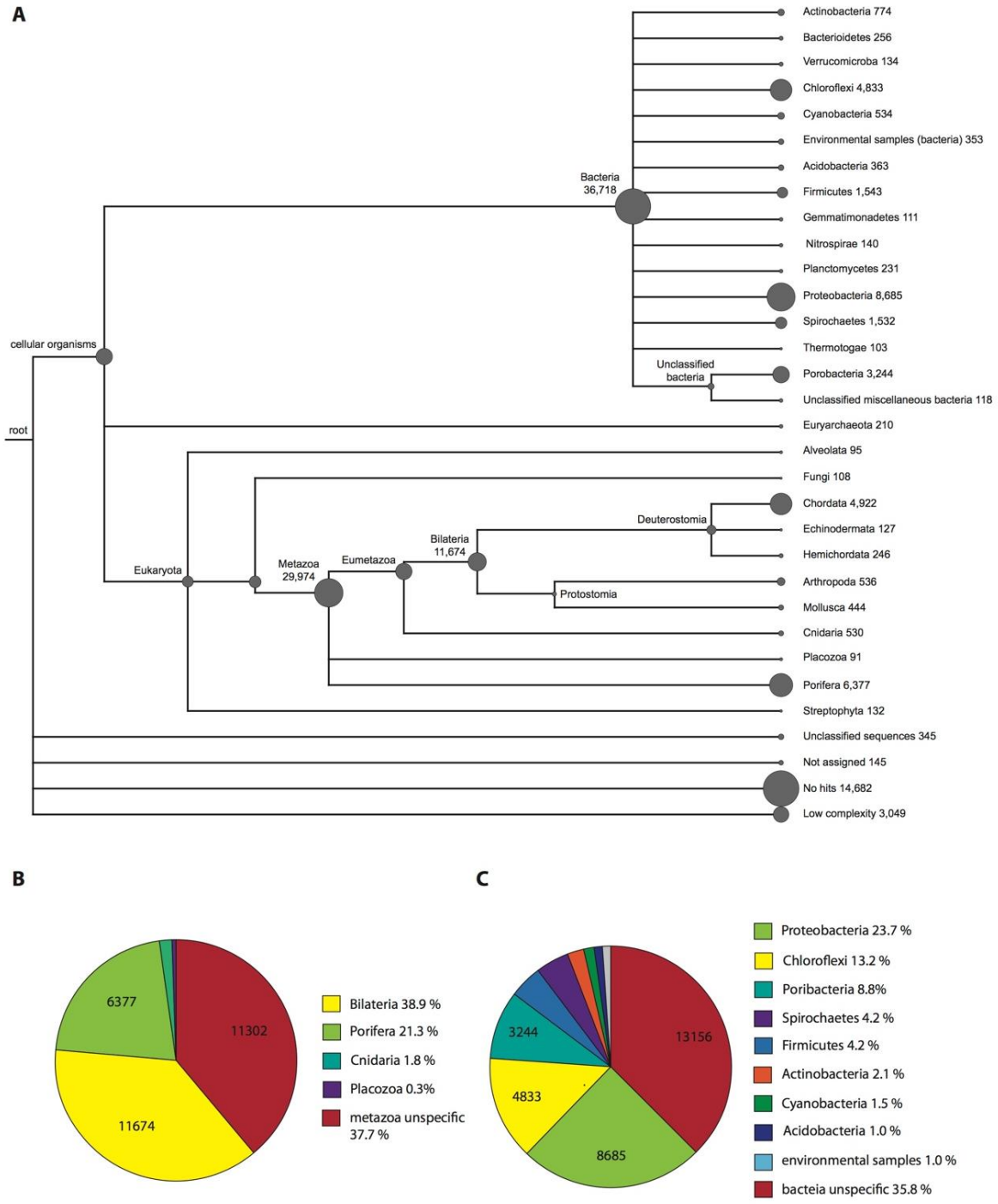

FIGURE 1 | Taxonomic assignment of *Vaceletia* sp. contigs using MEGAN. (**A**) A phylogenetic tree reflecting the relationships and abundance of the taxonomic assignments. (**B**) Further phylogenetic breakdown of contigs assigned to a metazoan origin (**C**) Further phylogenetic breakdown of contigs assigned to a bacterial origin.

2.4.3 Metabolic Interactions between *Vaceletia* sp. and Its Microbial Community

It has been suggested, that the evolutionary history and the ecological success of sponges is closely intertwined with their microbial community [11]. Sponge-microbe symbioses are likely evolutionary ancient, and have probably co-evolved to yield benefits to both sponge and microbe. Microbes living within sponges are provided with a nutrient rich environment through the products of the host's metabolism [55]. In return, microbes can provide the sponge host with access to novel metabolic pathways and chemical defences [18, 55–59]. Taking a closer look at the metabolic pathways present in the *Vaceletia* sp. transcriptome, a significant proportion of all metabolic pathways are derived from contigs originating from bacteria (Fig. 2). Comparing bacterial and metazoan features in these metabolic pathways reveals similar results to those reported by Fiore et al for the sponge *Xetospongia muta* and its symbiotic community [18]. Because the metabolism of fatty acids has received some attention in sponges we focused on this aspect of the sponge-microbe metabolic relationship.

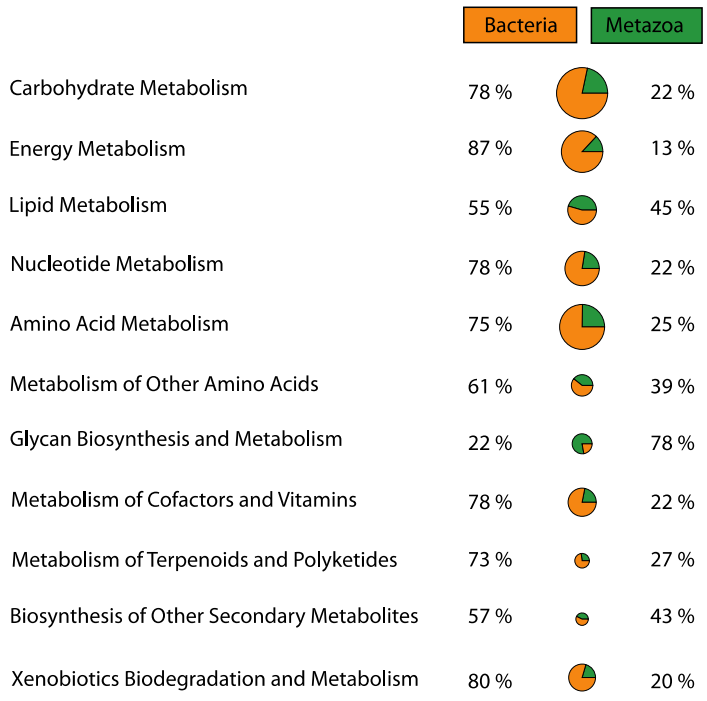

FIGURE 2 | Proportions of metazoan derived and bacterial derived contigs involved in a variety of metabolic pathways.

Lipid Metabolism: Biosynthesis of Short Chain Fatty Acids

Sponges are a rich source of unusual lipids that play a functional and structural role in their membranes. It has been proposed that a significant amount of these lipids are of bacterial origin [60]. The fatty acid (FA) compounds of *Vaceletia* sp. show that this sponge is rich in

short chain FAs, mid-chain branched fatty acids (MBFAs) and the sponge specific demospongic acids (Supplementary Fig. S1). Cell fraction experiments demonstrated the occurrence of short chain FAs in both sponge cells and bacteria, with increased levels in the bacterially derived matter [61, 62]. In animals, fatty acid biosynthesis is catalyzed by the monomodular type I fatty acid synthase (FAS), whereas in bacteria it is usually catalyzed by individual enzymes, the type II FAS [63]. Mapping *Vaceletia* sp.'s transcriptome against the KEGG lipid metabolism pathways shows that most components of the short chain FA biosynthesis pathway are present (Fig. 3). All compounds are represented by bacterial derived contigs and are mapped to the type II FAS, suggesting a purely bacterial origin for short-chain FAs in *Vaceletia*.To determine whether this is generally the case for sponges, the transcriptomes of *I. fasciculate*, *C. nucula*, *P. ficiformis, S. lacustris*, *P. suberitoides, X. testudinaris, S. cateri, S. coacatum* and *C. candelabrum* were also mapped against the KEGG lipid metabolism pathway (Fig.4). Only *C. candelabrum* gave similar results to *Vaceletia* sp. with three components represented by both bacterial and metazoan derived contigs. Both sponges are HMA sponges with the highest percentage of bacterial derived contigs of all transcriptomes. In *X. testudinaris,* also a HMA sponge, the majority of components are present but in contrary to *Vaceletia* they map to the type I FAS pathway, implying a metazoan origin of short chain FAs. *X. testudinaris* is the only sponge to possess an animal type FAS. The lack of bacterial derived components might also be the result of the extremely low percentage of bacterially derived contigs (1.4%) in that transcriptome. Although the short chain FA biosynthesis pathway is incomplete or missing in the other sponge transcriptomes, a general trend is observable. HMA sponges tend to have more bacterial derived pathway components then LMA sponges (Fig. 4). No literature regarding the classification of *P. suberitoides, A. vastus* and *S. coactum* as HMA or LMA sponges is available, but the low percentage of bacterial derived sequences (0.8%, 1.1% and 1.7%, respectively) suggests that they are LMA sponges. Genomic data from *A. queenslandica* indicates an incomplete short chain biosynthesis pathway, also suggesting that this sponge is not itself capable of producing short chain fatty acids [64]. Additionally, we identified contigs containing a KS domain from *Vaceleti*a sp., *Amphimedon queenslandica, Stylissa cateri*, *Petrosia ficiformis, Spongilla lacustris*, *Pseudospongosorites suberitoides, Xestospongia testudinaris, Sycon coactum* and *Corticum candelabrum*. This domain is present in type I FASs, in the component FabF of the type II FASs and in polyketide synthases (PKSs), a protein class that is evolutionary linked to type I and type II FASs and resembles the architecture of

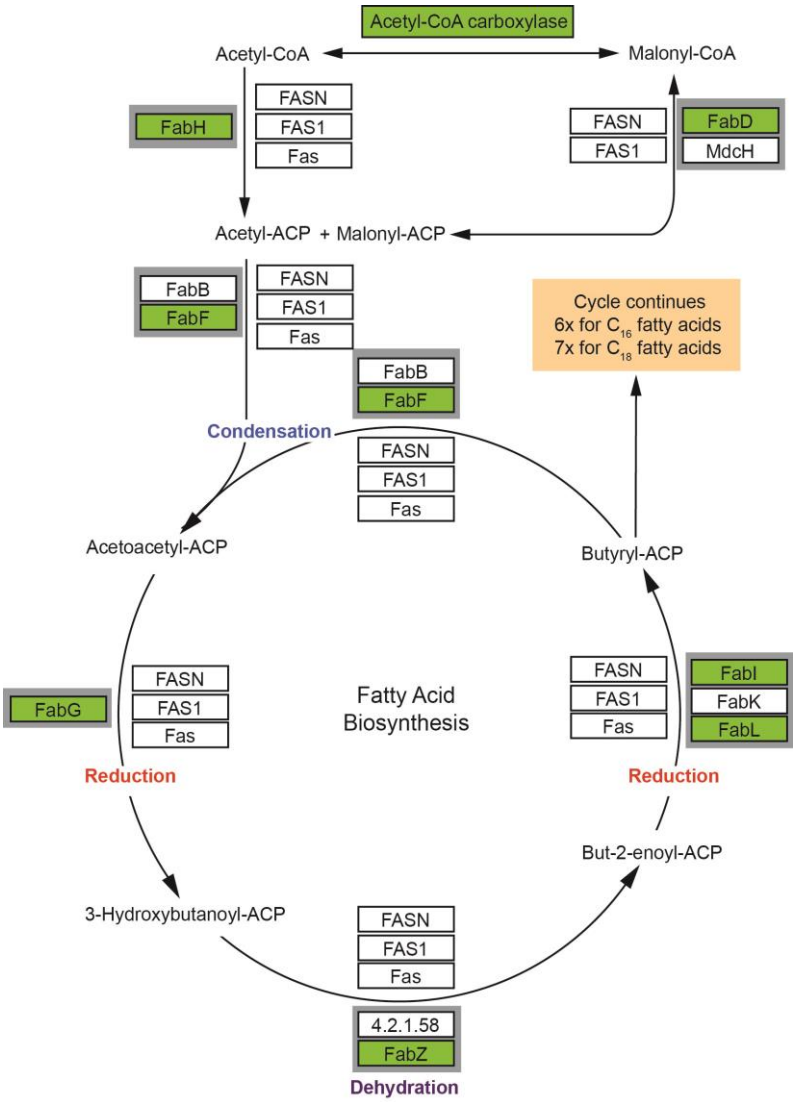

FIGURE 3 | Condensed short chain fatty acid biosynthesis pathway. Coverage of the short chain fatty acid transcriptome biosynthesis pathway**.** MEGAN was used to visualize contigs on the KEGG map. Rectangles represent proteins of the pathway while arrows represent signaling routes. Pathway components showing significant similarity to contigs in the *Vaceletia* sp. are highlighted in green. Type II FAS is highlighted in grey. Condensation of malonyl –ACP with acetyl-ACP initiates fatty acid synthesis. Successive steps of reduction, dehydration and reduction lead to chain elongation via the type I FAS used by animals or the type II FAS used by bacteria. Butyryl-ACP, with the addition of malonyl-ACP, then undergoes subsequent cycles of elongation. Each round elongates the FA to have two more carbon units.

these [65]. Bayesian and maximum likelihood phylogenetic analyses were conducted on the conserved KS domains. Both phylogenetic reconstructions show, that the type II sequences identified via the KEGG pathway-mapping cluster together with type II FAS sequences from bacteria within the Type II FAS clade (Supplementary Fig. S2). No sequence recovered by our analysis falls within the animal type I FAS clade. However, our results suggest that a variety of sponges rely on the supply of short chain FAs from their endobiotic microbial communities. Whether these communities present a final FA product

to the host, or whether the sponge farms and then harvests these communities for their FA resources is as yet unknown.

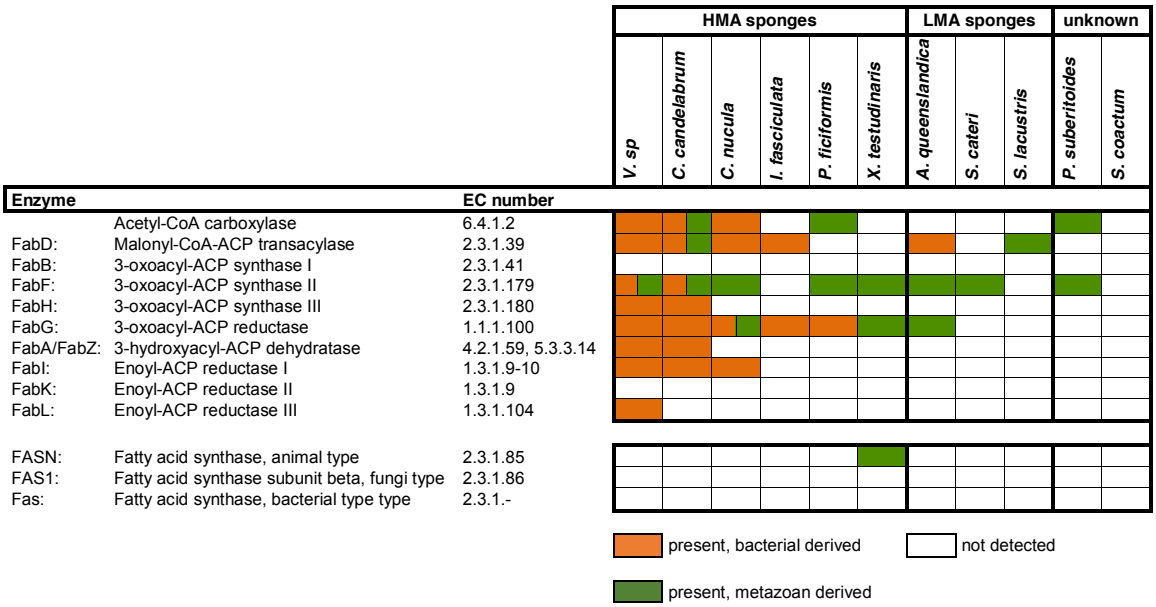

FIGURE 4 | Presence and absence of metazoan and/or bacterial derived short-chain fatty acid biosynthesis components. Components were identified by mapping the transcriptomes to the corresponding KEGG map.

Lipid Metabolism: Biosynthesis of Methyl-Branched Fatty Acids

Another prominent compound class of *Vaceletia's* biomarker content are the mid-chain branched fatty acids (Supplementary Fig. S1). MBFAs are often abundant compounds in HMA sponges and are believed to have a bacterial origin [66]. A recent study positively correlated the presences of a sponge specific polyketide synthase (PKS) in bacteria-rich sponges with the presence of MBFAs and proposed Poribateria as the potential producers of these compounds [20]. As *Vaceletia* sp. contains large amounts of MBFAs (Supplementary Fig. S1) and hosts Poribacteria (Fig. 1), we examined the bacteria derived subset of *Vaceltia*'s transcriptome for the occurrence of a bacterial-type PKS named "sponge symbiont ubiquitous PKS" (Sup-type PKS) [67]. This group of genes is commonly found in sponges, but is also highly specific to sponges [67], and has also been found in the genome of a representative Poribacterium [68]. Various *Vaceletia* sp. contigs share significant sequence similarity with the SupA gene isolated from a bacterial symbiont of the sponge *Theonella swinhoei*. Two of these contigs (Supplementary Table S2) contain all of the previously identified domains [67], including the methyltransferase

domain, which most probably is responsible for the mid-chain methyl branch. In *Mycobacterium tuberculosis* complex lipids are part of the cell wall and seem to be potent virulence factors by modulating the host cell function [69]. The question has been raised if Sup genes might play a similar role in sponge symbionts and thus be important for establishing and maintaining symbioses in sponges [67], but so far the function of MBFAs in sponge symbionts is still unknown.

Lipid Metabolism: Biosynthesis of Demospongic Acids

The most abundant FA compounds of *Vaceletia's* biomarker content are demospongic acids (Supplementary Fig. S1). Demospongic acids are non-methylene interrupted fatty acids with an atypical 5,9-diunsaturation pattern within a carbon number range from C_{20} to C_{34} [70]. They are named after their first proven presence in different demosponges and although not exclusive to sponges [70], they are a characteristic feature of demosponges and hexactinellids [71]. Early studies with ^{14}-C-labbled FAs on the sponge *Microciona prolifera* demonstrated that demospongic acids are biosynthesized via elongation of short chain precursors mainly derived from exogenous sources, followed by desaturation (Fig. 5 A) [72–74]. Probably the most remarkable feature of the Δ5, 9-FA biosynthesis is the desaturation process as it was found that the double bond can be introduced at either the Δ5 or Δ9 loci, meaning that the second double bond may be inserted on either site of the existing one, which is contrary to the formation of polyunsaturated FAs in animals [72, 73]. These studies demonstrate that sponges must possess an active elongation and desaturation enzyme system, however the underlying molecular processes remain largely unstudied [75]. The end product of the short chain fatty acid synthesis via FAS is palmitic acid (C_{16}) and to a lesser extent myristic (C_{14}) and stearic (C_{18}) acid. To produce FAs with chain length > C_{18} animals use a system involving four membrane-bound enzymes [76]. *In silico* searches of the *Vaceletia* sp. transcriptome for these enzymes, namely elongation of very long chain fatty acids (Elovl) proteins, ketoacyl-CoA reductase (KAR), β-hydroxyacyl-CoA dehydratase (HADC) and *trans*-2,3-enoyl-CoA reductase (TER), reveal a variety of contigs (Fig 5 B, Supplementary Table S3). As elongation is followed by desaturation in sponges we additionally searched for Δ5,9 desaturases in *Vaceletias*'s transcriptome and found contigs showing high similarities to Δ5 desaturase and stearoyl-CoA desaturase (Supplementary Table S3). Stearoyl-CoA desaturase is an enzyme that is ubiquitous present in living organism [77] showing Δ9 desaturase activity. Almost all compounds are represented by metazoan derived contigs confirming that sponges possess

an active elongation and desaturation enzyme system analogous to what is known from animals to produce endogenous demospongic acids and other long chain saturated and unsaturated FAs.

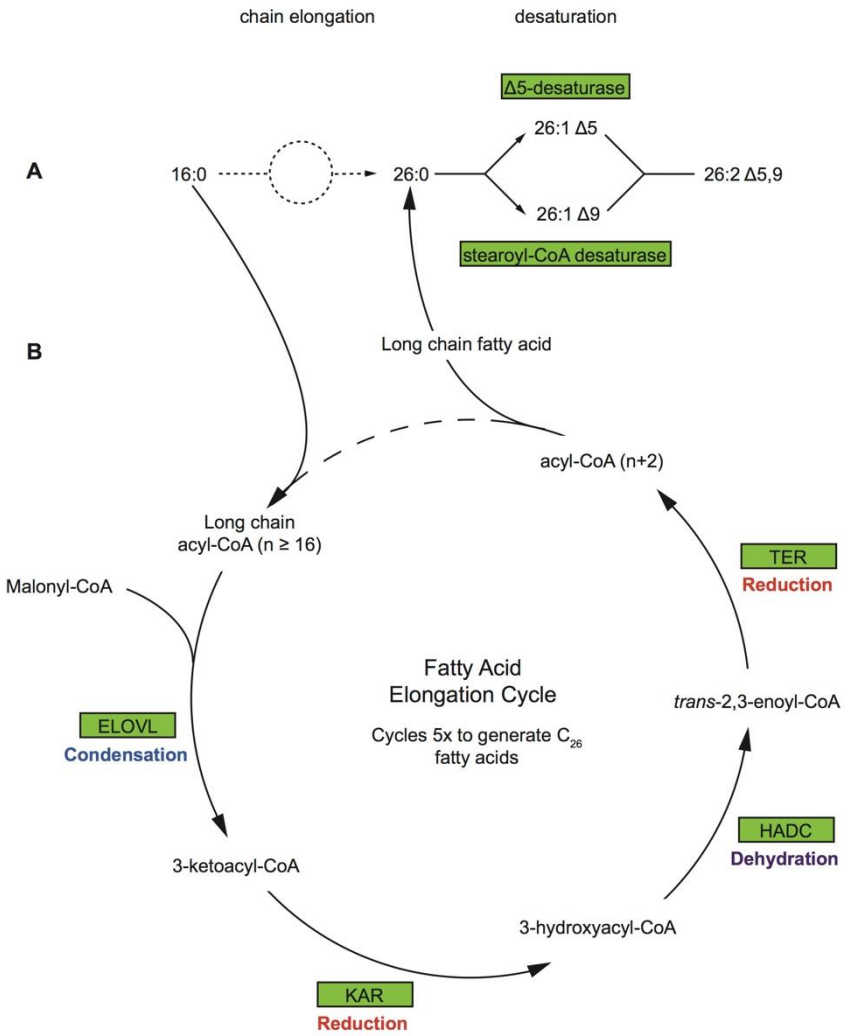

FIGURE 5 | Biosynthetic pathway of demospongic acids and fatty acid elongation pathway (A) As highlighted demonstrated in the sponge *Microciona prolifera* [72–74]. **(B)** The FA elongation cycle consists of four steps and each round elongates the FA to have two more carbon units. Rectangles represent proteins of the pathway while arrows represent signaling routes. Pathway components showing significant similarity to contigs in the *Vaceletia* sp. transcriptome are in green. ELOVL = elongation of very long chain fatty acids; KAR = ketoacyl-CoA reductase; HADC = β-hydroxyacyl-CoA dehydratase; TER = trans-2,3-enoyl-CoA reductase.

Lipid Metabolism: Biosynthesis of Sterols

Sterols occur ubiquitously among eukaryotes and fulfill important biologically functions; they are important for membrane structure and fluidity and function as precursors to signaling molecules or hormones. Sponges are a rich source of a great variety of sterol

compounds with unusual structural features [78] that sometimes are so unique that they might be used to provide taxonomical information of these organisms [79]. Sponges can acquire sterols in different ways; by *de novo* biosynthesis, by dietary uptake, by modifications of dietary sterols or by biosynthesis with associated microorganisms [78, 80].

Mapping *Vaceletia's* transcriptome against the KEGG steroid biosynthesis pathway shows that the majority of the components to synthesize cholesterol are present and are mostly represented by metazoan derived contigs (Supplementary Table S4). However, components in the KEGG pathway to produce lanosterol are mapped to bacterial derived contigs (Supplementary Table S4). Lanosterol is an intermediate product of the steroid biosynthesis pathway and is generally regarded as the precursor to all animal sterols [78]. Sterol biosynthesis is considered a eukaryotic key feature, nevertheless sterol production has been observed recently in a few bacterial species [81–83]. Interrogation of *Vaceletia'*s transcriptome indicate that the sterol precursor is not synthesized by the sponge *de novo* but rather acquired by dietary uptake or produced by associated microorganism and then further modified by the sponge. Searching transcriptomes of *I. fasciculate*, *C. nucula*, *P. ficiformis, S. lacustris*, *P. suberitoides, X. testudinaris, S. cateri, S. coacatum* and *C. candelabrum* against the KEGG steroid biosynthesis pathway confirms the results of a previous study [50] that the sterol precursor is either synthesized by the sponge *de novo* or acquired by dietary uptake. These transcriptomes show no evidence that associated microorganisms might somehow be involved in the biosynthesis of sterols in these sponges (Supplementary Table S4).

Characterizing sterols and their biosynthetic origin is interesting and important as their diagenetically derived counterparts, steranes, can be used by geochemists as "molecular fossils". One unique C_{30} sponge sterol that is of particular interest is 24-isopropylcholesterol and its geological counterpart 24-isopropylcholestane [84]. 24-isopropylcholestane is widely accepted as a sponge-specific biomarker. It is found abundantly in certain Neoproterozoic to Early Cambrian rocks (~650-540 mya) representing the oldest evidence for animal life in the geological record [50, 85]. 24-isopropylcholesterol is produced by certain demosponges [85] and in trace amounts by pelagophyte algae [86] by using the enzyme sterol-24/28-methyltransferase (SMT). Molecular clock analysis suggest that sponge SMT evolved before the Phanerozoic

whereas SMTs of the algaea evolved approximately 100 mya after that, indicating that 24-isopropylcholestane indeed is a sponge specific biomarker [50]. Across eukaryotes a strong correlation between C-24 side chain alkylation and the number of copies of the gene SMT exists in the genome, implicating that each SMT copy is responsible for one alkylation step [50]. Sponges seem to be the exception as fewer SMT copies as expected were consistently recovered to produce C_{29} and C_{30} sterols [50]. Hence, to produce 24-isopropylcholesterol sponges need two SMT copies (named SMT1 and SMT2) whereas other eukaryotes would need at least three SMT copies [50]. BLAST searches for SMT genes recovered two potential SMT1 contigs from the *Vaceletia* transcriptome. These contigs show significant similarities to a predicted cycloartenol-C-24-metyltransferase from *Populus euphratica* (eudicot) and to a hypothetical protein of *Sphaeroforma artica* (unicellular organism), respectively. Phylogenetic analyses cluster contig CLC_28234 with SMTs from trypanosomes (parasitic protozoans) and contig idb_20522 with Ichthyosporea (parasitic Holozoa) (Supplementary Figure S3). These results suggest that the detected SMT genes are not sponge derived but rather a contamination from an endo- or epibiont.

2.4.4 Characterization of the Expressed Immune System of *Vaceletia* sp.

Due to the high proportion of bacterially derived contigs from the *Vaceletia* sp. transcriptome, genes that regulate interactions between the sponge host and these microbes must also be expressed and present in this transcriptome. Our interest in the evolution of mechanisms that control metazoan biomineralization previously uncovered deep relationships between sponges and their microbial communities [6], and also revealed connections with biomineralization that have deep evolutionary origins [7]. In order to better understand how *Vaceletia* sp. manages its microbial community, with the prospect of understanding how this relationship may influence the construction of its highly calcified skeleton [4], we also surveyed and characterized some components of *Vaceletia*'s innate immune system.

The first contact point between entering bacteria and the sponge host are so-called germ line encoded pattern recognition receptor (PRR) [87]. PRRs detect and bind microbial ligands known as microbial- or pathogen-associated molecular pattern (MAMPs and PAMPs, respectively), which typically trigger a signal transduction cascade that leads to the transcription of immune response effector genes [88]. Ligands that are recognized by PRR include microbial glycolipids and glycoproteins, peptidoglycans of gram-positive and

gram-negative bacteria, and lipopolysaccharides of gram-negative bacteria [87] as well as recognition mechanisms for fungi and viral DNA [89].

Toll-like Receptors

Toll-like receptors (TLR) are among the most prominent and best characterized members of PRRs in animals. They are membrane bound receptors canonically composed of an intracellular Toll/interleukin-1 receptor (TIR) domain and extracellular leucine-rich-repeat (LRR) domains, which are responsible for the extracellular pattern recognition. So far, no conventional TLRs have been found in sponges [12, 90], and we find no conventional TLRs in the transcriptome of *Vaceletia* sp. We detected several sequences with similarity to TLRs, and domain searches revealed that the majority of these sequences contain a single TIR domain but are devoid of the characteristic LRR domains and therefore lack the MAMP-binding site. Instead of the LRR domain, three contigs possess an extracellular Immunoglobulin (IG) domain (Supplementary Table S5). In *Amphimedon queenslandica* the presence of TLR related receptors with an intracellular TIR domain and an extracellular Immunoglobulin (Ig) domain instead of LRR domains have been reported [64, 90]. In general Toll-like receptor proteins are reduced in number in sponges. Surveys of eight sponge transcriptomes revealed the presence of toll-like receptor 2 in only three sponges, *Ircinia fasciculata*, *Petrosia ficiformis* and *Corticum candelabrum* [33] (Fig. 6). Although it has been suggested that the Ig domains may interact with microorganisms, the lack of diversity of these receptors suggest that they play little or no role as PRRs in sponges [90]. The myeloid differentiation primary response gene 88 (MyD88) is another TIR domain containing protein with a characteristic DEATH domain and could be identified in the *Vaceletia* sp. transcriptome (Supplementary Table S5). MyD88 functions downstream of TLRs and has previously been identified in other sponges [33, 64, 91] (Fig. 6).

Nucleotide-Binding Domain and Leucine-Rich Repeat Proteins (NLR)

TLRs are responsible for detecting extracellular MAMPs and PAMPs. Most cytosolic counterparts of TLRs are NLRs (Nucleotide-binding domain and Leucine-rich repeat protein also known as Nucleotide Oligorimerisation Domain (NOD)-like receptors). NLRs react to bacteria that invade the cell and to bacterial products remaining after phagocytosis [92], or to endogenous damage-associated molecular patterns (DAMPS) that indicate cellular stress or injury [93]. Surprisingly, intracellular NLRs are not only important for the

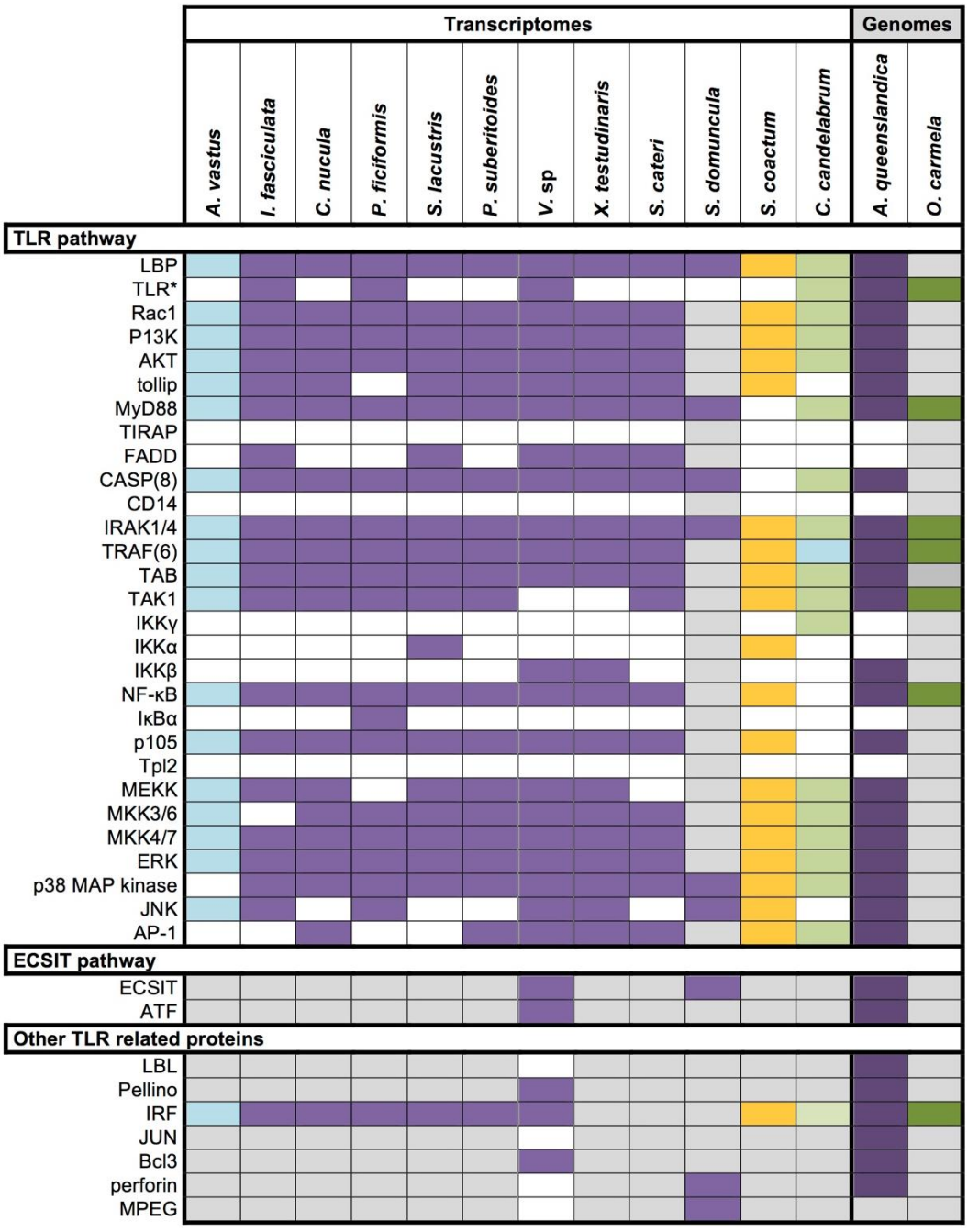

FIGURE 6 | Overview of innate immunity components in Porifera Data derives from recent papers [33, 90, 91, 106] and our own analysis. Different colors represent different sponge class; blue = Hexactinellid; purple = Demosponge; yellow = Calcarea; green = Homoscleromorpha; Colored square = component present; white square = component absent; grey square = unknown; * = includes unconventional Toll-like receptor with IG domains instead of LRR domains.

detection of intracellular bacteria, but can also react to extracellular bacteria or their components [94]. How extracellular bacteria are detected is not yet fully understood, but based on this ability it has been hypothesized that NLRs play an important role in distinguishing pathogenic from commensal bacteria [94]. The *A. queenslandica* genome encodes a surprisingly large and diverse repertoire of NLRs with 135 genes possessing the NLR's typical NACHT domain and C-terminal LRRs [95]. Of these, approximately one third have the typical tripartite architecture with an N-terminal DEATH or CARD domain [95]. This unexpected abundance and diversity might indicate an NLR-based immune system response that is sophisticated and highly specialized [96]. In contrary to *A. queenslandica*, *Vaceletia* sp. seems to have a much reduced (if any) NLR-based immune system response. We could find no contig that contained both NACHT and LRR domains, and only three contigs possessed a NACHT domain and a C-terminal DEATH domain but lacked detectable LRRs (Supplementary Table S5).

Toll-Like Receptor Signaling Pathway

The recognition of PAMPs through TLRs results in the initialization of signaling cascades such as the Toll-like receptor signaling pathway. Depending on which adapter molecule (e.g. MyD88 or TRIF-related adapter molecule (TRAM)) is recruited, the activation of the pathway leads to the production of inflammatory cytokines or the induction of type I interferon [89]. In addition to the pathway leading to nuclear localization of NF-κB, the TLR signaling cascade can also activate the Jun N-terminal kinase (JNK) and p38 mitogen-activated protein kinase (MAPK) resulting in the activation of AP1 (activating protein 1). NF-κB and AP1 are both transcription factors, which when activated result in the transcription of genes involved in immune responses. The link between these pathways seems to require the participation of the adaptor protein ECSIT (evolutionary conserved signaling intermediate in Toll pathway) [97], which have also been shown to link the TLR signaling pathway to the TGF-β (transforming growth factor-beta)/BMP (bone morphogenetic protein) pathway [98].

To obtain a more comprehensive picture of what is already known for sponge innate immunity components we surveyed the literature and summarized the present state of knowledge in Fig. 6. All components listed were either obtained by transcriptomic, genomic or experimentally focused studies. Additionally, we mapped the transcriptomes of *Aphrocallistes vastus*, *Ircinia fasciculate*, *Chondrilla nucula*, *Petrosia ficiformis*, *Spongilla*

lacustris, *Pseudospongosorites suberitoides*, *Xestospongia testudinaris*, *Stylissa cateri*, *Sycon coactum* and *Corticum candelabrum* against the KEGG TLR signaling pathway. A recent study of the genome and expressed sequence tags of *A. queenslandica* showed that the majority of the molecules involved in TLR signaling are present [90]. Mapping the *Vaceletia* sp. transcriptome against the KEGG TLR signaling pathway shows that most of the components in the pathway from MyD88 to NF-κB and AP-1 are present, with all mapped contigs being of metazoan origin (Fig. 7). TLR, FADD and MEKK3/6 were not detected by KEGG, however manual BLASTing of these components against the *Vaceletia* sp. transcriptome indicated the presence of contigs with significant similarities to these components (Supplementary Table S6). The other sponge transcriptomes that were mapped to the KEGG TLR signaling pathway show a similar pattern of presences and absence of pathway components (Fig. 6).

As discussed above, canonical TLRs are missing from the *Vaceletia* sp. transcriptome. Wiens and colleagues [91] proposed an alternative pattern recognition system for antimicrobial defense, a cell surface protein named SLIP (sponge lipopolysaccharide (LPS)-interacting protein), which was suggested to be a substitute for the missing TLRs as

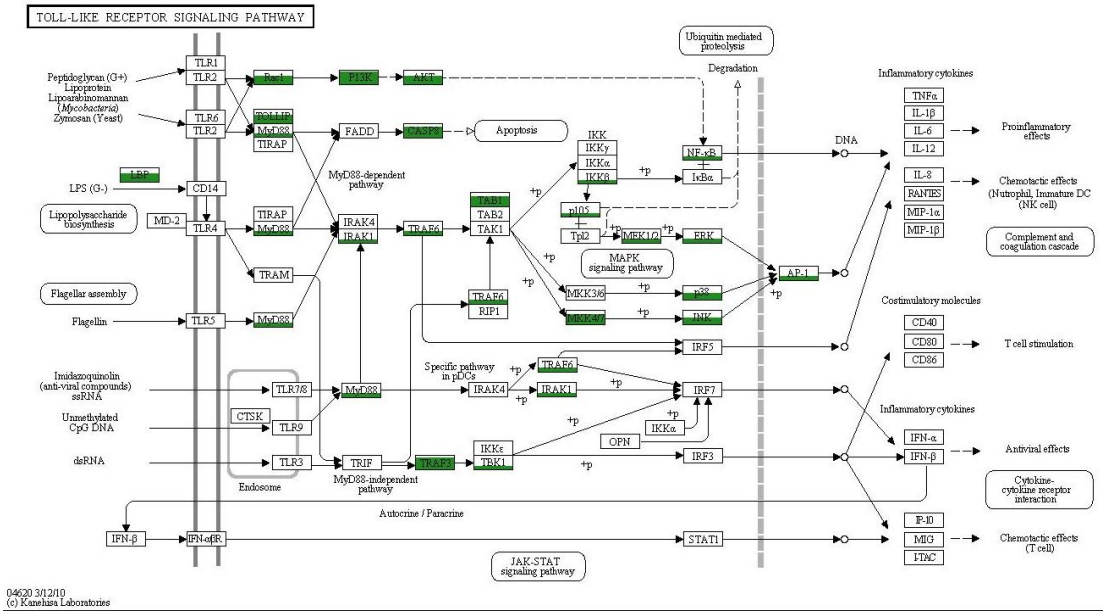

FIGURE 7 | Coverage of the TLR signaling pathway. MEGAN was used to visualize contigs on the KEGG map. Rectangles represent proteins of the pathway while arrows represent signaling routes. Pathway components showing significant similarities to contigs in the Vaceletia sp. transcriptome are highlighted in green. Different fill levels indicate the number of contigs assigned to the corresponding pathway component.

it interacts *in vivo* with MyD88. The *Vaceletia* sp. transcriptome contains contigs that share similarity with SLIP (CAI68017.1), however without experimental validation it is impossible to say if this protein substitutes for the missing TLRs.

Although it is not entirely clear how sponges recognize PAMPs, the presence of a mostly complete TLR signaling pathway in all investigated sponges to date suggests this pathway plays an important role in the sponge immune response. To gain deeper insight into how sponges manage their relationships with bacteria, functional analyses are now needed to dissect the roles that these TLR components play.

2.4.5 Eukaryotic-Like Domains

Recent metagenomic studies on sponges have revealed an abundance of eukaryotic-like proteins in sponge-associated bacteria [17, 55, 57, 58, 68, 99]. Notably, genes encoding proteins containing ankyrin repeats (ANKs) and tetratricopeptide repeats (TPRs) are abundant. These classes of proteins have been found in various facultative and obligate symbionts, as well as in intracellular pathogens and are thought to play a role in host microbe interactions by interfering with eukaryotic protein-protein interactions. For example, recent work provides evidence that ANK proteins of an uncultured gamma-proteobacterial sponge symbiont can interfere with phagocytosis which could provide a mechanism for bacteria to avoid digestion by their host [100]. *Vaceletia* sp. contigs assigned by MEGAN to a bacterial origin were searched for the occurrence of eukaryotic-like proteins. This subset was rich in proteins containing ANK and TPR domains (Supplementary Table S7). Other eukaryotic-like proteins with potential host-bacteria communication functions are also present (Supplementary Table S7). NHL domain (named after NCL-1, HT2A and Lin-41) containing proteins were abundant and are thought to play a role in protein-protein interactions. Fibronectin and Cadherin domains, likely to be involved in cell adhesion [101], were also found in the bacterially derived subset. A recent study demonstrated that sponge symbionts generally possess more eukaryotic-like proteins than bacteria from the surrounding seawater [55]. The presence of a variety of eukaryotic-like proteins within *Vaceletia*'s bacterially derived transcriptome supports that observation. Although evidence is accumulating that eukaryotic-like domains are enriched in sponge endobiotic bacteria, the role of these proteins remains unknown and awaits the development of functional assays.

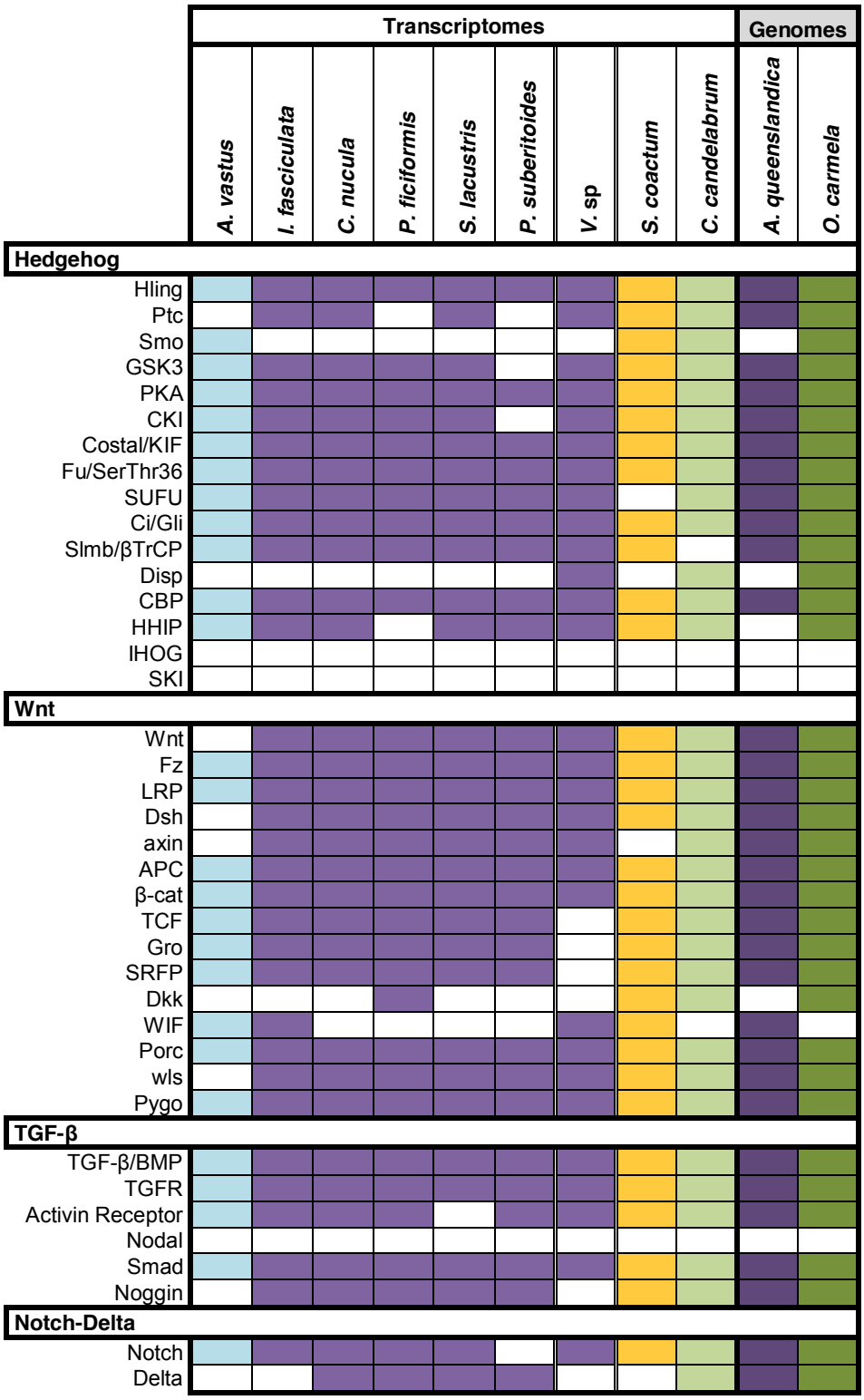

FIGURE 8 | Conserved signaling pathway components identified in *Vaceletia* sp. Presence/absence of genes belonging to the Hedgehog, TGF-β, Wnt, and Notch/Delta pathways in *Vaceletia* sp. compared to the presence of these genes in other sponges recently reported from Riesgo et al. [33]. Different colors represent different sponge class; blue = Hexactinellid; purple = Demosponge; yellow = Calcarea; green = Homoscleromorpha; Colored square = component present; white square = component absent.

2.4.6 Conserved Developmental Signaling Pathways and Homeobox Genes

It is now well recognised that the genetic repertoire of sponges is more complex than their morphological simple bodyplan implies. To check the completeness of the *Vaceletia* sp. transcriptome we surveyed it for the occurrence of important signaling pathway components belonging to the Hedgehog, TGF-β, Wnt, and Notch/Delta pathways that have recently been reported from other sponge transcriptomes [33]. Components were identified using KEGG mapping and/or BLASTx searches. Our results show that almost all signaling components known from other sponges are present in *Vaceletia* sp. (Fig. 8). Some components of the Wnt pathway appear to be missing. Sequences showing similarity to TCF, Gro and SFRP are present but could not be unambiguously identified.

Homeobox genes play crucial roles in developmental processes in all animals. The genome of the demosponge *Amphimedon queenslandica* contains several NK genes linked in a cluster but no Hox or ParaHox genes [102]. Due to the lack of Hox and ParaHox genes in both *Amphimedon* and the ctenophore *Mnemiopsis leidyi* the origin of Hox and ParaHox genes have been proposed to have occurred after the divergences of sponges and ctenophores from all other animals [102, 103]. On the other hand, the lack of Hox and ParaHox genes has been interpreted as gene loss as the genome of *Amphimedon* possesses distinct Hox and ParaHox neighborhoods (so called ghost loci) [104]. The presence of the ParaHox gene Cdx in the calcisponges *Sycon ciliatum* and *Leucosolenia complicata* supports the ghost locus hypothesis and pushes the origin of Hox and ParaHox genes prior to the divergence of sponges from the rest of the animals [105].

We have identified 40 homeobox containing genes in the *Vaceletia* transcriptome (Supplementary Table S8). Of these, eleven show significant similarity to ANPT-, nine to PRD-, six to LIM-, three to POU-, two to SINE- and nine to TALE-class genes. However, this classification of *Vaceletia*'s homeobox genes relies solely on similarity searches and should only be considered an initial classification. Further lines of evidence (e.g. careful phylogenetic analyses) need to be performed before these classifications could be considered complete. The current knowledge of homeobox containing genes in sponges is summarized in Fig. 9. As already described by [105], the ANTP-class gene repertoire of the two calcisponges is very similar but differs notable from the repertoire of the two demosponges. Neither Hox nor ParaHox genes are present in the transcriptome of *Vaceletia*. Both *Amphimedon* and *Vaceletia* contain genes of all homeobox gene classes

but differ in their composition (Fig. 9). No further data (except for the ANTP-class) is currently available for the Calcispongiae. The number of homeobox genes identified in *Vaceletia* also confirms the significantly reduced demosponge homeobox gene content in comparison to eumatazoan lineages [102].

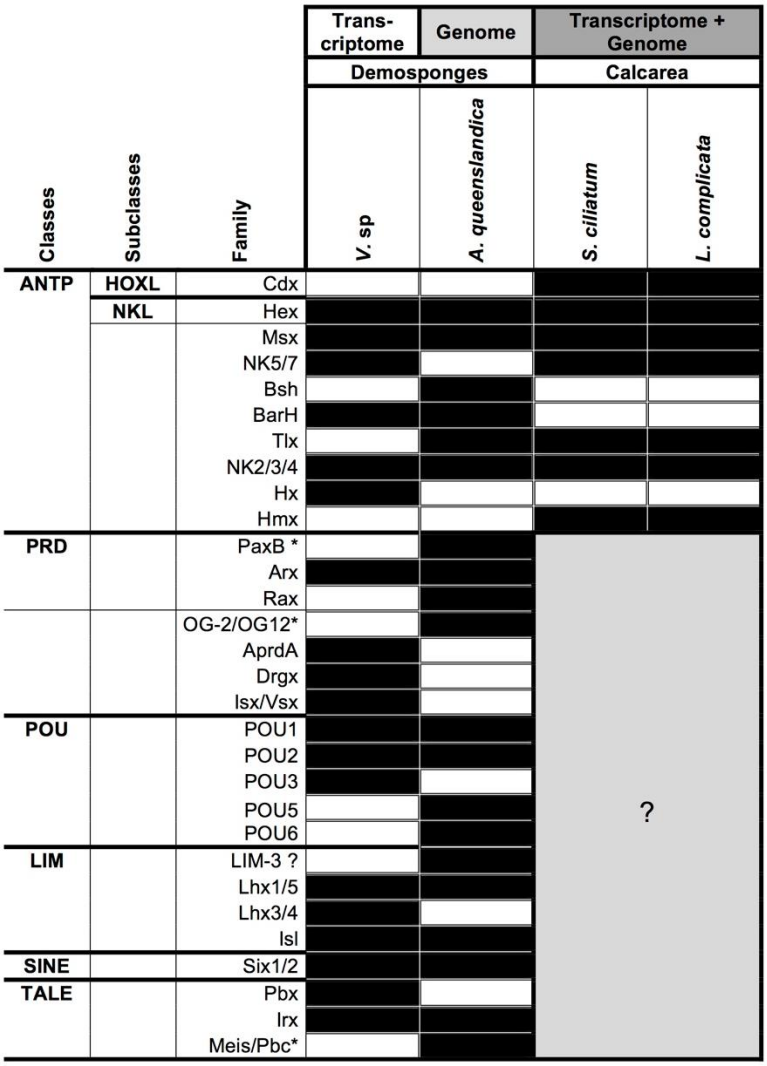

FIGURE 9 | Presence/absence of homeobox genes in sponges Colored square = component present; white square = component absent; grey square = unknown. Data for *A. queenslandica* and the two Calcisponges was taken from recent publications [102, 105]. ANTP = Antennapedia; HOXL = HOX-like; NKL = NK-like; PRD = name derives from the Paired gene of *Drosophila*; POU = name derives from the mammalian Pou1F1, Pou2F1, Pou2F2 and nematode unc-86 genes; LIM = derives from the nematode lin-11, mammalian Isl1 and nematode mec-3 genes; SINE = named after the *Drosophila* gene sine oculis; TALE = Three Amino Acid Loop Extension; *Gene name not present in the Homeobox Database [51, 52].

2.5 Conclusion

In characterizing and comparing the transcriptome of *Vaceletia* sp. to other sponge transcriptomes and genomes we demonstrate that this genus is likely to be interacting with its extensive microbial community in a variety of ways. Our data further highlights the important role the sponge innate immune system is likely to play in managing its microbial community, and the possible role that bacterial eukaryotic-like proteins may also play in this interaction. Metabolic interactions between sponge-associated bacteria and the sponge host are also likely to be more complex than currently appreciated. Our data also highlights the discrepancy between our understanding of how complex these interactions are, and how these interactions are functionally fulfilled. Many other interesting questions remain unanswered such as how is bacterial growth within a host regulated, to what extent are these relationships mutualistic, commensal or parasitic, or how do sponges determine the composition of their bacterial community? As sponge-microbe symbioses are likely to be one of the most ancient within the animal kingdom addressing these questions will assist our understanding of how these relationships can evolve.

Abbreviations

ANK: ankyrin repeats, ANTP: Antennapedia, AP1: activating protein 1, BLAST: Basic Local Alignment Search Tool, BMP: bone morphogenetic protein, Cds: coding sequences, DAMPS: damage-associated molecular patterns, DGGE: Denaturing gradient gel electrophoresis, ECSIT: evolutionary conserved signaling intermediate in Toll pathway, Elovl: elongation of very long chain fatty acids, FA: fatty acid, FAS: fatty acid synthase, GC/MS: gas chromatography/mass spectrometry, HADC: β-hydroxyacyl-CoA dehydratase, HMA: high microbial abundance, HOXL: HOX-like, IG: Immunoglobulin, JNK: Jun N-terminal kinase, KAR: ketoacyl-CoA reductase, KEEG: Kyoto Encyclopedia of Genes and Genomes, LIM: derives from the nematode lin-11, mammalian Isl1 and nematode mec-3 genes, LMA: low microbial abundance, LPS: lipopolysaccharide, LRR: leucine-rich-repeat, MAMP: microbial-associated molecular pattern, MAPK: p38 mitogen-activated protein kinase, MBFA: mid-chain branched fatty acid, My D88: myeloid differentiation primary response gene 88, NKL: NK-like, NLR: Nucleotide-binding domain and Leucine-rich repeat protein, NOD: Nucleotide Oligorimerisation Domain, PAMP: pathogen-associated molecular pattern, PKS: polyketide synthase, POU: name derives from the mammalian Pou1F1, Pou2F1, Pou2F2 and nematode unc-86 genes, PRD:

name derives from the Paired gene of Drosophila, PRR: Pattern Recognition Receptor, SINE: named after the Drosophila gene sine oculis, SLIP: sponge lipopolysaccharide-interacting protein, SMT: sterol-24/28-methyltransferase, Sup-type PKS: sponge symbiont ubiquitous PKS, TALE: Three Amino Acid Loop Extension, TCF: T-cell specific transcription factor, TER: trans-2,3-enoyl-CoA reductase, TGF-β: transforming growth factor- β, TIR: Toll/interleukin-1 receptor, TLR: toll-like receptor, TPR: tetratricopeptide repeats, TRAM: TRIF-related adapter molecule, TRIF: TIR-domain-containing adapter-inducing interferon-β.

Ethics Statement

All experiments described here comply with the current laws and regulations of the country they were conducted in.

Authors' contributions

JG carried out bioinformatic analyses and wrote the manuscript. NC conducted the *de novo* transcriptome assemblies and the BUSCO analysis. DJJ collected the samples for transcriptome sequencing, conceived and supervised the study, and planned and drafted the manuscript. All authors read and approved the final manuscript.

Funding

This research was funded by a grant from the Deutsche Forschungsgemeinschaft to DJ (JA2108/2-1), and from the Open Access Publication Fund provided by the University of Göttingen.

Acknowledgements

The Transcriptome and Genome Analysis Laboratory (University of Göttingen) provided Illumina sequencing services. Samples for transcriptome sequencing were collected during the "Deep DownUnder" expedition (DFG project Wo896/7-1), and we acknowledge Professor Gert Wörheide's (LMU, Munich) role in leading that expedition. Professor Volker Thiel and Dr. Martin Blumenberg are gratefully acknowledged for assistance with the GC/MS work. Samples for GC/MS analysis were collected through work funded by DFG projects Mi 157/10; Re 665/4,8; Th 713/1.

Supplementary material

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmars.2017.00081/full#supplementary-material or via the attached CD.

References

1. Pick, K. S., Philippe, H., Schreiber, F., Erpenbeck, D., Jackson, D. J., Wrede, P. et al. (2010) **Improved phylogenomic taxon sampling noticeably affects nonbilaterian relationships.** *Mol*. *Biol*. *Evol*., 27(9), 1983-1987, doi:10.1093/molbev/msq089.

2. Pisani, D., Pett, W., Dohrmann, M., Feuda, R., Rota-Stabelli, O., Philippe, H. et al. (2015) **Genomic data do not support comb jellies as the sister group to all other animals.** *Proc*. *Natl*. *Acad*. *Sci*. *USA*, 112(50), 15402-15407, doi:10.1073/pnas.1518127112.

3. Yin, Z., Zhu, M., Davidson, E. H., Bottjer, D. J., Zhao, F., Tafforeau, P. (2015) **Sponge grade body fossil with cellular resolution dating 60 Myr before the Cambrian.** *Proc*. *Natl*. *Acad*. *Sci*. *USA*, 112(12), E1453-E1460, doi:10.1073/pnas.1414577112.

4. Germer, J., Mann, K., Wörheide, G., Jackson, D. J. (2015) **The skeleton forming proteome of an early branching metazoan: a molecular survey of the biomineralization components employed by the coralline sponge *Vaceletia* sp.** *PLoS One*, 10(11), e0140100, doi:10.1371/journal.pone.0140100.

5. Jackson, D. J., Macis, L., Reitner, J., Degnan, B. M., Wörheide, G. (2007) **Sponge paleogenomics reveals an ancient role for carbonic anhydrase in skeletogenesis.** *Science*, 316(5833), 1893-1895, doi:10.1126/science.1141560.

6. Jackson, D. J., Thiel, V., Wörheide, G. (2010) **An evolutionary fast-track to biocalcification.** *Geobiology*, 8(3), 191-196, doi:10.1111/j.1472-4669.2010.00236.x.

7. Jackson, D. J., Macis, L., Reitner, J., Wörheide, G. (2011) **A horizontal gene transfer supported the evolution of an early metazoan biomineralization strategy.** *BMC Evol*. *Biol*., 11, 238, doi:10.1186/1471-2148-11-238.

8. Fernandez-Valverde, S. L., Degnan, B. M. (2016) **Bilaterian-like promoters in the highly compact *Amphimedon queenslandica* genome.** *Sci*. *Rep*., 6, 22496-22496, doi:10.1038/srep22496.

9. Levin, M., Anavy, L., Cole, A. G., Winter, E., Mostov, N., Khair, S. et al. (2016) **The mid-developmental transition and the evolution of animal body plans.** *Nature*, 531(7596), 637-641, doi:10.1038/nature16994.

10. Nakanishi, N., Sogabe, S., Degnan, B. M. (2014) **Evolutionary origin of gastrulation: insights from sponge development.** *BMC Biology*, 12(1), 26, doi:10.1186/1741-7007-12-26.

11. Taylor, M. W., Radax, R., Steger, D., Wagner, M. (2007) **Sponge-Associated microorganisms: evolution, ecology, and biotechnological potential.** *Microbiol. Mol. Biol. Rev.*, 71(2), 295-347, doi:10.1128/MMBR.00040-06.

12. Hentschel, U., Piel, J., Degnan, S. M., Taylor, M. W. (2012) **Genomic insights into the marine sponge microbiome.** *Nat. Rev. Microbiol.*, 10(9), 641-654, doi:10.1038/nrmicro2839.

13. Reitner, J., Wörheide, G., Lange, R., Thiel, V. (1997) **Biomineralization of calcified skeletons in three Pacific coralline demosponges - an approach to the evolution of basal skeletons.** *Cour. Forsch-Inst. Senckenberg* 201, 371-383.

14. Webster, N. S., Thomas, T. (2016) **The sponge hologenome.** *mBio*, 7(2), e00135-16, doi:10.1128/mBio.00135-16.

15. Hentschel, U., Hopke, J., Horn, M., Friedrich, A. B., Wagner, M., Hacker, J. et al. (2002) **Molecular evidence for a uniform microbial community in sponges from different oceans.** *Appl. Environ. Microbiol.*, 68(9), 4431-4440, doi:10.1128/AEM.68.9.4431-4440.2002.

16. Simister, R. L., Deines, P., Botté, E. S., Webster, N. S., Taylor, M. W. (2012) **Sponge-specific clusters revisited: a comprehensive phylogeny of sponge-associated microorganisms.** *Eviron. Microbiol.*, 14(2), 517-524, doi:10.1111/j.1462-2920.2011.02664.x.

17. 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(12), 1557-1567, doi:10.1038/ismej.2010.74.

18. Fiore, C. L., Labrie, M., Jarett, J. K., Lesser, M. P. (2015) **Transcriptional activity of the giant barrel sponge, *Xestospongia muta* Holobiont: molecular evidence for metabolic interchange.** *Front. Microbiol.*, 6, 364, doi:10.3389/fmicb.2015.00364.

19. Piel, J., Hui, D., Wen, G., Butzke, D., Platzer, M., Fusetani, N. et al. (2004) **Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge *Theonella swinhoei*.** *Proc. Natl. Acad. Sci. USA*, 101(46), 16222-16227, doi:10.1073/pnas.0405976101.

20. Hochmuth, T., Niederkrüger, H., Gernert, C., Siegl, A., Taudien, S., Platzer, M. et al. (2010) **Linking chemical and microbial diversity in marine sponges: possible role for Poribacteria as producers of methyl-branched fatty acids.** *ChemBioChem*, 11(18), 2572-2578, doi:10.1002/cbic.201000510.

21. Wilson, M. C., Mori, T., Rückert, C., Uria, A. R., Helf, M. J., Takada, K. et al. (2014) **An environmental bacterial taxon with a large and distinct metabolic repertoire.** *Nature*, 506(7486), 58-62, doi:10.1038/nature12959.

22. Pawlik, J. R. (2011) **The chemical ecology of sponges on caribbean reefs: natural products shape natural systems.** *BioScience*, 61(11), 888-898, doi:10.1525/bio.2011.61.11.8.

23. Wilkinson, C. R., Garrone, R., Vacelet, J. (1984) **Marine sponges discriminate between food bacteria and bacterial symbionts: electron microscope radioautography and *in situ* evidence.** *Proc. R. Soc. Lond. B. Biol. Sci.*, 220(1221), 528-519, doi:10.1098/rspb.1984.0018.

24. Müller, W., E. G., Müller, I. M. (2003) **Origin of the metazoan immune system: identification of the molecules and their functions in sponges.** *Integr. Comp. Biol.*, 43(2), 292-281, doi:10.1093/icb/43.2.281.

25. Vacelet, J. (1977) **Une nouvelle relique du Secondaire: un représentant actuel des Eponges fossiles Sphinctozoaires.** *Comptes Rendus De L'Academie Des Sciences Paris* (série D) 285, 509-511.

26. Wörheide, G. (2008) **A hypercalcified sponge with soft relatives: *Vaceletia* is a keratose demosponge.** *Mol. Phylogenet. Evol.*, 47(1), 433-438, doi:10.1016/j.ympev.2008.01.021.

27. Erpenbeck, D., Voigt, O., Wörheide, G., Lavrov, D. V. (2009) **The mitochondrial genomes of sponges provide evidence for multiple invasions by Repetitive Hairpin-forming Elements (RHE).** *BMC Genomics*, 10, 591, doi:10.1186/1471-2164-10-591.

28. Morrow, C., Cárdenas, P. (2015) **Proposal for a revised classification of the Demospongiae (Porifera).** *Front. Zool.*, 12(1), 7, doi:10.1186/s12983-015-0099-8.

29. Reitner, J. (1992) **Coralline Spongien: der Versuch einer phylogenetisch-taxonomischen Analyse.** *Berliner Geowisschenschaftliche Abhandlungen* (Reihe E) 1, 1-352.

30. Wörheide, G. (1998) **The reef cave dwelling ultraconservative coralline demosponge *Astrosclera willeyana* from the Indo-Pacific.** *Facies*, 38, 1-88, doi:10.1007/BF02537358.

31. Karlińska-Batres, K., Wörheide, G. (2013) **Microbial diversity in the coralline sponge *Vaceletia crypta*.** *Antonie Van Leeuwenhoek*, 103(5), 1041-1056, doi:10.1007/s10482-013-9884-6.

32. Jackson, D. J., Wörheide, G. (2014) **Symbiophagy and biomineralization in the „living fossil" *Astrosclera willeyana*.** *Autophagy*, 10(3), 408-415, doi:10.4161/auto.27319.

33. Riesgo, A., Farrar, N., Windsor, P. J., Giribet, G., Leys, S. P. (2014) **The analysis of eight transcriptomes from all poriferan classes reveals surprising genetic complexity in sponges.** *Mol. Biol. Evol.*, 31(5), 1102-1120, doi:10.1093/molbev/msu057.

34. Andrews, S. (2014) **FastQC: A quality control tool for high throughput sequence data. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.**

35. Bolger, A. M., Lohse, M., Usadel, B. (2014) **Trimmomatic: A flexible trimmer for Illumina Sequence Data.** *Bioinformatics*, 30(15), 2114-2120, doi:10.1093/bioinformatics/btu170.

36. Cerveau, N., Jackson, D. J. (2016) **Combining independent *de novo* assemblies optimizes the coding transcriptome for nonconventional model eukaryotic organisms.** *BMC Bioinformatics*, 17(1), 525, doi:10.1186/s12859-016-1406-x.

37. Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J. et al. (2013) ***De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis.** *Nat. Protoc.,* 8(8), 1494-1512, doi:10.1038/nprot.2013.084.

38. **https://www.qiagenbioinformatics.com**

39. Peng, Y., Leung, H. C. M., Yiu, S.-M., Lv, M.-J., Zhu, X.-G., Chin, F. Y. L. (2013) **IDBA-tran: a more robust de novo de Bruijn graph assembler for transcriptomes with uneven expression levels.** *Bioinformatics*, 29(13), i326-i334, doi:10.1093/bioinformatics/btt219.

40. Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R. et al. (2014) **Pfam: the protein families database.** *Nucleic Acids Res.*, 42(D1), D222-D230, doi:10.1093/nar/gkt1223.

41. Finn, R. D., Clements, J., Eddy, S. R. (2011) **HMMER web server: interactive sequence similarity searching.** *Nucleic Acids Res.*, 39(suppl), W29-W37, doi:10.1093/nar/gkr367.

42. Petersen, T. N., Brunak, S., von, H., Gunnar, Nielsen, H. (2011) **SignalP 4.0: discriminating signal peptides from transmembrane regions.** *Nature Methods*, 8(10), 785-786, doi:10.1038/nmeth.1701.

43. Krogh, A., Larsson, B., von Heijne, G., Sonnhammer, E. L. L. (2001) **Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes.** *J. Mol. Biol.*, 305(3), 567-580, doi:10.1006/jmbi.2000.4315.

44. Huson, D. H., Mitra, S., Ruscheweyh, H. J., Weber, N., Schuster, S. C. (2011) **Integrative analysis of environmental sequences using MEGAN4.** *Genome Res.*, 21(9), 1552-1560, doi:10.1101/gr.120618.111.

45. Ziemert, N., Podell, S., Penn, K., Badger, J. H., Allen, E., Jensen, P. R. (2012) **The Natural Product Domain Seeker NaPDoS: A Phylogeny Based Bioinformatic Tool to Classify Secondary Metabolite Gene Diversity.** *PLoS One*, 7(3), e34064, doi:10.1371/journal.pone.0034064.s007.

46. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W. et al. (2011) **Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.** *Mol Syst Biol*, 7, 539, doi:10.1038/msb.2011.75.

47. Castresana, J. (2000) **Selection of Conserved Blocks from Multiple Alignments for Their Use in Phylogenetic Analysis.** *Mol. Biol. Evol.*, 17(4), 540-552, doi:10.1093/oxfordjournals.molbev.a026334.

48. Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Hohna, S. et al. (2012) **MrBayes 3.2: Efficient Bayesian Phylogenetic Inference and Model Choice Across a Large Model Space.** *Syst. Biol.*, 61(3), 539-542, doi:10.1093/sysbio/sys029.

49. Stamatakis, A., Hoover, P., Rougemont, J. (2008) **A Rapid Bootstrap Algorithm for the RAxML Web Servers.** *Syst. Biol.*, 57(5), 758-771, doi:10.1080/10635150802429642.

50. Gold, D. A., Grabenstatter, J., de Mendoza, A., Riesgo, A., Ruiz-Trillo, I., Summons, R. E. (2016) **Sterol and genomic analyses validate the sponge biomarker hypothesis.** *Proc. Natl. Acad. Sci. USA*, 113(10), 2684-2689, doi:10.1073/pnas.1512614113.

51. Zhong, Y.-F., Butts, T., Holland, P. W. H. (2008) **HomeoDB: a database of homeobox gene diversity.** *Evol. Dev.*, 10(5), 516-518, doi:10.1111/j.1525-142X.2008.00266.x.

52. Zhong, Y.-F., Holland, P. W. H. (2011) **HomeoDB2: functional expansion of a comparative homeobox gene database for evolutionary developmental biology.** *Evol. Dev.*, 13(6), 567-568, doi:10.1111/j.1525-142X.2011.00513.x.

53. Simão, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., Zdobnov, E. M. (2015) **BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs.** *Bioinformatics*, 31(19), 3210-3212, doi:10.1093/bioinformatics/btv351.

54. Ryu, T., Seridi, L., Moitinho-Silva, L., Oates, M., Liew, Y. J., Mavromatis, C. et al. (2016) **Hologenome analysis of two marine sponges with different microbiomes.** *BMC Genomics*, 17(1), 158, doi:10.1186/s12864-016-2501-0.

55. Fan, L., Reynolds, D., Liu, M., Stark, M., Kjelleberg, S., Webster, N. S. et al. (2012) **Functional equivalence and evolutionary convergence in complex communities of microbial sponge symbionts.** *Proc. Natl. Acad. Sci. USA*, 109(27), E1878-E1887, doi:10.1073/pnas.1203287109.

56. Kennedy, J., Marchesi, J. R., Dobson, A. D. W. (2007) **Metagenomic approaches to exploit the biotechnological potential of the microbial consortia of marine sponges.** *Appl*. *Microbiol*. *Biotechnol*., 75(1), 11-20, doi:10.1007/s00253-007-0875-2.

57. Li, Z. Y., Wang, Y. Z., He, L. M., Zheng, H. J. (2014) **Metabolic profiles of prokaryotic and eukaryotic communities in deep-sea sponge *Neamphius huxleyi* indicated by metagenomics.** *Sci*. *Rep*., 4, 3895, doi:10.1038/srep03895.

58. Liu, M., Fan, L., Zhong, L., Kjelleberg, S., Thomas, T. (2012) **Metaproteogenomic analysis of a community of sponge symbionts.** *ISME J*., 6(8), 1515-1525, doi:10.1038/ismej.2012.1.

59. Radax, R., Rattei, T., Lanzen, A., Bayer, C., Rapp, H. T., Urich, T. et al. (2012) **Metatranscriptomics of the marine sponge *Geodia barretti*: tackling phylogeny and function of its microbial community.** *Eviron*. *Microbiol*., 14(5), 1308-1324, doi:10.1111/j.1462-2920.2012.02714.x.

60. Gillan, F. T., Stoilov, I. L., Thompson, J. E., Hogg, R. W., Wilkinson, C. R., Djerassi, C. (1988) **Fatty acids as biological markers for bacterial symbionts in sponges.** *Lipids*, 23(12), 1139-1145, doi:10.1007/BF02535280.

61. Zimmerman, M. P., Thomas, F. C., Thompson, J. E., Djerassi, C., Streiner, H., Evans, E. et al. (1989) **The distribution of lipids and sterols in cell types from the marine sponge *Pseudaxinyssa* sp.** *Lipids*, 24(3), 210-216, doi:10.1007/BF02535236.

62. Zimmerman, M. P., Hoberg, M., Ayanoglu, E., Djerassi, C. (1990) **Cell separation of *Tethya aurantia*, an analytical study of embryonic and differentiated sponge cells.** *Lipids*, 25(7), 383-390, doi:10.1007/BF02537981.

63. Lu, Y.-J., Zhang, Y.-M., Rock, C. O. (2004) **Product diversity and regulation of type II fatty acid synthases.** *Biochem*. *Cell Biol*., 82(1), 145-155, doi:10.1139/o03-076.

64. Srivastava, M., Simakov, O., Chapman, J., Fahey, B., Gauthier, M. E. A., Mitros, T. et al. (2010) **The *Amphimedon queenslandica* genome and the evolution of animal complexity.** *Nature*, 466(7307), 720-726, doi:10.1038/nature09201.

65. Jenke-Kodama, H. (2005) **Evolutionary Implications of Bacterial Polyketide Synthases.** *Mol*. *Biol*. *Evol*., 22(10), 2027-2039, doi:10.1093/molbev/msi193.

66. Thiel, V., Jenisch, A., Wörheide, G., Löwenberg, A., Reitner, J., Michaelis, W. (1999) **Mid-chain branched alkanoic acids from „living fossil" demosponges: a link to ancient sedimentary lipids?** *Org*. *Geochem*., 30(1), 1-14, doi:10.1016/S0146-6380(98)00200-9.

67. Fieseler, L., Hentschel, U., Grozdanov, L., Schirmer, A., Wen, G., Platzer, M. et al. (2007) **widespread occurrence and genomic context of unusually small polyketide**

synthase genes in microbial consortia associated with marine sponges. *Appl*. *Environ*. *Microbiol*., 73(7), 2144-2155, doi:10.1128/AEM.02260-06.

68. Siegl, A., Kamke, J., Hochmuth, T., Piel, J. O. R., Richter, M., Liang, C. et al. (2010) **Single-cell genomics reveals the lifestyle of *Poribacteria*, a candidate phylum symbiotically associated with marine sponges.** *ISME J*., 5(1), 61-70, doi:10.1038/ismej.2010.95.

69. Cox, J. S., Chen, B., McNeil, M., Jacobs, W. R. (1999) **Complex lipid determines tissue-specific replication of *Mycobacterium tuberculosis* in mice.** *Nature*, 402(6757), 79-83, doi:10.1038/47042.

70. Kornprobst, J.-M., Barnathan, G. (2010) **Demospongic acids revisited.** *Mar*. *Drugs*, 8(10), 2569-2577, doi:10.3390/md8102569.

71. Litchfield, C., Greenberg, A. J., Noto, G., Morales, R. W. (1976) **Unusually high levels of C24 - C30 fatty acids in sponges of the class demospongiae.** *Lipids*, 11(7), 567-570, doi:10.1007/BF02532903.

72. Djerassi, C., Lam, W. K. (1991) **Phospholipid studies of marine organisms. Part 25. Sponge phospholipids.** *Accounts Chem*. *Res*., 24(3), 69-75, doi:10.1021/ar00003a002.

73. Hahn, S., Stoilov, I. L., Ha, T. B. T., Raederstorff, D., Doss, G. A., Li, H. T. et al. (1988) **Biosynthetic studies of marine lipids. 17. The course of chain elongation and desaturation in long-chain fatty acids of marine sponges.** *J*. *Am*. *Chem*. *Soc*., 110(24), 8117-8124, doi:10.1021/ja00232a025.

74. Morales, R. W., Litchfield, C. (1977) **Incorporation of 1-14C-Acetate into C26 fatty acids of the marine sponge *Microciona prolifera*.** *Lipids*, 12(7), 570-576, doi:10.1007/BF02533383.

75. Monroig, Ó., Tocher, D., Navarro, J. (2013) **Biosynthesis of polyunsaturated fatty acids in marine invertebrates: recent advances in molecular mechanisms.** *Mar*. *Drugs*, 11(10), 3998-4018, doi:10.3390/md11103998.

76. Leonard, A. E., Pereira, S. L., Sprecher, H., Huang, Y.-S. (2004) **Elongation of long-chain fatty acids.** *Prog*. *Lipid Res*., 43(1), 36-54, doi:10.1016/S0163-7827(03)00040-7.

77. Castro, L. F. C., Wilson, J. M., Gonçalves, O., Galante-Oliveira, S., Rocha, E., Cunha, I. (2011) **The evolutionary history of the stearoyl-CoA desaturase gene family in vertebrates.** *BMC Evol*. *Biol*., 11(1), 132, doi:10.1186/1471-2148-11-132.

78. Djerassi, C., Silva, C. J. (1991) **Biosynthetic studies of marine lipids. 41. Sponge sterols: Origin and biosynthesis.** *Accounts Chem*. *Res*., 24(12), 371-378, doi:10.1021/ar00012a003.

79. Bergquist, P. R., Karuso, P., Cambie, R. C., Smith, D. J. (1991) **Sterol composition and classification of the Porifera.** *Biochem*. *Syst*. *Ecol*., 19(1), 17-24, doi:10.1016/0305-1978(91)90109-D.

80. Goad, L. J. (1981) **Sterol biosynthesis and metabolism in marine invertebrates.** *Pure Appl*. *Chem*., 53(4), 837-852, doi:10.1351/pac198153040837.

81. Bird, C. W., Lynch, J. M., Pirt, F. J., Reid, W. W., Brooks, C. J. W., Middleditch, B. S. (1971) **Steroids and squalene in *Methylococcus capsulatus* grown on methane.** *Nature*, 230(5294), 473-474, doi:10.1038/230473a0.

82. Bode, H. B., Zeggel, B., Silakowski, B., Wenzel, S. C., Reichenbach, H., Möller, R. (2003) **Steroid biosynthesis in prokaryotes: identification of myxobacterial steroids and cloning of the first bacterial 2, 3 (S)-oxidosqualene cyclase from the myxobacterium *Stigmatella aurantiaca*.** *Mol*. *Microbiol*., 47(2), 471-481, doi:10.1046/j.1365-2958.2003.03309.x.

83. Desmond, E., Gribaldo, S. (2009) **phylogenomics of sterol synthesis: insights into the origin, evolution, and diversity of a key eukaryotic feature.** *Genome Biol*. *Evol*., 1(0), 364-381, doi:10.1093/gbe/evp036.

84. McCaffrey, M. A., Moldowan, J. M., Lipton, P. A., Summons, R. E., Peters, K. E., Jeganathan, A. et al. (1994) **Paleoenvironmental implications of novel C 30 steranes in Precambrian to Cenozoic age petroleum and bitumen.** *Geochim*. *Cosmochim*. *Ac*., 58(1), 529-532, doi:10.1016/0016-7037(94)90481-2.

85. Love, G. D., Grosjean, E., Stalvies, C., Fike, D. A., Grotzinger, J. P., Bradley, A. S. et al. (2008) **Fossil steroids record the appearance of Demospongiae during the Cryogenian period.** *Nature*, 457(7230), 718-721, doi:10.1038/nature07673.

86. Antcliffe, J. B. (2013) **Questioning the evidence of organic compounds called sponge biomarkers.** *Palaeontology*, 56(5), 917-925, doi:10.1111/pala.12030.

87. Gordon, S. (2002) **Pattern Recognition Receptors.** *Cell*, 111(7), 927-930, doi:10.1016/S0092-8674(02)01201-1.

88. Janeway, C. A., Medzhitov, R. (2002) **Innate immune recognition.** *Annu*. *Rev*. *Immunol*., 20(1), 197-216, doi:10.1146/annurev.immunol.20.083001.084359.

89. Akira, S., Uematsu, S., Takeuchi, O. (2006) **Pathogen recognition and innate immunity.** *Cell*, 124(4), 783-801, doi:10.1016/j.cell.2006.02.015.

90. Gauthier, M. E. A., Du Pasquier, L., Degnan, B. M. (2010) **The genome of the sponge *Amphimedon queenslandica* provides new perspectives into the origin of Toll-like and interleukin 1 receptor pathways.** *Evol*. *Dev*., 12(5), 519-533, doi:10.1111/j.1525-142X.2010.00436.x.

91. Wiens, M., Korzhev, M., Krasko, A., Thakur, N. L., Perovi -Ottstadt, S., Breter, H. J. et al. (2005) **Innate immune defense of the sponge *Suberites domuncula* against bacteria involves a MyD88-dependent signaling pathway induction of a perforin-like molecule.** *J*. *Biol*. *Chem*., 280(30), 27949-27959, doi:10.1074/jbc.M504049200.

92. Franchi, L., Eigenbrod, T., Muñoz-Planillo, R., Nunez, G. (2009) **The inflammasome: a caspase-1-activation platform that regulates immune responses and disease pathogenesis.** *Nat*. *Immunol*., 10(3), 241-247, doi:10.1038/ni.1703.

93. Stuart, L. M., Paquette, N., Boyer, L. (2013) **Effector-triggered versus pattern-triggered immunity: how animals sense pathogens.** *Nat*. *Rev*. *Immunol*., 13(3), 199-206, doi:10.1038/nri3398.

94. Ferrand, J., Ferrero, R. L. (2013) **Recognition of Extracellular Bacteria by NLRs and Its Role in the Development of Adaptive Immunity.** *Front*. *Immunol*., 4, 344, doi:10.3389/fimmu.2013.00344.

95. Yuen, B., Bayes, J. M., Degnan, S. M. (2014) **The characterization of sponge NLRs provides insight into the origin and evolution of this innate immune gene family in animals.** *Mol*. *Biol*. *Evol*., 31(1), 106-120, doi:10.1093/molbev/mst174.

96. Degnan, S. M. (2015) **The surprisingly complex immune gene repertoire of a simple sponge, exemplified by the NLR genes: A capacity for specificity?** *Dev*. *Comp*. *Immunol*., 506(7486), 58-62, doi:10.1016/j.dci.2014.07.012.

97. Kopp, E., Medzhitov, R., Carothers, J., Xiao, C., Douglas, I., Janeway, C. A. et al. (1999) **ECSIT is an evolutionarily conserved intermediate in the Toll/IL-1 signal transduction pathway.** *Genes Dev*., 13, 2059-2071, doi:10.1101/gad.13.16.2059.

98. Xiao, C., Shim, J.-h., Klüppel, M., Zhang, S. S.-M., Dong, C., Flavell, R. A. et al. (2003) **Ecsit is required for Bmp signaling and mesoderm formation during mouse embryogenesis.** *Genes Dev*., 17(23), 2933-2949, doi:10.1101/gad.1145603.

99. 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(1), e87353, doi:10.1371/journal.pone.0087353.

100. Nguyen, M. T., Liu, M., Thomas, T. (2014) **Ankyrin-repeat proteins from sponge symbionts modulate amoebal phagocytosis.** *Mol Ecol*, 23(6), 1635-1645, doi:10.1111/mec.12384.

101. Hoffmann, C., Ohlsen, K., Hauck, C. R. (2011) **Integrin-mediated uptake of fibronectin-binding bacteria.** *Eur*. *J*. *Cell Biol*., 90(11), 891-896, doi:10.1016/j.ejcb.2011.03.001.

102. Larroux, C., Fahey, B., Degnan, S. M., Adamski, M., Rokhsar, D. S., Degnan, B. M. (2007) **The NK homeobox gene cluster predates the origin of Hox genes.** *Curr*. *Biol*., 17(8), 706-710, doi:10.1016/j.cub.2007.03.008.

103. Ryan, J. F., Pang, K., Mullikin, J. C., Martindale, M. Q., Baxevanis, A. D. (2010) **The homeodomain complement of the ctenophore *Mnemiopsis leidyi* suggests that Ctenophora and Porifera diverged prior to the ParaHoxozoa.** *Evol*. *Dev*., 1(1), 9, doi:10.1186/2041-9139-1-9.

104. Ramos, O. M., Barker, D., Ferrier, D. E. K. (2012) **Ghost loci imply Hox and ParaHox existence in the last common ancestor of animals.** *Curr*. *Biol*., 22(20), 1951-1956, doi:10.1016/j.cub.2012.08.023.

105. Fortunato, S. A. V., Adamski, M., Ramos, O. M., Leininger, S., Liu, J., Ferrier, D. E. K. et al. (2014) **Calcisponges have a ParaHox gene and dynamic expression of dispersed NK homeobox genes.** *Nature*, 514(7524), 620-623, doi:10.1038/nature13881.

106. Wiens, M., Korzhev, M., Perovic-Ottstadt, S., Luthringer, B., Brandt, D., Klein, S. et al. (2007) **Toll-like receptors are part of the innate immune defense system of sponges (Demospongiae: Porifera).** *Mol*. *Biol*. *Evol*., 24(3), 792-804, doi:10.1093/molbev/msl208.

Chapter 3:

Full title:

The Skeleton Forming Proteome of an Early Branching Metazoan: A Molecular Survey of the Biomineralization Components Employed by the Coralline Sponge *Vaceletia* sp.

Short title:

The Skeleton Forming Proteome of a Sponge

Juliane Germer*, Karlheinz Mann*, Gert Wörheide, Daniel J. Jackson

*These authors contributed equally to this work

correspondence: djackso@uni-goettingen.de

PLoS ONE (2016), 10(11)

DOI: 10.1371/journal.pone.0140100, Open Access

3.1 Abstract

The ability to construct a mineralized skeleton was a major innovation for the Metazoa during their evolution in the late Precambrian/early Cambrian. Porifera (sponges) hold an informative position for efforts aimed at unravelling the origins of this ability because they are widely regarded to be the earliest branching metazoans, and are among the first multi-cellular animals to display the ability to biomineralize in the fossil record. Very few biomineralization associated proteins have been identified in sponges so far, with no transcriptome or proteome scale surveys yet available. In order to understand what genetic repertoire may have been present in the last common ancestor of the Metazoa (LCAM) that may have contributed to the evolution of the ability to biocalcify we have studied the skeletal proteome of the coralline demosponge *Vaceletia* sp. and compare this to other metazoan biomineralizing proteomes. We bring some spatial resolution to this analysis by dividing *Vaceletia's* aragonitic calcium carbonate skeleton into "head" and "stalk" regions. With our approach we were able to identify 40 proteins from both the head and stalk regions with many of these sharing some similarity to previously identified gene products from other organisms. Among these proteins are known biomineralization compounds, such as carbonic anhydrase, spherulin, extracellular matrix proteins and very acidic proteins. This report provides the first proteome scale analysis of a calcified poriferan skeletal proteome, and its composition clearly demonstrates that the LCAM contributed several key enzymes and matrix proteins to its descendants that supported the metazoan ability to biocalcify. However lineage specific evolution is also likely to have contributed significantly to the ability of disparate metazoan lineages to biocalcify.

3.2 Introduction

Biomineralization is a phenomenon that can be found throughout the tree of life. Its appearance in the metazoan (animal) fossil record coincides with a rapid increase in their morphological diversity, suggesting that the evolution of this ability was one key factor that supported the Cambrian explosion (~540 mya). Much effort has therefore been aimed at elucidating the genetic and molecular mechanisms that underlie the ability to biomineralize in disparate animal phyla. It has been proposed that the metazoan ability to build mineral skeletal elements evolved at least twenty times independently [1]. However, this estimate makes assumptions regarding the morphological homology of skeletal elements in disparate taxa, and assumes simplistic models of evolutionary gain/loss of mineralized elements while disregarding the underlying molecular mechanisms that fabricate these structures. Available skeletal proteome datasets from metazoans such as molluscs [2–4], sea urchins [5, 6] and brachiopods [7, 8] go some way towards addressing this issue, but to study the origins of metazoan biomineralization it is crucial to investigate the biomineralizing proteome of an early branching metazoan.

Sponges (Phylum Porifera) have traditionally been considered to be the earliest branching surviving metazoan lineage (reviewed in [9, 10]). However, resolving deep metazoan relationships, especially those among the non-bilaterian taxa Porifera, Ctenophora, Cnidaria and Placozoa, is still a challenging task (see [11]) and the branching order close to the root of the animal Tree of Life is not unequivocally accepted. Recent studies using molecular phylogenetic analyses, transcriptomic and genomic data either confirm [12–14] or reject [15–18]) the view of sponges as the sister group to all other animals. Resolving the phylogeny of the non-bilaterian phyla is crucial to understand the evolution of metazoan traits such as epithelia, nerves and muscles, as well as biomineralization.

Despite the ongoing discussions about placement in the metazoan tree, sponges are among the first multi-cellular eukaryotes (animals) represented in the fossil record to display a “biologically controlled” mode of biomineralization [1]. During the Tommotian Age (beginning 530 MYA) the Archaeocyatha, an assemblage of organisms which most authorities now agree were an extinct class of sponges [19], began to leave evidence in the fossil record of a mode of heavy calcification that is poorly represented among living sponges. As the planet’s first metazoan reef builders, Archaeocyathids were ecologically important, globally distributed, and were taxonomically diverse with hundreds of recognised

species [20–22]. The Archaeocyathids have been extinct since the Cambrian, however superficial similarities in some skeletal features of these ancient animals have been described from a single living ('sphinctozoan'-like) taxon *Vaceletia*, which first appears in the Middle Triassic [23]. Based on these superficial morphological similarities some authors have argued that *Vaceletia* may be a modern Archaeocyath [23, 24]. However, this is very likely not the case since molecular data has shown that the genus *Vaceletia* belongs to the Dictyoceratida within the Class Demospongiae [25–27]. Nonetheless, taxon *Vaceletia* represents an early branching metazoan with a possible ancient mode of biomineralization and this makes it an ideal candidate to deepen our understanding of how the ability to biomineralize may have first arisen in sponges. Taxon *Vaceletia* has been regarded as a monospecific genus with the single type species *Vaceletia crypta* [28]. However, several different growth forms have been discovered in the last decades (see discussion in [25]) and their taxonomy is not fully resolved yet. In this present study we obtained data from a yet to be described likely new colonial-branching *Vaceletia* species from Osprey Reef [see 25, 29].

At length scales ranging from the cm to the nm, *Vaceletia* sp. exhibits exquisite biological control over the formation of its $CaCO_3$ skeleton. The overall structure comprises a series of chambers terraced one on top of the next (Fig. 1). The skeleton is aragonitic $CaCO_3$, with some features of the process of skeleton formation previously described [23, 30, 31]. Briefly, an organic framework is first constructed which is thought to be successively substituted by crystalline aragonite. The organic framework consists of proteins and polysaccharides rich in galactose, glucose and fucose, the latter suggesting that bacterial EPS (exopolymeric substances) may be involved in the biocalcification process. This is likely given that the bacterial biomass of an individual *Vaceletia crypta* can be as high as 50% [23]. Related to this observation we have recently demonstrated that another coralline sponge directly employs its' bacterial community in a biomineralization role [32, 33]. Furthermore Uriz et al [34] suggest that microbial endosymbionts are directly responsible for the precipitation of the calcium carbonate skeleton in the sponge genus *Hemimycale*. Despite this previous work, little is known about the molecular basis of sponge biomineralization in general. The lack of information from this phylogenetically informative group motivated us to address this problem using *Vaceletia* sp. as a model and transcriptomic and mass spectrometry-based proteomics as tools to address the problem. This approach has allowed us to generate a dataset representing what is likely to be the majority of the *Vaceletia* sp. skeleton forming proteome. To our knowledge this is the first such proteome reported for a sponge.

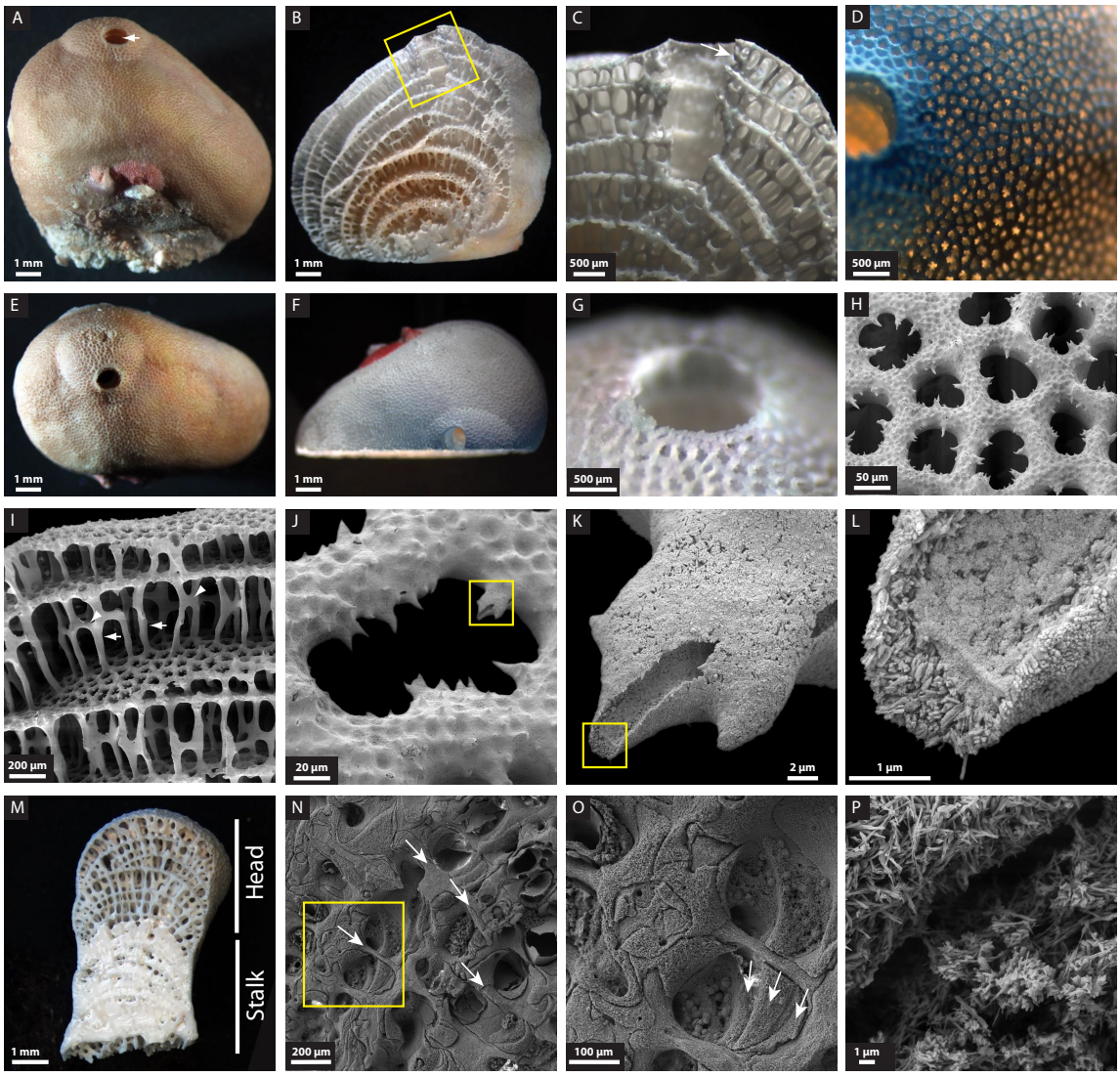

Fig. 1. General morphological features of *Vaceletia* sp. and its $CaCO_3$ skeleton. (A) A lateral view of a fixed animal. The exhalent siphon (arrow) is clearly visible. (B) A sagittal section view after treatment with NaOCl and grinding to reveal the interior structure of the skeleton. (C) Magnification of the boxed section in B illustrates the structure of the siphon and pillars which support each terraced chamber. The ontogenetically youngest chamber is at the top of the animal (arrow). (D) An apical view with transmitted light through the specimen following NaOCl treatment to highlight the elaborate structure of the ostia. (E) An apical view of the intact animal. (F) An apical view of the animal following treatment with NaOCl and grinding to the sagittal plane. (G) Magnified view of the siphon following treatment with NaOCl. (H) SEM image of the ostia illustrating the unique pattern of inward facing spines. (I) SEM image roughly equivalent to the boxed section in B. Pillars (arrows) support each chamber, and are reinforced by radial spines (arrowhead). (J) Magnification of an ostium. (K) Magnification of the damaged inward facing spine boxed in J. (L) Magnification of the tip of the damaged spine boxed in K. Individual crystals of aragonite are clearly visible. (M) A sagittal section view after treatment with NaOCl and grinding shows the head and hypercalcified stalk regions. (N) SEM image of the stalk region after being etched with EDTA. Note that the pillars of the skeleton are still visible (arrows). (O) Magnification of the boxed section in N shows that the chambers are mineralized in layers (arrows). Note that not all chambers are mineralized entirely. (P) Both pillars and mineralized chambers are constructed by needles of aragonite.

3.3 Materials and methods

3.3.1 Sample collection

Specimens of *Vaceletia* sp. were collected during the Deep Down Under Expedition (www.deepdownunder.de) at Osprey- and Bougainville Reefs (Coral Sea, Australia) in depth from 5 to 24 m by SCUBA diving. Samples for RNA extractions were preserved in RNAlater and stored at -20 °C. Samples for protein extractions were freeze-dried and stored at -20 °C. Before processing, all samples were carefully inspected under a microscope for contaminating organisms which were carefully removed. Samples for protein extraction were then separated into head and stalk material.

3.3.2 Generation of a *Vaceletia* sp. transcriptome

Total RNA derived from the head of one individual was extracted using the miRNeasy Kit (Qiagen) according to the manufacturer's instructions. Single-end and paired-end Illumina sequencing was conducted using the MiSeq and HiSeq 2000 platforms respectively. A *de novo* transcriptome assembly was performed using Trinity [35], and this dataset was used to conduct all proteomic surveys. All contigs that yielded matches to the MS/MS data can be found in Supplementary file 1.

3.3.3 Matrix preparation

Pools of calcified *Vaceletia* sp. head and stalk pieces (approximately 2 g/pool for head pieces and 4 g/pool for stalk pieces) were treated with sodium hypochlorite (14 % active Cl_2; GPR Rectapur, VWR Chemicals, Germany; 10 ml/g) for 2 h at room temperature with a 5 min ultrasound treatment and changes of hypochlorite solution every 30 min. These pieces were then washed thoroughly with de-ionized water and air-dried. Next, the pieces were placed into a double-layered plastic bag and crushed into smaller pieces in a wrench to liberate the internal structures of the skeleton. The fragments were treated once again with sodium hypochlorite as described above. This treatment was performed in disposable 50 ml centrifuge tubes and skeleton fragments were collected by centrifugation for 5 min at 3000 x g between changes of hypochlorite solution and between washing (3 x) with de-ionized water. The dried skeleton pieces were then demineralization in 50 % acetic acid (20 ml/g of sponge skeleton) over night at 4-6 °C and the resulting suspension was dialyzed (Spectra/Por 6 dialysis membrane, molecular weight cut-off 2000; Spectrum Europe, Breda, The Netherlands) successively against 3 x 1l of 10 % acetic acid and 3 x 1l of 5 % acetic acid at

4-6 °C. The suspension was then lyophilized. The resulting organic matrices were analysed by SDS-PAGE using precast 4-12 % Novex Bis-Tris gels in MES buffer using reagents and protocols supplied by the manufacturer (Invitrogen, Carlsbad, California), except that 1 % β-mercaptoethanol was used as a reducing agent in the sample buffer. Samples were suspended in sample buffer (200 µg / 30 µL), heated to 70 ^{o}C for 10 min and centrifuged for 5 min at 15,700 x g in a 5415D Eppendorf centrifuge to remove sample buffer-insoluble material. The molecular weight marker was Novex Sharp Pre-stained (Invitrogen). Gels were either stained with the Coomassie using the Colloidal Blue staining kit or silver stained with SilverXpress (both Invitrogen).

3.3.4 Peptide preparation

Reduction, carbamidomethylation and enzymatic cleavage of matrix proteins were performed using a modification of the FASP (Filter-aided sample preparation) method [36] as outlined below. Aliquots of 200-300 µg of matrix were suspended in 300 µL of 0.1 M Tris, pH 8, containing 6 M guanidine hydrochloride and 0.01 M dithiothreitol (DTT). This mixture was heated to 56 °C for 60 min, cooled to room temperature, and centrifuged at 14,000 x g in an Eppendorf bench-top centrifuge 5415D for 15 min. The supernatant was loaded into an Amicon Ultra 0.5 ml 30 K filter device (Millipore; Tullagreen, Ireland). DTT was removed by centrifugation at 14,000 x g for 15 min and washing with 2 x 1 vol of the same buffer. Carbamidomethylation was done in the device using 0.1 M Tris buffer, pH 8, containing 6 M-guanidine hydrochloride and 0.05 mM iodoacetamide and incubation for 45 min in the dark. Carbamidomethylated proteins were washed with 0.05 M ammonium hydrogen carbonate buffer, pH 8, containing 2 M urea, and centrifugation as before. Trypsin (2 µg, Sequencing grade, modified; Promega, Madison, USA) was added in 40 µL of 0.05 M ammonium hydrogen carbonate buffer containing 2 M urea and the devices were incubated at 37 °C for 16 h. Peptides were collected by centrifugation and the filters were washed twice with 40 µL of 0.05 M ammonium hydrogen carbonate buffer and twice with 1 % trifluoroacetic acid in 5 % acetonitrile. The acidic peptide solution (pH 1-2) was applied to C18 Stage Tips [37] and the eluted peptides were vacuum-dried in an Eppendorf concentrator.

3.3.5 LC-MS and data evaluation

Peptide mixtures were analysed by on-line nanoflow liquid chromatography using the EASY-nLC 1000 system (Proxeon Biosystems, Odense, Denmark, now part of Thermo

Fisher Scientific) with 50 cm capillary columns of an internal diameter of 75 µM filled with 1.8 µM Reprosil-Pur C18-AQ resin (Dr. Maisch GmbH, Ammerbuch-Entringen, Germany). Peptides were eluted with a linear gradient from 5-30 % buffer B (80 % acetonitrile in 0.1 % formic acid) in 90 min, 30-60 % B in 5 min and 60-95 % B in 5 min at a flow rate of 250 nL/min and a temperature of 50 °C. The eluate was electro-sprayed into an Orbitrap Q Exactive (Thermo Fisher Scientific, Bremen, Germany) using a Proxeon nanoelectrospray ion source. The instrument was operated in a HCD top 10 mode essentially as described [38]. The resolution was 70,000 for full scans and 7,500 for fragments (both specified at m/z 400). Ion target values were 1e6 and 5e4ms, respectively. Dynamic exclusion time was 20 sec or 10 sec. MS runs were monitored using the SprayQC quality monitoring system [39]. Raw files were processed using the Andromeda search engine-based version 1.5.0.8 of MaxQuant (http://www.maxquant.org/) with enabled second peptide, iBAQ, and match between runs (match time window 0.5 min; alignment time window 20 min) options [40, 41]. The sequence database was combined with the reversed sequences for FDR calculation and sequences of common contaminants, such as human keratins and mammalian cytoskeletal proteins. Carbamidomethylation was set as fixed modification. Variable modifications were oxidation (M), N-acetyl (protein), pyro-Glu/Gln (N-term), phospho (STY), and hydroxyproline. The initial mass tolerance for full scans was 7 ppm and 20 ppm for MS/MS. Two missed cleavages were allowed and the minimal length required for a peptide was seven amino acids. Maximal FDR for peptide spectral match, proteins and site was set to 0.01. The minimal score for peptides was 60 and the minimal delta score for modified peptides was 17. Identifications with only one or two sequence-unique peptides identified at least 10 and three times, respectively, were routinely validated using the MaxQuant Expert System software [42] considering the assignment of major peaks, occurrence of uninterrupted y- or b-ion series of at least four consecutive amino acids, preferred cleavages N-terminal to proline bonds, the possible presence of a2/b2 ion pairs and immonium ions, and mass accuracy. The iBAQ (intensity-based absolute quantification) [43] option of MaxQuant was used to calculate, based on the sum of peak intensities, the approximate share of each protein in the total proteome, including identifications, which were not accepted after manual validation. This enabled us to discern between minor and major proteins.

Sequence similarity searches were performed with FASTA (http://www.ebi.ac.uk/Tools/sss/fasta/) [44] against current releases of the Uniprot Knowledgebase (UniProtKB). Other bioinformatics tools used were Clustal Omega for

sequence alignments (http://www.ebi.ac.uk/Tools/msa/clustalo/) [45], InterPro (http://www.ebi.ac.uk/interpro) [46] for domain predictions, SignalP 4.1 (http://www.cbs.dtu.dk/services/SignalP/) [47] for signal sequence prediction, and TMHMM 2.0 (http://www.cbs.dtu.dk/services/TMHMM-2.0) [48] for transmembrane sequence prediction. Amino acid composition and theoretical pI were determined using the ProtParam tool provided by the Expasy server (http://web.expasy.org/protparam/) [49]. Intrinsically disordered protein structure was predicted using IUPred (http://iupred.enzim.hu/) [50]. Subcellular location predictions were based on sequence similarities to known proteins, domain predictions, signal sequence predictions, and transmembrane segment predictions.

3.3.6 Spherulin sequence alignment and phylogenetic analysis

An alignment of Spherulin homologs was constructed using the dataset described in Jackson et al [33]. The *Vaceletia* sp. sequence was included in this collection of sequences and aligned as previously described. Phylogenetic analyses were conducted using MrBayes v. 3.2.3 and the following parameters: lset nst=6 rates=invgamma; prset applyto = all; mcmcp nruns=4, ngen= 1000000, relburnin = yes, burninfrac = 0.25, printfreq = 1000, samplefreq = 100, nchains = 4, temp = 0.2, savebrlens = yes. The first 25 % of these trees were discarded as burn in.

3.3.7 Histology

Fixed *Vaceletia sp.* material was decalcified, dehydrated, embedded in paraffin and sectioned at 2-5 µM. Sections were deparaffinised and stained using alcian-blue fast red dye [51].

3.3.8 Comparison of *Vaceletia* sp. biomineral proteins to other sponge transcriptomes

TBLASTX based comparison of selected *V.sp* proteins were made against eight previously published sponge transcriptomes [52]. For sequence similarity searches, recovered sponge sequences were mapped against the NCBI-uniprot/swissprot database using BLASTX. All Blast searches were done using an e-value cut-off of 1e-5. HMMER v3.1b2 (www.hmmer.org) and CD-Search [53] were used to screen for protein domains against the Pfam 28.0 Protein Family database [54] and the CDD database v3.14 [55], respectively.

3.4 Results and Discussion

The yields of organic matrix/g of skeleton were 2.4 mg for head and 2.2 mg for stalk. This value was in good agreement with matrix yields of invertebrate biomineral matrices reported previously [4–6, 8, 56]. However, PAGE analysis of the matrix proteins yielded a different outcome (Fig. 2). Coomassie Brilliant Blue staining showed only very few faint bands that became more prominent with silver staining. This indicated that most of the matrix was either not soluble in PAGE sample buffer or that most of the matrix was not protein. For protein cleavage and peptide isolation under denaturing conditions we used FASP [36], a gel-independent method. The number of identified proteins was low. The head matrix yielded 203 proteins (Supplementary file 2: ProteinGroups_HEAD) and the stalk matrix yielded 105 proteins (Supplementary file 3: ProteinGroups_STALK), with 19 identifications unique to stalk matrix in this initial list. In agreement with the relatively low number of proteins we identified very few sequence-unique peptides. In head matrix these were 610 (Supplementary file 4: Peptides_HEAD) and in stalk we obtained only 215 (Supplementary file 5: Peptides_STALK). Furthermore, 43% of the head matrix proteins and 50% of those of stalk matrix were identifications with only one sequence-unique peptide. Such identifications are not commonly accepted in mass spectrometry-based proteomics, at least with samples containing predominantly or exclusively protein. However, many of these peptides were identified many times. Thus, for instance, entry C53634_gi_i1_1, encoding an uncharacterized very acidic protein, was identified with a single sequence-unique peptide

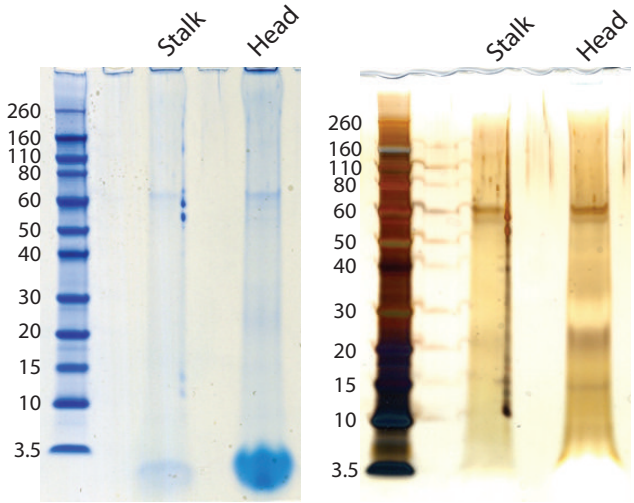

Fig. 2. PAGE separation of stalk (S) and Head (H) skeleton matrix. The same gel was first stained with Coomassie Brilliant Blue (CBB, left) and then with silver (right). The molecular weight of marker proteins is given in kDa.

that was identified 213 times altogether. Inspection of the sequence contained in this entry indicated that the identified peptide was most probably the only one that could be detected at all. Therefore this identification clearly was a valid one. In other cases the reasons for

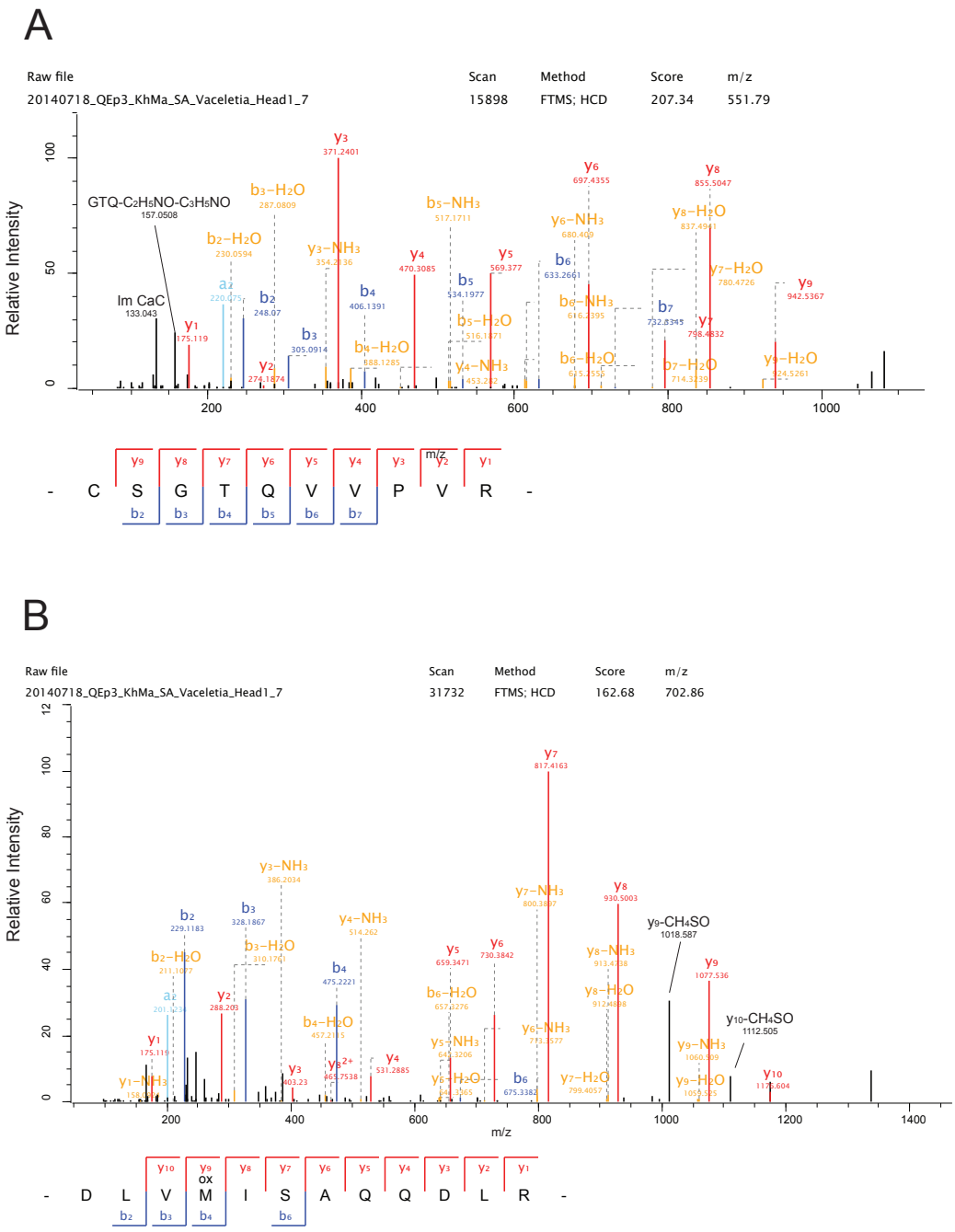

Fig. 3. Selected spectra of single sequence-unique peptide identifications. Y-ions are shown in red, b-ions in blue, a-ions in light blue, b- and y ions showing loss of water or ammonia are shown in orange, ions annotated with the help of the MaxQuant Expert System are shown in black. (A) peptide of entry c102844_g1_i1_3. Two fragments annotated with the help of the Expert system are the immonium ion of carbamidomethylated cysteine (Im CaC) and an internal fragment at m/z 157.0508 derived from the tripeptide GTQ. (B) peptide of entry C41414_g3_i1_2. Ions y10 and y9 show the loss of CH_4SO typical for oxidized methionine residues (Met-sulfoxides).

identification of only one peptide were less obvious and could have included errors in the database, unanticipated modifications, or the scarcity of protein in these samples. Therefore we decided to provisionally accept identifications with one sequence-unique peptide if this was identified more than 10 times and after manual validation of the spectra with the help of the Expert System that is part of the MaxQuant software package [42]. Figure 3 shows some typical annotated spectra of this kind. After elimination of identifications not conforming to these criteria, and combining identifications apparently belonging to the same protein, we obtained a list of 122 accepted protein identifications (Supplementary file 6: *Vaceletia* sp. skeleton matrix proteins). Identifications that were not accepted are provided in Supplementary files 2-5.

All proteins identified in the stalk were present in the head, but some proteins identified in the head were unique to that location. Based on iBAQ values that yield the percentages of proteins normalized to the sum of iBAQ intensities of all identified proteins in a sample, 40 of the 121 identified head proteins constitute more than 90 % of the total identified head proteome and 35 of the 72 identified stalk proteins (all of which are present in the head proteome) constitute more than 87 % of the stalk total identified proteome (Table 1). We will only consider these 40 "major" proteins further as they are likely to represent the key components of *Vaceletia*'s biomineral proteome, however all 122 isotigs (consisting of 181 contigs) are provided in the supplementary material. In general the majority of these proteins share similarity with proteins in UniProt and/or contain recognisable protein domains (Table 1); eleven of the 40 most abundant proteins did not return hits against UniProt. Of these 40 major proteins approximately 50% apparently differ in their abundance within the head and stalk regions (19 out of 40; Fig. 4).

The most abundant protein in the *Vaceletia* sp. skeletal proteome (Contig 7761) is found at levels more than 10 times that of the next most abundant (Table 1) and shares significant similarity with the Astrosclerins, a family of alpha carbonic anhydrases (α-CAs) previously identified in another coralline demosponge, *Astrosclera willeyana* [57]. CAs catalyse the reversible reaction of CO_2 and water to HCO_3^- (which can then react with free Ca^{2+} to form $CaCO_3$) and are known to play an important role in invertebrate biomineralization [58, 59]. Astrosclerin is directly involved in the formation of the hypercalcified aragonitic skeleton of *A. willeyana* and is also highly expressed in that sponge [57]. The *Vaceletia* sp. α-CA homolog constitutes more than 30% of the head proteome and more than 50% of the stalk

Table 1. The major proteins of the *Vaceletia* sp. head and stalk proteome: 40 proteins (with an iBAQ percentage more than 0.1) constitute more than 90% of the head and more than 87% of the stalk proteome.

Contig	Similarity to best match	E-value	Protein features	Isoelectric point	% of total in head/stalk (iBAQ)
C7761_g1_i1_1	A6YCJ0 (Sponge)	4.9e-29	Similar to astrosclerin-2; Domain: α-; 11% L; pI 5.6; shares 1 peptide with c94004_g1_i1_2	5.6	31.9 / 51.6
C38723_g1_i1_3	K1WIY3 (Cyanobacteria)	1.8e-8	Similar to Na-Ca exchanger/integrin-β4; domains: Na-Ca-exchanger/integrin_β4; TMH, PM	4.0	8.8 / 3.2
C99840_g1_i1_1	None	-	(10% G, 11% I, 11% V); TMH	5.2	5.8 / 0.8
C53634_g1_i1_3	None	-	(18% D, 12% E, 13% I, 10% V)	3.5	5.6 / 8.2
C36962_g2_i1_6	W4Y3E1 (Urchin)	1.5e-25	Sp-Srcr85; domain: SRCR, PM	4.5	5.6 / 3.4
C23124_g1_i2_3 C23124_g1_i1_3	None	-	domain: Na-Ca_exchanger/intrgrin_β4	4.0	5.0 / -
C94004_g1_i1_2	None	-	-	-	4.1 /-
C77644_g1_i1_3	H2Y8G7 (Ascidian)	3.5e-4	domain: fibrinogen_α,β,γ_C_term_glob, subdomain_2; THM; EC	6.0	2.9 / -
C32287_g1_i1_1	I1G7C7 (Sponge)	3.2e-37	(10% I, 10% S); see also C31462_g1_i1_1	5.7	2.2 / 1.2
C29357_g1_i1_2	A0A022L1D0 (Actinobacteria)	1.1e-6	Uncharacterized collagen (fragment) /α1,6-glucosidase; domain: triple_helical, EC	9.2	1.9 / 1.8
C22072_g1_i1_3	None	-	15 % L; THM	9.2	1.5 / -
C54677_g1_i1_2	Q5QBF8 (Insect)	1.5e-65	Ubiquitin; IC, EC	-	1.3 / 2.7
C3544_g1_i1_1	Q2KT50 (Diatom)	2.4e-32	Actin 2	8.9	1.1 / 1.7
C37591_g1_i3_5	B5X2X5 (Bony Fish)	7.2e-13	Spondin-2; Domain: spondin; EC	8.6	1.1 / 1.2
C1963_g1_i2_2	K1QSR0 (Oyster)	6.4e-7	Similar to angiopoietin-4; domain: fibrinogen_α,β,γ_C_term_glob, subdomain_1; EC	5.7	1.1 / 0.4
C20021_g1_i1_2	H2AZL (Frog)	1.1e-40	Histone H2A	9.7	0.9 / 5.2
C36962_g2_i3_6 C80079_g1_i1_2	W4XYX3 (Urchin)	3.0e-20	Sp-Srcr71; domain: SRCR	4.1	0.9 / 0.2

Contig	Similarity to best match	E-value	Protein features	Isoelectric point	% of total in head/stalk (iBAQ)
C41075_g1_i4_4 C41075_g1_i2_4	V5YU14 (Starfish)	5.9e-160	β-actin; shares 4 peptides with c14026_g1_i1_3 and 1 with c3544_g1_i1 and c21396_g1_i1_4, IC	-	0.8 / 1.9
C32738_g1_i3_3	I1G9M3 (Sponge)	8.1e-8	Uncharacterized; 11% L; TMH	5.5	0.8 / 0.1
C64227_g1_i1_3	None	-	Uncharacterized; domain: fibrinogen _α,β,γ_C_term_glob; 10% L; TMH	8.2	0.7 / 1.1
C102844_g1_i1_3	S9WWY6 (Mammal)	7.4e-12	Similar to neurotrypsin (fragment); domain: SRCR; 13% G, 13% S, 10% V; PM	8.7	0.7 / 0.8
C21396_g1_i1_4	G9I1P2 (Bony Fish)	1.2e-47	Cytoplasmic β-actin (fragment); domain: actin_related (aa1-115); shares peptide with c3544_g1_i1_1 and c41075_g1_i4_4/c41075_g1_i2_4; IC	-	0.6 / 0.6
C80614_g1_i1_3	V6GWB1 (Spirochaetes)	4.6e-18	Similar to peroxidasin; domain: haem_peroxidase	8.8	0.6 / 0.3
C97612_g1_i1_1	K1QE34 (Oyster)	8.5e-11	Similar to DBH-like monooxygenase protein 2-like protein; domain: DOMON	4.2	0.5 / 0.4
C41117_g3_i2_4 C41117_g3_i3_5 C41117_g3_i1_5 C41117_g3_i4_6	K7LZT4 (Soybean)	2.1e-39	Histone H4; IC	-	0.4 / 0.8
C40964_g7_i1_1 C40964_g7_i2_1 C40964_g7_i4_2	K1R2Z9 (Oyster)		Uncharacterized/hemicentin-1; domains: metallopeptidase, disintegrin, EGF_3, 6x TSP1; 12% G, 11% S; EC, PM	5.6	0.4 / 0.2
C3160_g1_i1_2 C3160_g1_i2_2	H2V0I8 (Bony Fish)	9.8e-8	domains: ConA_lectin/LamG, EGF-like;	6.1	0.4 / 0.2
C35050_g1_i1_1	None	-	Uncharacterized; domain: Na-Ca-exchanger/integrin_β4	4.4	0.3 / 0.5
C40249_g1_i3_3 C40249_g1_i2_3 C40249_g1_i1_3	A7S664 (Sea Anemone)	2.3e-12	Uncharacterized; domain: VWA	4.9 -5.7	0.3 / 0.3

Contig	Similarity to best match	E-value	Protein features	Isoelectric point	% of total in head/stalk (iBAQ)
C38115_g2_i1_3 C38115_g1_i1_5	I1GHA4 (Sponge)	1.1e-81 9.5e-71	Enolase; domains: enolase_N-term, enolase_C-term; TMH; IC, PM		0.3 / 0.2
C38911_g1_i3_1	None		Uncharacterized; pI 5.4; domain: PTHR24637; TMH	-	0.3 / 0.2
C100960_g1_i1_4	I1FHH5 (Sponge)	5.2e-5	Similar to Hedgeling/uncharacterized; domain: VWA; PM		0.3 / 0.1
C27354_g1_i1_6 C34006_g1_i1_6 C34006_g1_i2_6	F6VY96 (Mammal)	1.4e-50	Histone H3 (fragment)	-	0.2 / 0.8
C41693_g1_i7_5 C41693_g1_i3_6	I1EQR1 (Sponge)	5.1e-3	Uncharacterized; domain: fibrinogen_ α,β,γ_C_term_glob ; EC	5.6	0.2 / 0.1
C41731_g1_i3_5	None	-	(10% I, 12% L, 10% S), TMH	9.2	0.2 / 0.1
C35925_g1_i3_2	None	-	Uncharacterized; domains: IG; 12% S; pI 6.3; TMH	6.3	0.2 / 0.1
C41377_g2_i1_1	None	-	Uncharacterized; domains: IG (58-137), DUF4440 (172-278); shares peptides with c41377_g2_i2_1; TMH	8.5	0.2 / 0.1
C41584_g1_i4_5 C41584_g1_i2_5 C41584_g1_i5_5	G8HT99 (Stony Coral)	1.7e-5	Uncharacterized/similar to cytochrome c oxidase subunit 3; 10% L, 13% S; shares 1 peptide with C41584_g1_i8_4; TMH	6.6	0.2 / 0.1
C32545_g1_i1_1 C32545_g1_i2_1	H6TI88_9METZ (Sponge)	3.7e-33	Spherulin; SSP (aa26/27); EC	4.7	0.1/-
C103979_g1_i1_6	B5XCM2 (Bony Fish)	8.0e-40	Calmodulin; domain: EFh_pair, shares 1 peptide with C27518_g1_i1_4H; IC	-	0.1/0.1

IC = intracellular; EC = extracellular; PM = plasma membrane. TMH = predicted trans-membrane helix.

proteome, suggesting that this protein is also a key component of *Vaceletia*'s biomineralization toolkit.

Jackson and co-workers also identified another protein involved in biomineralization in *Astrosclera willeyana* that is present in *Vaceletia* sp.'s skeleton. Spherulin is expressed in the same spherulite forming cells as Astrosclerin, and was most likely acquired via a horizontal gene transfer (HGT) event from a prokaryote [33]. *Vaceletia* sp. isotig 32545 (possibly represented by two contigs) shares significant similarity with the Awi-spherulin. Interestingly it is only present in *Vaceletia* sp.'s head proteome in minor quantities (0.1% of total head iBAQ). *Vac*-spherulin only returns hits against bacterial proteins with similarities to sugar transporters. This finding supports the hypothesis that an HGT event delivered this gene into the genome of a common ancestor of the demosponges *Astrosclera willeyana, Vaceletia* sp.*, Amphimedon queenslandica, Chondrilla nucula, Spongilla lacustris* and the hexactinellid *Aphrocallistes vastus* and was subsequently co-opted to a biomineralization role in calcifying sponges. Our phylogenetic analyses support this interpretation with all sponge spherulins clustering together and the bacterial orthologues forming well separated clades (Supplementary file 7). We would like to point out here that although *A. willeyana*

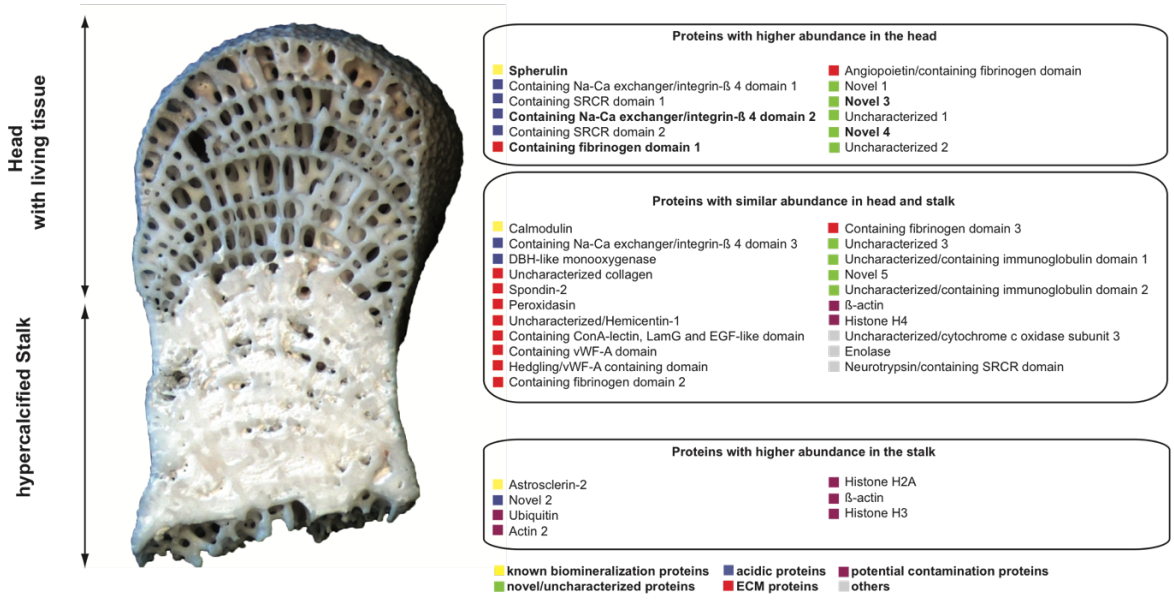

Fig. 4. Schematic representation of *Vaceletia* sp. head and stalk region with iBAQ estimates of protein abundances. 12 proteins are enriched in the head, while 7 proteins are enriched in the stalk and 21 proteins have an equal abundance in head and stalk. Protein abundances were considered as different when iBAQ estimations between head and stalk were ≥ 0.5. Proteins in bold were only detected in the head proteome

and *Vaceletia* sp. display very different skeletal morphologies, they apparently share at least two important biomineralization proteins (Astrosclerin and Spherulin). These underlying molecular commonalities should be taken into account when considering the broad evolutionary picture of biomineralization and the apparent plasticity of skeletal morphologies. To elucidate this intriguing question further more data from *Astrosclera willeyana* and other calcifying sponges with divergent skeletal morphologies is required.

3.4.1 Extracellular matrix proteins (ECM)

Some proteins extracted from *Vaceletia* sp.'s skeletal proteome can be identified as ECM proteins or share similarity with previously characterised biomineralization proteins. c40964 shows similarity to hemicentin-1, a cell adhesion protein that recently has been reported as a soluble organic matrix protein (SOMP) from the coral *Acropora millepora* [60], while c80614 shows similarities to peroxidasin which has been suggested to cross-link proteins in the extracellular space [61]. c37591 contains two spondin domains which are known to function as extracellular neuroregulators by guiding axon growth [62]. Recently, spondins have also been found to play a role in processes associated with bone mineralization; F-spondin seems to be involved in the regulation of bone metabolisms resulting in a high bone mass phenotype in F-spondin deficient mice [63]. All spondins involved in bone metabolism contain six thrombospondin-type 1 domains at their C-terminus, which are absent in the detected *Vaceletia* sp. protein. Spondin-2, also called mindin, has been proposed to function in mice as a pattern-recognition ECM molecule involved in opsonisation of microbes for phagocytosis, and is therefore essential to the initiation of the innate immune response [64]. *Vac*-c37591 may therefore either be directly involved in the biomineralization process or provide the skeleton with the ability to resist microbial action.

Although many of the proteins we have identified in *Vaceletia* sp.'s skeleton share no overall similarity with proteins in UniProt, they do contain recognisable domains that are associated with extracellular matrix and/or biomineralization proteins. For example ConA_lectin, laminin G and epidermal growth factor-like (EGF) domains have been reported from the mineralizing matrix of scleractinian corals [60]. Two *Vaceletia* sp. proteins (c40249, c100960) contain a von Willebrand factor type A (vWF-A) domain, which are present in biomineral-associated proteins from several organisms. For example Pif, an acidic protein derived from the pearl oyster *Pinctada fucata* possesses a vWF-A domain that has been shown to play an important role in the formation of nacre [65]. In the calcareous sponge

Leucosolenia complicata, a vWF-A domain is located at the N-terminus of a carbonic anhydrase that plays an important role in sclerite formation [58].

Collagen, a fibrous protein, and Chitin, an amino-polysaccharide, play important roles in a variety of biomineralization systems as they can act as templates upon which mineralization takes place [66]. For example corals use collagens as a mineralization framework (reviewed in [66]), molluscs can employ chitin [67], while keratose sponges use sponge specific collagens to form their organic skeleton [68]. We detected several collagens in *Vaceletia* sp.'s mineral proteome (c27773, c29357, c34834, c36461, c40551, c58706), with one of these (c29357) categorised as a major protein. We also detected a protein with a chitin-binding domain that is only present in the head region albeit at low abundance (c34763). Chitin is known to act as a mineral framework in Verongida sponges [66, 69].

3.4.2 Acidic proteins

Acidic proteins have long been recognised in many organic matrices associated with biomineralization, and it is assumed that they play a major role in this process. Acidic proteins possess a pI < 4.5 [70] and are often rich in negatively charged residues such as aspartic and glutamic acid. They are thought to serve many purposes such as promoting the nucleation of calcium carbonate, determining the growth axes and inhibiting crystal growth [71]. The seven acidic proteins detected in *Vaceletia* sp.'s skeletal proteome account for a large proportion of the total proteome (27 % of the total head proteome and 16 % of the total stalk proteome). Three of these proteins contain a sodium-calcium-exchanger domain/integrin domain, suggesting a role for binding and delivering calcium to the site of mineralization. Two other acidic proteins contain a scavenger receptor cysteine-rich domain. Proteins containing this domain have been reported from the sea urchin test and spine proteomes [5]. Another acidic protein shows no significant similarity to any UniProt entry but is enriched in glutamic and aspartic acid (18% and 12 % respectively). Given their abundance in *Vaceletia* sp.'s skeletal proteome we visualised the spatial distribution of acidic molecules by using an alcian blue + direct-red stain on sections of *Vaceletia* sp.. This staining reveals the insoluble organic framework of the skeleton and an abundance of acidic mucus substances throughout the head and stalk regions with more intense staining in the head region (Fig. 5A-D).

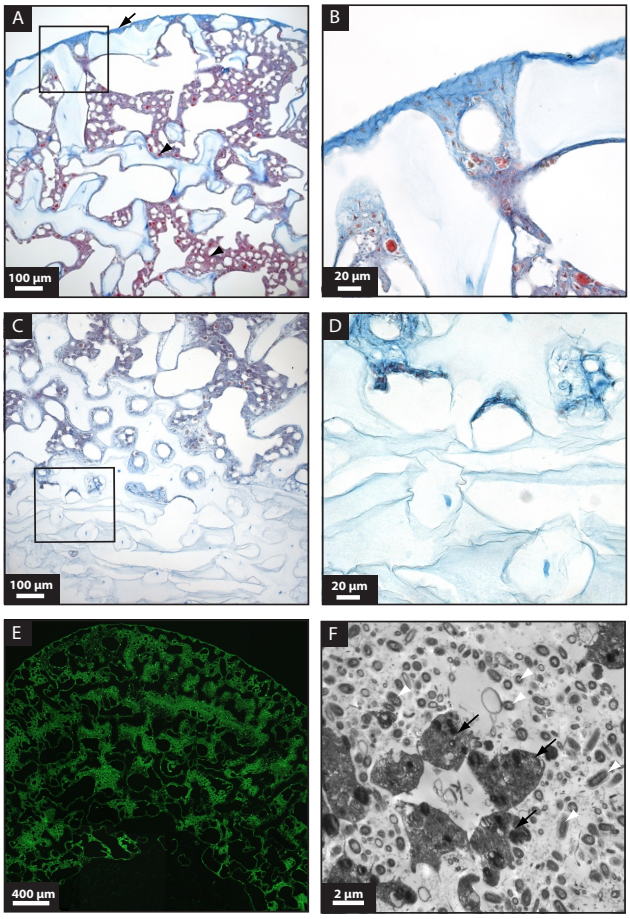

Figure 5. General histological features of decalcified *Vaceletia* sp. (A) Alcian blue stained section of a sagittal sectioned individual illustrating the head region. Blue staining reveals acidic mucopolysacharrides and likely reflects the location of previously calcified pillars (see Figure 1I). Intense red stain reveals sponge larvae (arrowhead) and sponge tissue. (B) Magnification of the boxed section in A illustrates the more intense blue staining in the outermost head region where the acidic substance is produced. (C) Alcian blue staining in the stalk region is less intense than in the head region. (D) Magnification of the boxed section in C shows that the previously calcified stalk region contains no red stained sponge tissue. (E) Sagittally sectioned individual shows the autofluorescent sponge tissue in the head region and a lack of cells in the stalk region. (F) TEM image of sponge mesohyl filled with darkly stained sponge cells (black arrows) and diverse and abundant bacteria (white arrowheads).

3.4.3 Uncharacterized proteins

Six proteins extracted from *Vaceletia* sp.'s skeleton contain no recognisable domains and show no similarity to UniProt entries or show similarity to uncharacterized proteins. In some cases it was possible to predict a trans-membrane (TM) domain in these novel proteins. Proteins with TM domains have been found in other datasets of skeletal organic matrix proteins [56, 60, 72]. It has been proposed that membrane bound and trans-membrane

proteins may contribute directly to the biomineralization process via their extracellular domains. However without functional assays (either *in vitro* or *in vivo*) the roles that these proteins play in biomineralization remain unknown. This is the situation for many proteomic studies of invertebrate biominerals and highlights the growing need for the development of such assays.

3.4.4 Potential contaminants

Intracellular proteins such as actin, myosin and tubulin are often considered to be contaminations in biomineral protein datasets [73]. According to current working models of biomineralization such skeletal proteomes should only include proteins that are intimately associated with (or occluded within) the mineral phase. Intracellular components such as those listed above have been suggested to derive from cellular debris that remains following inadequate cleaning of the biomineral (i.e. a technical artefact), or that has been inadvertently occluded within the biomineral during its formation [73]. We detected several proteins that fall within this category, namely actin, ubiquitin and histones. In all cases, the abundance of these proteins is higher in the stalk than in the head. *Vaceletia* sp.'s mode of growth may explain this observation. The living tissue of the head lays down new mineral material in the 'outermost' region of the animal (Fig. 1 C), and as growth ensues older more proximal regions of the skeleton continue to mineralize until they are completely fused into the stalk region (Fig. 1 M-P). There is no living tissue in the stalk (Fig. 5 C-E), so we assume that there would be some degree of apoptosis or unregulated cell death in the region that borders the head and stalk regions. We suggest that cell death in this region may be the source of the higher abundance of these 'contaminating' proteins. Further investigation using cell death and proliferation markers would help to resolve this issue. Of course the alternative interpretation is that these proteins may be genuine biomineralization components. Indeed it has been previously shown that actin and ubiquitin may be involved in the formation of mineralized body parts [74–76]. However we prefer the former interpretation given the lack of living tissue in the stalk region.

3.4.5 Microbes apparently play a minor proteomic role in skeleton formation in *Vaceletia* sp.

Many sponges are host to a high diversity and large quantity of microbes and species of *Vaceletia* are no exception [77] (Fig. 5 F). Uriz et al. recently demonstrated that specific bacteria play a direct role in the biomineralization strategy of demosponges belonging to the

genus *Hemimycale* [34]. The skeleton of *Hemimycale* sponges is composed of calcitic spherules that are produced by endosymbiontic 'calcibacteria'. Given the deep evolutionary association between sponges and microbes, it was not surprising that a HGT event was detected within the skeleton of another calcifying demosponge *A. willeyana* [33], and that we can identify this same gene product in the skeleton of *Vaceletia* sp.. Unexpectedly, and despite the fact that *Vaceletia* sp. is host to a vast community of various microbes, our proteome data contains no evidence that proteins synthesised by symbiotic microorganisms are directly involved in the process of biomineralization within this sponge. However this does not exclude the possibility that microbes may contribute to *Vaceletia* sp.'s biomineralization strategy via other metabolic pathways. Unfortunately developing this line of research further is challenged by the technical limitations of working with these kinds of animals; they are difficult to maintain (let alone culture) in aquaria over long periods of time (necessary in order to observe meaningful skeletal growth), they are found in remote localities, and few functional molecular tools have been developed for them.

3.4.6 Comparison of *Vaceletia* sp. biomineral proteins to other sponge transcriptomes

Very little work has been done to investigate the molecular biomineralization strategy of sponges in general, and the absence of any other sponge biomineral proteomes prevents the investigation of any potential broad commonalities employed by sponges to build their skeletons. Riesgo et al. [52] recently reported the characterisation of eight sponge transcriptomes and while these datasets were not focused on the identification of biomineralization proteins we conducted a survey of these resources using 20 *V. sp.* biomineralization proteins that were selected on the basis of their high abundance (≥ 1 % of the total head proteome) or their potential role in the biomineralization process. Unlike *Vaceletia* six of the eight sponges employ silica as a primary biomineral (*Aphrocallistes vastus*, *Chondrilla nucula*, *Petrosia ficiformis, Spongilla lacustris*, *Pseudospongosorites suberitoides*, *Corticum candelabrum*), while *Ircinia fasciculata* possesses a solely fibrous skeleton and *Sycon coacatum* is the only species to use calcium carbonate to build its skeleton.

Of the 20 *V.sp* proteins approximately 50% shared similarity (at an e-value threshold of 1e-5) with one or more proteins derived from the eight sponge transcriptome datasets (Supplementary file S8). A key component of *Vaceletia's* biomineralization toolkit, a carbonic anhydrase similar to Astrosclerin, is present in seven out of the eight

transcriptomes. Besides the role of CA in biomineralization, CA enzymes are also involved in a variety of other metabolic processes such as CO_2 transport and pH and ion regulation [78, 79]. CA has been identified as a key enzyme employed in the biomineralization strategy of another *Sycon* species *S. ciliatum* [58], and it is therefore likely to be involved in the mineralization process of *S. coacatum*.

Interestingly, we were able to detect the previously described horizontally transfered gene *Spherulin* [33] in the hexactinellid sponge *A. vastus* and in two demosponges, *C. nucula* and *S. lacustris* from the dataset of Riesgo [52], but could not detect it in the calcifying *S. coacatum*. It is tempting to speculate that besides playing a role in sponge calcification [33] Spherulin may also be involved in biosilification. However, the function of Spherulin remains unknown and without further data this must remain speculation. The absence of Spherulin in *S. coacatum, A. vastus*, *P. ficiformis, . suberitoides*, *C. candelabrum* and *I. fasciculata* may either indicate the loss of this gene in these species or a lack of expression in the Riesgo [52] transcriptome datasets.

The majority of the 20 *V. sp.* biomineralizing proteins used in this comparison share similarity to domains present in contigs represented in all 8 of the Riesgo transcriptomes (Supplementary file S8). However on the basis of these sequence similarity results it is problematic to infer any genuine homology to the *V. sp.* biomineralizing proteins we report here; while proteins may share recognisable domains that confer a similar function to the entire protein, this does not necessarily imply that those proteins share an evolutionary history and so we interpret the results of these comparisons with caution.

3.5 Conclusion

The proteome that we report here for *Vaceletia* sp. is the first comprehensive biomineralization dataset from a sponge. As reported for other biomineralization proteomes it contains proteins known to play roles in biomineralization, and novel proteins that display no similarity to known proteins. The presence of deeply conserved biomineralization proteins such as α-CA illustrates that the Last Common Ancestor of the Metazoa (LCAM) did indeed contribute some biocalcification tools to its descendants, and that therefore there is likely to be considerable conservation in the molecular details of skeleton formation across the Metazoa despite divergent skeletal morphologies. The data we reported here suggests that different mineralization processes are taking place in the head and stalk regions, and

that bacteria apparently contribute minimal proteinaceous resources to the construction of *Vaceletia* sp.'s skeleton. Skeletogenic proteome surveys such as reported here are an important resource that serve to both expand our knowledge of the protein repertoires animals use to biomineralize, and how this ability evolved. However, the lack of functional assays available to study in detail the role that these proteins play remains a major challenge to the field of biomineralogy.

Acknowledgements:

We would like to acknowledge Matthias Mann (MPI of Biochemistry, Martinsried), for generous support, Gaby Sowa (MPI) for preparing the capillary columns, Korbinian Mayr and Igor Paron (both MPI) for keeping the mass spectrometers in excellent condition. Dorothea Hause-Reitner provided expert SEM and TEM support and Wolfgang Dröse assisted with histology work.

Author contributions:

Conceived and designed the experiments: DJJ and KM. Performed the experiments: KM, DJJ and JG. Analysed the data: KM, JG and DJJ. Contributed reagents/materials/analysis tools: KM, DJJ, GW. Wrote the paper: KM, JG and DJJ. Sample collection: DJJ and GW. Data interpretation and manuscript revision: KM, JG, GW and DJJ.

Supplementary material

The Supplementary Material for this article can be found online at: http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0140100#sec019 or via the attached CD.

References

1. Knoll, A. H. (2003) **Biomineralization and evolutionary history.** *Reviews in Mineralogy and Geochemistry*, 54(1), 329-356, doi:10.2113/0540329.

2. Jackson, D. J., McDougall, C., Green, K., Simpson, F., Wörheide, G., Degnan, B. M. (2006) **A rapidly evolving secretome builds and patterns a sea shell.** *BMC Biol*, 4, 40, doi:10.1186/1741-7007-4-40.

3. Marie, B., Marie, A., Jackson, D. J., Dubost, L., Degnan, B. M., Milet, C. et al. (2010) **Proteomic analysis of the organic matrix of the abalone *Haliotis asinina* calcified shell.** *Proteome Sci.*, 8, 54, doi:10.1186/1477-5956-8-54.

4. Mann, K., Edsinger-Gonzales, E., Mann, M. (2012) **In-depth proteomic analysis of a mollusc shell: acid-soluble and acid-insoluble matrix of the limpet *Lottia gigantea*.** *Proteome Sci.*, 10(1), 28, doi:10.1186/1477-5956-10-28.

5. Mann, K., Poustka, A. J., Mann, M. (2008) **The sea urchin (*Strongylocentrotus purpuratus*) test and spine proteomes.** *Proteome Sci.*, 6, 22, doi:10.1186/1477-5956-6-22.

6. Mann, K., Wilt, F. H., Poustka, A. J. (2010) **Proteomic analysis of sea urchin (*Strongylocentrotus purpuratus*) spicule matrix.** *Proteome Sci.*, 8, 33, doi:10.1186/1477-5956-8-33.

7. Immel, F., Gaspard, D., Marie, A., Guichard, N., Cusack, M., Marin, F. (2015) **Shell proteome of rhynchonelliform brachiopods.** *J. Struct. Biol.*, 190(3), 360-366, doi:10.1016/j.jsb.2015.04.001.

8. Jackson, D. J., Mann, K., Häussermann, V., Schilhabel, M., Lüter, C., Griesshaber, E. et al. (2015) **The *Magellania venosa* biomineralizing proteome: a window into brachiopod shell evolution.** *Genome Biol. Evol.*, doi:10.1093/gbe/evv074.

9. Wörheide, G., Dohrmann, M., Erpenbeck, D., Larroux, C., Maldonado, M., Voigt, O. et al. (2012) **Deep phylogeny and evolution of sponges (phylum Porifera).** *Adv. Mar. Biol.*, 61, 1-78, doi:10.1016/B978-0-12-387787-1.00007-6.

10. Dohrmann, M., Wörheide, G. (2013) **Novel Scenarios of Early Animal Evolution - Is It Time to Rewrite Textbooks?** *Integr. Comp. Biol.*, 53(3), 503-511, doi:10.1093/icb/ict008.

11. Philippe, H., Brinkmann, H., Lavrov, D. V., Littlewood, D. T. J., Manuel, M., Wörheide, G. et al. (2011) **Resolving Difficult Phylogenetic Questions: Why More Sequences Are Not Enough.** *PLoS Biol.*, 9(3), e1000602-, doi:10.1371/journal.pbio.1000602.

12. Philippe, H., Derelle, R., Lopez, P., Pick, K., Borchiellini, C., Boury-Esnault, N. et al. (2009) **Phylogenomics Revives Traditional Views on Deep Animal Relationships.** *Curr. Biol.*, 19(8), 706-712, doi:10.1016/j.cub.2009.02.052.

13. Pick, K. S., Philippe, H., Schreiber, F., Erpenbeck, D., Jackson, D. J., Wrede, P. et al. (2010) **Improved phylogenomic taxon sampling noticeably affects nonbilaterian relationships.** *Mol. Biol. Evol.*, 27(9), 1983-1987, doi:10.1093/molbev/msq089.

14. Srivastava, M., Simakov, O., Chapman, J., Fahey, B., Gauthier, M. E. A., Mitros, T. et al. (2010) **The *Amphimedon queenslandica* genome and the evolution of animal complexity.** *Nature*, 466(7307), 720-726, doi:10.1038/nature09201.

15. Moroz, L. L., Kocot, K. M., Citarella, M. R., Dosung, S., Norekian, T. P., Povolotskaya, I. S. et al. (2014) **The ctenophore genome and the evolutionary origins of neural systems.** *Nature*, 510(7503), 109-114, doi:10.1038/nature13400.

16. Ryan, J. F., Pang, K., Schnitzler, C. E., Nguyen, A. D., Moreland, R. T., Simmons, D. K. et al. (2013) **The Genome of the Ctenophore *Mnemiopsis leidyi* and Its Implications for Cell Type Evolution.** *Science*, 342(6164), 1242592-1242592, doi:10.1126/science.1242592.

17. Sperling, E. A., Peterson, K. J., Pisani, D. (2009) **Phylogenetic-Signal Dissection of Nuclear Housekeeping Genes Supports the Paraphyly of Sponges and the Monophyly of Eumetazoa.** *Mol. Biol. Evol.*, 26(10), 2261-2274, doi:10.1093/molbev/msp148.

18. Whelan, N. V., Kocot, K. M., Halanych, K. M. (2015) **Employing Phylogenomics to Resolve the Relationships among Cnidarians, Ctenophores, Sponges, Placozoans, and Bilaterians.** *Integr. Comp. Biol.*, doi:10.1093/icb/icv037.

19. Debrenne, F., Zhuravlev, A. Y., Kruse, P. D. (2002) **Class Archaeocyatha Bornemann, 1884.** In: *Systema Porifera*, edited by Hooper, J. N. A., Van Soest, R. W. M., Willenz, P., Boston, MA: Springer US; 1539-1699.

20. Debrenne, F., Reitner, J. (2001) **Sponges, cnidarians, and ctenophores.** In: *The ecology of the Cambrian radiation*, edited by Zhuravlev, A. Y., Riding, R., New York: Columbia University Press; 301-325.

21. Rowland, S. M. (2001) **Archaeocyaths: a history of phylogenetic interpretation.** *Journal of Paleontology*, 75(6), 1065-1078, doi:10.1017/S0022336000017133.

22. Kerner, A., Debrenne, F., Vignes-Lebbe, R. (2011) **Cambrian archaeocyathan metazoans: revision of morphological characters and standardization of genus descriptions to establish an online identification tool.** *Zookeys*, (150), 381-395, doi:10.3897/zookeys.150.1566.

23. Reitner, J., Wörheide, G., Lange, R., Thiel, V. (1997) **Biomineralization of calcified skeletons in three Pacific coralline demosponges - an approach to the evolution of basal skeletons.** *Cour. Forsch-Inst. Senckenberg* 201, 371-383.

24. Reitner, J., Wörheide, G. (2002) **Non-Lithistid Fossil Demospongiae - Origins of their Palaeobiodiversity and Highlights in History of Preservation.** In: *Systema Porifera: A*

Guide to the Classification of sponges, edited by Hooper, John, van Soest, R. W. M., New York: Springer; 52-68.

25. Wörheide, G. (2008) **A hypercalcified sponge with soft relatives: *Vaceletia* is a keratose demosponge.** *Mol. Phylogenet. Evol.*, 47(1), 433-438, doi:10.1016/j.ympev.2008.01.021.

26. Erpenbeck, D., Voigt, O., Wörheide, G., Lavrov, D. V. (2009) **The mitochondrial genomes of sponges provide evidence for multiple invasions by Repetitive Hairpin-forming Elements (RHE).** *BMC Genomics*, 10, 591, doi:10.1186/1471-2164-10-591.

27. Morrow, C., Cárdenas, P. (2015) **Proposal for a revised classification of the Demospongiae (Porifera).** *Front. Zool.*, 12(1), 7, doi:10.1186/s12983-015-0099-8.

28. Vacelet, J. (1977) **Une nouvelle relique du Secondaire: un représentant actuel des Eponges fossiles Sphinctozoaires.** *Comptes Rendus De L'Academie Des Sciences Paris* (série D) 285, 509-511.

29. Wörheide, G., Reitner, J. (1996) **"Living fossil" sphinctozoan coralline sponge colonies in shallow water caves of the Osprey Reef (Coral Sea) and the Astrolabe Reefs (Fiji Islands).** In: *Göttinger Arbeiten zur Geologie und Palaeontologie*, edited by Reitner J, Neuweiler F, F, G., Göttingen: 145-148.

30. Reitner, J., Wörheide, G., Lange, R., Schumann-Kindel, G. (2001) **Coralline demosponges, a geobiological portrait.** *Bull. Tohoku Univ. Museum* 1, 229-235.

31. Vacelet, J. (2002) **Recent 'Sphinctozoa', Order Verticillitida, Family Verticillitidae Steinmann, 1882.** In: *Systema Porifera: A Guide to the Classification of Sponges*, edited by Hooper, J. N. A., Van Soest, R. W. M., New York: Springer; 1097-1098.

32. Jackson, D. J., Thiel, V., Wörheide, G. (2010) **An evolutionary fast-track to biocalcification.** *Geobiology*, 8(3), 191-196, doi:10.1111/j.1472-4669.2010.00236.x.

33. Jackson, D. J., Macis, L., Reitner, J., Wörheide, G. (2011) **A horizontal gene transfer supported the evolution of an early metazoan biomineralization strategy.** *BMC Evol. Biol.*, 11, 238, doi:10.1186/1471-2148-11-238.

34. Uriz, M. J., Agell, G., Blanquer, A., Turon, X., Casamayor, E. O. (2012) **Endosymbiotic calcifying bacteria: a new cue to the origin of calcification in metazoa?** *Evolution*, 66(10), 2993-2999, doi:10.1111/j.1558-5646.2012.01676.x.

35. Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J. et al. (2013) ***De novo* transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis.** *Nat. Protoc.,* 8(8), 1494-1512, doi:10.1038/nprot.2013.084.

36. Wiśniewski, J. R., Zougman, A., Nagaraj, N., Mann, M. (2009) **Universal sample preparation method for proteome analysis.** *Nat. Methods*, 6(5), 359-362,

doi:10.1038/nmeth.1322.

37. Rappsilber, J., Mann, M., Ishihama, Y. (2007) **Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips.** *Nat Protoc*, 2(8), 1896-1906, doi:10.1038/nprot.2007.261.

38. Michalski, A., Damoc, E., Hauschild, J. P., Lange, O., Wieghaus, A., Makarov, A. et al. (2011) **Mass spectrometry-based proteomics using Q Exactive, a high-performance benchtop quadrupole Orbitrap mass spectrometer.** *Mol Cell Proteomics*, 10(9), M111.011015, doi:10.1074/mcp.M111.011015.

39. Scheltema, R. A., Mann, M. (2012) **SprayQc: a real-time LC-MS/MS quality monitoring system to maximize uptime using off the shelf components.** *J Proteome Res*, 11(6), 3458-3466, doi:10.1021/pr201219e.

40. Cox, J., Mann, M. (2008) **MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification.** *Nat. Biotechnol.*, 26(12), 1367-1372, doi:10.1038/nbt.1511.

41. Cox, J., Neuhauser, N., Michalski, A., Scheltema, R. A., Olsen, J. V., Mann, M. (2011) **Andromeda: a peptide search engine integrated into the MaxQuant environment.** *J Proteome Res*, 10(4), 1794-1805, doi:10.1021/pr101065j.

42. Neuhauser, N., Michalski, A., Cox, J., Mann, M. (2012) **Expert system for computer-assisted annotation of MS/MS spectra.** *Mol Cell Proteomics*, 11(11), 1500-1509, doi:10.1074/mcp.M112.020271.

43. Schwanhäusser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J. et al. (2011) **Global quantification of mammalian gene expression control.** *Nature*, 473(7347), 337-342, doi:10.1038/nature10098.

44. Goujon, M., McWilliam, H., Li, W., Valentin, F., Squizzato, S., Paern, J. et al. (2010) **A new bioinformatics analysis tools framework at EMBL-EBI.** *Nucleic Acids Res.*, 38(suppl 2), W695-W699, doi:10.1093/nar/gkq313.

45. Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W. et al. (2011) **Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega.** *Mol Syst Biol*, 7, 539, doi:10.1038/msb.2011.75.

46. Hunter, S., Jones, P., Mitchell, A., Apweiler, R., Attwood, T. K., Bateman, A. et al. (2012) **InterPro in 2011: new developments in the family and domain prediction database.** *Nucleic Acids Res.*, D306-12, doi:10.1093/nar/gkr948.

47. Petersen, T. N., Brunak, S., von, H., Gunnar, Nielsen, H. (2011) **SignalP 4.0: discriminating signal peptides from transmembrane regions.** *Nature Methods*, 8(10), 785-786, doi:10.1038/nmeth.1701.

48. Sonnhammer, E. L. L. (1998) **A Hidden Markov Model for Predicting**

Transmembrane Helices in Protein Sequences. *Proc. of Sixth Int. Conf. on Intelligent Systems for Molecular Biology*, 175-182,

49. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M., Appel, R. et al. (2005) **Protein Identification and Analysis Tools on the ExPASy Server.** In: *The Proteomics Protocols Handbook*, edited by Walker, J., Totowa, NJ: Humana Press; 571-607.

50. Dosztányi, Z., Csizmok, V., Tompa, P., Simon, I. (2005) **IUPred: web server for the prediction of intrinsically unstructured regions of proteins based on estimated energy content.** *Bioinformatics*, 21(16), 3433-3434, doi:10.1093/bioinformatics/bti541.

51. Romeis, B. (1989) **Mikroskopische Technik; 17., neubearbeitete Auflage.** München, Wien, Baltimore: Urban & Schwarzenberg; 443.

52. Riesgo, A., Farrar, N., Windsor, P. J., Giribet, G., Leys, S. P. (2014) **The analysis of eight transcriptomes from all poriferan classes reveals surprising genetic complexity in sponges.** *Mol. Biol. Evol.*, 31(5), 1102-1120, doi:10.1093/molbev/msu057.

53. Marchler-Bauer, A., Bryant, S. H. (2004) **CD-Search: protein domain annotations on the fly.** *Nucleic Acids Res.*, W327-W331, doi:10.1093/nar/gkh454.

54. Finn, R. D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R. Y., Eddy, S. R. et al. (2014) **Pfam: the protein families database.** *Nucleic Acids Res.*, 42(D1), D222-D230, doi:10.1093/nar/gkt1223.

55. Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y. et al. (2015) **CDD: NCBI's conserved domain database.** *Nucleic Acids Res.*, 43(D1), D222-D226, doi:10.1093/nar/gku1221.

56. Mann, K., Jackson, D. J. (2014) **Characterization of the pigmented shell-forming proteome of the common grove snail *Cepaea nemoralis*.** *BMC Genomics*, 15, doi:10.1186/1471-2164-15-249.

57. Jackson, D. J., Macis, L., Reitner, J., Degnan, B. M., Wörheide, G. (2007) **Sponge paleogenomics reveals an ancient role for carbonic anhydrase in skeletogenesis.** *Science*, 316(5833), 1893-1895, doi:10.1126/science.1141560.

58. Voigt, O., Adamski, M., Sluzek, K., Adamska, M. (2014) **Calcareous sponge genomes reveal complex evolution of α-carbonic anhydrases and two key biomineralization enzymes.** *BMC Evol. Biol.*, 14, 230, doi:10.1186/s12862-014-0230-z.

59. Le Roy, N., Jackson, D. J., Marie, B., Ramos-Silva, P., Marin, F. (2014) **The evolution of metazoan alpha-carbonic anhydrases and their roles in calcium carbonate biomineralization.** *Front. Zool.*, 11(1), 1-16, doi:10.1186/s12983-014-0075-8.

60. Ramos-Silva, P., Kaandorp, J., Huisman, L., Marie, B., Zanella-Cléon, I., Guichard, N. et al. (2013) **The skeletal proteome of the coral *Acropora millepora*: the evolution of**

calcification by co-option and domain shuffling. *Mol. Biol. Evol.*, 30(9), 2099-2112, doi:10.1093/molbev/mst109.

61. Bhave, G., Cummings, C. F., Vanacore, R. M., Kumagai-Cresse, C., Ero-Tolliver, I. A., Rafi, M. et al. (2012) **Peroxidasin forms sulfilimine chemical bonds using hypohalous acids in tissue genesis.** *Nat Chem Biol*, 8(9), 784-790, doi:10.1038/nchembio.1038.

62. Feinstein, Y., Borrell, V., Garcia, C., Burstyn-Cohen, T., Tzarfaty, V., Frumkin, A. et al. (1999) **F-spondin and mindin: two structurally and functionally related genes expressed in the hippocampus that promote outgrowth of embryonic hippocampal neurons.** *Development*, 126(16), 3637-3648,

63. Palmer, G. D., Attur, M. G., Yang, Q., Liu, J., Moon, P., Beier, F. et al. (2014) **F-spondin deficient mice have a high bone mass phenotype.** *PLoS One*, 9(5), e98388, doi:10.1371/journal.pone.0098388.

64. He, Y. W., Li, H., Zhang, J., Hsu, C. L., Lin, E., Zhang, N. et al. (2004) **The extracellular matrix protein mindin is a pattern-recognition molecule for microbial pathogens.** *Nat. Immunol.*, 5(1), 88-97, doi:10.1038/ni1021.

65. Suzuki, M., Saruwatari, K., Kogure, T., Yamamoto, Y., Nishimura, T., Kato, T. et al. (2009) **An acidic matrix protein, Pif, is a key macromolecule for nacre formation.** *Science*, 325(5946), 1388-1390, doi:10.1126/science.1173793.

66. Ehrlich, H. (2010) **Chitin and collagen as universal and alternative templates in biomineralization.** *International Geology Review*, 52(7-8), 661-699, doi:10.1080/00206811003679521.

67. Falini, G., Fermani, S. (2004) **Chitin mineralization.** *Tissue engineering*, 10(1-2), 1-6, doi:10.1089/107632704322791646.

68. Junqua, S., Robert, L., Garrone, R., Pavans de Ceccatty, M., Vacelet, J. (1974) **Biochemical and morphological studies on collagens of horny sponges. *Ircinia* filaments compared to spongines.** *Connect. Tissue Res.*, 2(3), 193-203, doi:10.3109/03008207409152244.

69. Ehrlich, H., Maldonado, M., Spindler, K. D., Eckert, C., Hanke, T., Born, R. et al. (2007) **First evidence of chitin as a component of the skeletal fibers of marine sponges. Part I. Verongidae (demospongia: Porifera).** *J Exp Zool B Mol Dev Evol*, 308(4), 347-356, doi:10.1002/jez.b.21156.

70. Marin, F., Luquet, G. (2007) **Unusually acidic proteins in biomineralization.** In: *Handbook of biomineralization: Biological aspects and structure formation*, edited by Bäuerlein, E., Weinheim: Wiley; 273-290.

71. Wheeler, A. P., George, J. W., Evans, C. A. (1981) **Control of calcium carbonate nucleation and crystal growth by soluble matrix of oyster shell.** *Science*, 212(4501),

1397-1398, doi:10.1126/science.212.4501.1397.

72. Mann, K., Poustka, A. J., Mann, M. (2008) **In-depth, high-accuracy proteomics of sea urchin tooth organic matrix.** *Proteome Sci.*, 6, 33, doi:10.1186/1477-5956-6-33.

73. Ramos-Silva, P., Marin, F., Kaandorp, J., Marie, B. (2013) **Biomineralization toolkit: the importance of sample cleaning prior to the characterization of biomineral proteomes.** *Proc Natl Acad Sci U S A*, 110(24), E2144-6, doi:10.1073/pnas.1303657110.

74. Allemand, D., Tambutté, É., Girard, J.-P., Jaubert, P. (1998) **Organic Matrix Synthesis in the Scleratinian Coral *Stylophora pistillata*: Role in Biomineralization and Potential Target of the Organotin Tributyltin.** *J. Exp. Biol.*, 201(13), 2001-2009,

75. Fang, D., Pan, C., Lin, H., Lin, Y., Xu, G., Zhang, G. et al. (2012) **Ubiquitylation functions in the calcium carbonate biomineralization in the extracellular matrix.** *PLoS One*, 7(4), e35715, doi:10.1371/journal.pone.0035715.

76. Rahman, M. A., Shinjo, R., Oomori, T., Wörheide, G. (2013) **Analysis of the proteinaceous components of the organic matrix of calcitic sclerites from the soft coral *Sinularia* sp.** *PLoS One*, 8(3), e58781, doi:10.1371/journal.pone.0058781.

77. Karlińska-Batres, K., Wörheide, G. (2013) **Microbial diversity in the coralline sponge *Vaceletia crypta*.** *Antonie Van Leeuwenhoek*, 103(5), 1041-1056, doi:10.1007/s10482-013-9884-6.

78. Henry, R. P. (1996) **Multiple Roles of Carbonic Anhydrase in Cellular Transport and Metabolism.** *Ann. Rev. Physiol*, 58, 523-538, doi:10.1146/annurev.ph.58.030196.002515.

79. Hewett-Emmett, D., Tashian, R. E. (1996) **Functional Diversity, Conservation, and Convergence in the Evolution of the alpha-, beta-, and gamma-Carbonic Anhydrase Gene Families.** *Mol. Phylogenet. Evol.*, 5(1), 50-77, doi:10.1006/mpev.1996.0006.

Chapter 4:

Investigating the expression of biomineralization gene candidates in *Vaceletia* sp. using *in situ* hybridization experiments

4.1 Introduction

Biomineralization is the ability of living organisms to construct a rigid skeleton. It is a widespread phenomenon among metazoans: representatives of most animal clades are able to fabricate and deposit biominerals [1]. During the Cambrian Explosion, beginning around 543 mya, most extant animal lineages appeared in a relatively short time window. This diversification coincides with a major increase of skeletal elements, which emerged almost simultaneously in many metazoan phyla, and which are well documented within the fossil record [2]. New ecological interactions and an increase of predation pressure facilitated the rapid increase of metazoan morphology and the evolution of protection armor. This suggests that the ability to biomineralize was one key factor that supported the Cambrian Explosion [2]. To obtain a complete picture of the rapid evolution of complex animals it is important to decipher the genetic and molecular mechanism that guide the assemblage of mineralized body parts. Because sponges are among the earliest branching metazoan taxa, and among the first animals represented in the fossil record to display a biologically controlled mode of biomineralization [2], it is crucial to include them when addressing the question of the origin and evolution of metazoan biomineralization. To date there is still little information about the genetic and molecular mechanisms that underlie the biomineralization processes of this important phylum.

The sponge *Vaceletia* is a so-called 'living-fossil' which first appears in the Middle Triassic [3]. This hypercalcifying taxon has been regarded as a monospecific genus with one described species, *Vaceletia crypta* [4]. However, it occurs in several different growth forms, colonial vs. solitary, which are likely to represent different species [5]. The taxon

Vaceletia is the only living representative that builds its skeleton in the so called 'sphinctozoan-like' way [3], which superficially resembles the skeleton of the first metazoan reef builder on earth, the Archaeocyathids (an extinct class of sponges), in some skeletal features. It has been proposed that *Vaceletia* might represent a modern archaeocyatha [3, 6] but as molecular data has shown that *Vaceletia* belongs to the Dictyoceratida within the Class Demospongiae [5] this is most likely not the case. Active mineralization in *Vaceletia* occurs in two different zones within the sponge: (i) in the uppermost part of the sponge where new chambers are produced resulting in sponge growth and (ii) within the older parts of the skeleton where existing chambers are subsequently mineralized resulting in building the hyper-calcified stalk [7].

We previously generated and characterized the transcriptome (chapter 2 [8]) and the skeleton forming proteome of the yet to be described colonial branching *Vaceletia* sp. (chapter 3 [9]). We could identify numerous skeletal forming protein candidates of the head and the stalk region of the sponge, including proteins that are known to play a role in biomineralization, such as carbonic anhydrase, as well as novel proteins that showed no similarity to known proteins [9]. Another coralline sponge, *Astrosclera willeyana*, employs its microbial community directly in the deposition of its skeleton by degrading a proportion of its microbial community via the autophagy pathway [10] and then using the organic residues as crystallization seeds [11]. In the sponge genus Hemimycale calcification is mediated by intracellular calcibacteria equipping the sponge surface with a rudimentary peripheral skeleton [12]. Given that *Vaceletia* sp. is the host of an abundant and diverse microbial community [8] we reasoned that it is highly probable that bacteria might also play a role in the fabrication of the sponges' skeleton. Transcriptome analysis showed that *Vaceletia* sp. very likely interacts in a variety of ways with its diverse and abundant microbial community (chapter 2 [8]). However, the proteomic survey revealed that bacteria apparently have no direct role in the biomineralization process of *Vaceletia* sp. (chapter 3 [9]).

The aim of this study was to further characterize the biomineralization processes in *Vaceletia* sp. by visualizing the spatial expression of the known and novel biomineralization gene candidates obtained from proteome analysis (chapter 3 [9]) by using the technique of *in situ* hybridization (ISH). This method allows suggesting a function of a particular gene transcript based on the location of its expression. Applied to

Vaceletia sp. this means if a candidate gene is expressed in sponge or bacterial cells that are in intimate association with the active biomineralization site a role in this process is likely. A recent study on biomineralization components of the calcareous sponge *Sycon ciliatum* demonstrated how this method can help to assign different functions to genes involved in the skeletogenesis process [13]. However, adapting *in situ* hybridization to a non-model organism is challenging. There are various factors that critically affect the success of this methods and need to be optimized for each experimental application. Here I report preliminary results for whole mount and slide *in situ* hybridization in *Vaceletia* sp..

4.2 Material and Methods

4.2.1 Sample collection, fixation and RNA extraction

Specimens of *Vaceletia* sp. were collected in 2009 during the Deep Down Under Expedition (http://www.deepdownunder.de) at Osprey- and Bougainville Reefs (Coral Sea, Australia) in depths from 5 to 24 m by SCUBA diving. Samples for RNA extraction were preserved in RNAlater and stored at -20°C. Samples for whole mount *in situ* hybridization (WMISH) were fixed as described [14] and stored in 75 % EtOH at -20°C.

Before processing, samples were carefully inspected under a stereomicroscope for contaminating organisms, which were carefully removed. Total RNA extraction was performed using the miRNeasy Kit (Qiagen) according to the manufacturer's instructions. RNA was checked by gel electrophoresis on a 1.5 % TAE agarose gel and by spectrometry using a Nanodrop spectrometer.

4.2.2 Decalcification

Individual heads of *Vaceletia* sp. were rehydrated through a graded EtOH series into 1X phosphate-buffered saline (PBS) and incubated in a decalcification solution of 350 mM ethylenediaminetetraacetic acid (EDTA), 4 % of paraformaldehyde (PFA) and 1X PBS buffer at room temperature for as long as necessary (often over two nights). The decalcification solution was changed approximately every 12 hours. Decalcified samples were then washed three times in 1X PBS-Tween (PBTw). Individual heads used for WMISH were dissected in half in 1X PBTw. Individual heads used for slide *in situ* hybridization (SISH) were embedded in Paraffin, sectioned at approximately 5 µm and placed on polysine slides (Menzel J2800AMNZ). All slides were dewaxed with two

washes of xylene for 5 min each and then rehydrated using a graded EtOH series into 1X PBSTw.

4.2.3 PCR, cloning and riboprobe synthesis

3'RACE cDNA libraries were constructed using CDS and UPM primers (Table 1) specifically designed for *Vaceletia* sp. to isolate 3'RACE fragments. Gene specific primers designed to amplify 16S, actin, carbonic anhydrase and biomineralization candidate genes extracted from the skeleton of *Vaceletia* sp. (chapter 3 [9]) were designed from Illumina RNA-Seq data (Table 2). Fragments were amplified from pooled cDNA RACE-libraries from different specimens. The resulting PCR products were cloned into the pGEM-T Easy vector (Promega) containing T7 and SP6 primer sites and verified by Sanger sequencing using standard molecular methods and reagents. These fragments were then amplified from plasmid DNA using M13 primers and purified using the QIAquick Gel Extraction Kit (Qiagen) according to the manufacture's instruction. Complementary riboprobes were generated by reverse transcription in a 10 µl reaction containing 0.25-0.5 volume PCR template, 1X reverse transcription buffer, 10 mM Dithiothreitol, 1X Digoxigenin-RNA labeling Mix (Roche. #11277073910) and 20 Units of the appropriate RNA polymerase (SP6 or T7; Promega, #P108 or #P207). The reactions were incubated at 37°C for 2-4 hours, precipitated with 0.1 volumes of 3 M sodium acetate (pH 5.2) and 3 volumes of 100 % ethanol for 15 minutes and centrifuged at 16,000 RCF for 15 min. The resulting pellets were washed with 75 % ethanol, dried and re-suspended in 20 µl RNase free water at 55°C. Probes were quantified using a Nanodrop spectrometer and qualitatively assessed by agarose gel electrophoresis using 500 ng to 1 µg of probe denatured in 95 % formamide at 75°C for 10 min. Probes were used at concentrations ranging from 100-500 ng/ml.

Table 1. Set of RACE primers (5' to 3') designed for *Vaceletia* sp.

3'RACE CDS	GAAGGGAATCGAGTGAGGTTGAGCTTTTTTTGTTTTTTTTTCTTTTTTTTTTVN
Short UPM	CAGCTACTAGGTGCATGTCGTA
Long UPM	CAGCTACTAGGTGCATGTCGTAGAAGGGAATCGAGTGAGGTTGAG

4.2.4 Proteinase K treatment

Most of the WMISH and SISH samples were briefly permeabilized with 10-30 µg/ml Proteinase K (Carl Roth, #7528) for 5 to 10 min at room temperature followed by two 5 min washes in 2 mg/ml glycine to stop the Pro-K activity. Glycine was washed out with three 5 min washes of PBTw.

Table 2. Primer sequences used to isolate gene fragments for riboprobe syntheses.

Gene	sequence of forward primer (5' to 3')	sequence of reverse primer (5' to 3')
16 S (27F/1492R)	AGAGTTTGATCMTGGCTCAG	CGGTTACCTTGTTACGACTT
Vsp_actin	GGAATGTGTAAGGCTGGCTTTG	TCTGTCTGCAATTCCTGGGTAC
carbonic anhydrase		
Vsp_CA_1	CAGCGATCCAGAAACGTGGAAGG	AACCATTGCCTTGCTGGTACATGC
Vsp_CA_2	CCGCCTGAAAGAACACCCCAGA	ACGAAGTGCAGTTCACCGGAAA
Vsp_CA_3	CGTGGAAGGAAGTTGAGGAGTGG	TCGATTGTGTGTTCTGATCCAAGCTC
Vsp_CA_4	CGGGCGTGGCTATTAATGTGGT	GCATCTCCATCAGTAATACTTCCAGTG
Vsp_CA_5	GTTCAAGCGGCTGAGGAAAA	AACCGAATGTCCGTTGTTGG
Vsp_CA_6	AGCCGCCTGAAAGAACACCCCA	TCCATATCTTCCTACTCCACTGGCCT
Vsp_CA_7	GCGTTCAAGCGGCTGAGGAAAA	ACCACTCCTCCACCCTCCTCTTGA
Vsp_C7761_1	CCCGATGGCTGGAGATGCTCTT	AAAACACTGCTGGCACCCTGAT
Vsp_C7761_2	TGGGGAGATAGCAACGAAGAAGGA	CCAAAGGGTGCTCTTATGCTCT
Vsp_C7761_2	AGGCGGCCTTCTAGTCAGTGA	AAATGCAAATGAATAACCTAAGTAGCT
Biomineralization candidates		
Vsp_C99840	AGCTTTCAATTTGATTCAGTTGGTGA	CAATTGGGACTGGTATATTGCC
Vsp_C53634	TTTGACATTGGAGTTGACAGTG	TGGTGCTGGTTCATCATTGTCT
Vsp_C36962_1	TGCAGTCAGACTAGTAGGAGGCAGT	CTTGACCAAAGCCAGCTGATCT
Vsp_C36962_2	ACAATGAGGGTAGAGTGGAGGT	CATCTTGAATGTGACTGCAACT

4.2.5 Hybridization and antibody binding

After rinsing with PBTw, samples were transferred into a 1 % (v/v) triethanolamine solution (TEA) and incubated for five minutes. This step was repeated once with the 1 % TEA solution and twice with a 1 % TEA + 0.3 % acetic anhydride solution. Samples were then washed three times with PBTw, post-fixed in 4 % PFA in PBTw for 10 min at room temperature and washed again 3 times in PBTw. After this post-fixation step, all samples were transferred into an InsituPro robot (Intavis) for the following hybridization, antibody incubation and washing steps. Samples were incubated with hybridization buffer for 15 min at room temperature before being brought to the intended hybridization temperature of 55 °C and incubated for additional two hours. Hybridization buffer was then replaced by 100 ng or 500 ng riboprobe in hybridization buffer and incubated for 16 hours at 55°C. Samples were then washed 3 times for 15 min with 4X wash buffer, 3 times for 15 min with 2X wash buffer, 3 times for 15 min with 1X wash buffer and once in 1X SSC with 0.1% Tween-20 for 15 min at 55 °C. After the washing steps, samples were allowed to cool down to room temperature and washed again with 1X SSC and 0.1% Tween for 15 min. The following steps were carried out at room temperature. Samples were washed in maleic acid buffer for 10 min and incubated twice for 1 hour and 45 min in 2 % block

solution before being incubated for 5 hours in block solution containing 1:X anti-DIG antibody conjugated to Alkaline Phosphatase (Roche, #11093274910). Unbound antibody was washed out by 15 rinses with PBTw for each 10 min. Samples were then removed from the robot and color development were conducted manually.

4.2.6 Color development

WMISH samples were washed twice in 1X alkaline phosphatase buffer with 0.1% Tween-20 (APTw) for 10 minutes each. Detection buffer was applied and color development was performed in the dark. The reaction was stopped with two washes of PBS. Samples were postfixed in 4% PFA in PBS over night at 4 °C. After fixation, samples were washed with 1X PBS and deionized H_2O followed by dehydration through a graded EtOH series. For SISH samples, slides and counter-slides were disassembled in a PBTw bath and washed twice in APTw for 10 minutes each. Minimal volume of detection buffer was applied and color development was performed in the dark. The reaction was stopped with two washes of PBS. Samples were postfixed in 4% PFA in PBS over night at 4°C. After fixation, samples were washed with 1X PBS and deionized H_2O followed by dehydration through a graduated EtOH series. Slide samples were mounted and counterstained with Roti-Mount FluorCare DAPI (Roth HP20.1).

4.2.7 Imaging

WMISH samples were imaged in a petri dish under a Zeiss stereo Discovery V8 microscope. Slides were photographed using Zeiss Axio Imager Z1 microscope running Zeiss camera software Axio Rel.4.8.

4.3 Results and Discussion

4.3.1 Isolation of biomineralization candidate genes

Standard primers used to generate RACE libraries are not suitable for *Vaceletia* sp.

Our recent study of the skeletal proteome of *Vaceletia* sp. has identified 40 proteins that most likely represent the key components the sponge employs in its mineralization process (chapter 3 [9]). In this study, some spatial resolution was achieved by dividing *Vaceletia*'s skeleton in head and stalk regions: approximately 50 % of the proteins differed in their abundance within the head and stalk regions (chapter 3 [9]). The most abundant proteins of this survey were used to design gene-specific primers (Table 2) with the goal to isolate

gene specific 3'RACE fragments. Initially, standard CDS (3'RACE CDS primer) and the anchor primer UPM (universal primer mix) were used to construct a cDNA 3'RACE library. 3'RACE PCR resulted in amplifying gene fragments that were also present in the UPM negative control. PCR optimization including the use of different polymerases, different template concentrations and different PCR conditions did not improve the specificity of the reaction. Both sequencing the unspecific product in the control reactions and blasting (blastN) the CDS and UPM primer sequences against the *Vaceletia* sp. transcriptome concordantly showed that several contigs of *Vaceletia* share significant similarity with the primer sequences used to construct the cDNA 3'RACE library. Therefore, it was necessary to design a new set of RACE primers (CDS, UPM long and UPM short) for *Vaceletia* sp. that showed no similarity to any of *Vaceletia*'s contigs (Table 1). The functionality of the new primer set was positively tested on the gastropod *Lymnea stagnalis*. The new primer set successfully resolved the problem of amplifying unspecific gene fragments in the UPM control reactions. However, amplifying gene specific 3'RACE fragments was not successful despite extensive PCR optimizations. When using internal gene specific primers instead of RACE primers, the PCR reactions worked well. An explanation could be that the cDNA synthesis was not functioning and genomic DNA (gDNA) was carried over from all preceding steps which would explain the missing 3'RACE fragments and the amplification of internal gene fragments. To test whether a cDNA library is contaminated with gDNA, PCR can be performed using primers that span an intron or that bind to an intron-exon boundary. Such primers should amplify different length fragments in gDNA and cDNA or should not yield any product at all in the cDNA sample. However, no genomic data is available for *Vaceletia* impeding the design of such primers.

RNA quality plays an important role for downstream applications

Working with RNA can be technically difficult as RNA is prone to degradation and standard RNA extraction protocols are not optimized for unusual tissue types like sponge tissues [15]. *Vaceletia* most probably contains pigments since the head appears grey-blue in living animals, as well as secondary compounds, and has a high load of symbiotic bacteria. These are all factors that could hinder RNA extraction, influence quantity and quality of extracted RNA and influence the performance of downstream applications [15]. Already the method of tissue preservation (liquid nitrogen or RNAlater) together with the time of storage can influence the purity and integrity of total RNA in demosponge tissues

[15, 16]. The quality and quantity of RNA extracted from *Vaceletia* was tested by spectrometry using a Nanodrop and by gel electrophoresis. The ratio absorbance of 260/280 and 260/230 are used to assess the purity of RNA. Expected values for pure RNA are approximately 2.0 and 2.0-2.2, respectively. The 260/280 ratio were in an acceptable range for *Vaceletia* but 260/230 ratios were consistently lower indicating the presence of contaminants such as polyphenols, polysaccharides, buffers or other impurities. Visual quality assessment via gel electrophoresis resulted in bands with smearing. This normally hints at degraded RNA, however sponges often show smearing when their RNA integrity is checked via gel electrophoresis but this does not necessarily mean that the RNA is degraded and that downstream applications are affected (pers. comment). For the extracted RNA of *Vaceletia* different RNA clean-up steps and/or serial dilution were performed before cDNA synthesis with the aim to get rid or dilute a potential inhibitor but had no positive effect on the outcome. Due to these difficulties, internal gene specific primers were used to produce riboprobes ranging from 180 bp to 820 bp.

4.3.2 *In situ* hybridization experiments

Whole mount *in situ* experiments with presumably highly expressed genes

Whole mount *in situ* hybridization as well as slide *in situ* hybridization were used with the goal to visualize the spatial expression of the biomineralization gene candidates in *Vaceletia* sp.. Since this was the first time ISH experiments were applied to *Vaceletia*, WMISH was performed against the bacterial 16S rRNA gene and actin. These genes were chosen because they should be abundantly expressed throughout the sponge tissue and give a strong signal. 16S riboprobes should bind to bacteria that are distributed throughout *Vaceletia* sp. mesohyl. Actin is a component of the cytoskeleton of eukaryotic cells. This gene should be expressed in sponge cells throughout the living tissue. WMISH against 16S showed an extensive staining of both, the living tissue and the organic residues of the previously hypercalcified stalk (Fig. 1 A,B). The magnification used did not allow to distinguish whether the signal within the tissue is produced by bacteria cells or sponge cells or is unspecific background. Sectioning of the sample to achieve a higher magnification did unfortunately fail due to experimental errors. WMISH against actin showed a staining of the organic residues of the previously calcified pillars (Fig. 1 D) and the previously hypercalcified stalk (Fig. 1 C,D) but no expression within the living tissue. This indicates a non-specific color reaction with the organic framework of the previously calcified parts of the sponge. Control experiments lacking riboprobe showed no non-

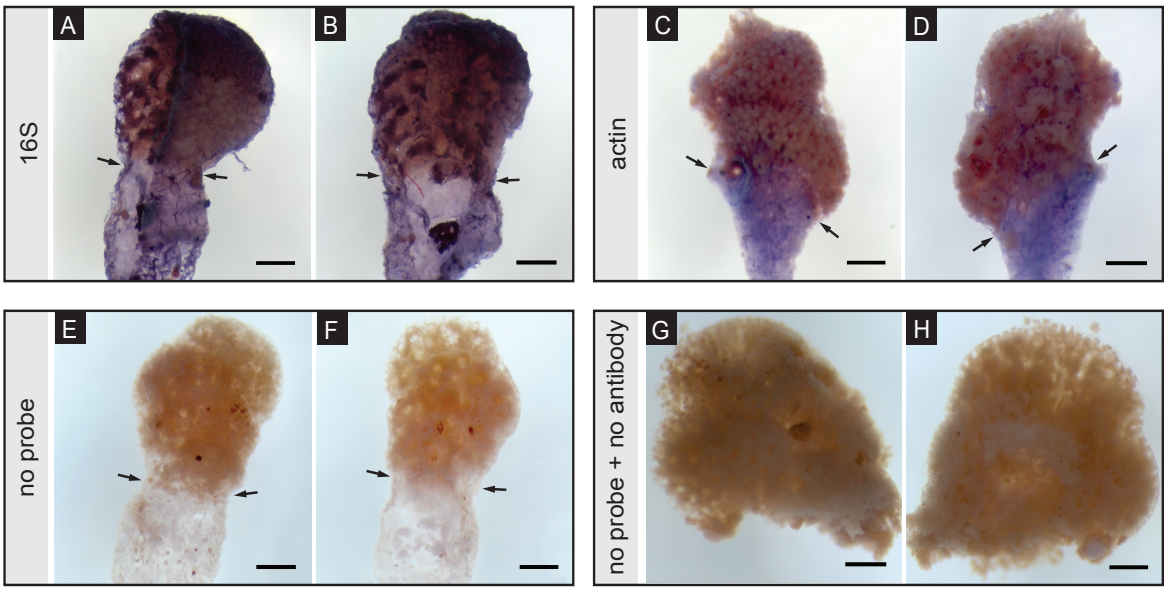

Figure 1. Decalcified, bisected heads of *Vaceletia* sp. used in WMISH. The arrows indicate the transition of the head region with living tissue to the previously hypercalcified stalk region without living tissue. **A** Lateral view of *Vaceletia* sp. outer surface and **B** inner surface following WIMSH against 16S showed extensive staining of the living tissue and the previously calcified organic framework of the skeleton. **C** Lateral view of *Vaceletia* sp. outer surface and **D** inner surface following WIMSH against actin showed staining of the previously calcified organic framework of the skeleton. **E-F** Control experiments lacking riboprobe and **G-H** control experiments lacking riboprobe and antibody showed no non-specific background staining. Scale bars = 500 µm.

specific color reaction (Fig. 1 E,F). Control experiments lacking riboprobe and antibody showed no staining of the head region (Fig. 1 G,H). All heads used in this control lacked the organic residue of the previously hypercalcified organic framework. The absent color reaction within the organic framework of the pillars indicates that the stalk would not show any staining either.

Slide *in situ* experiments with biomineralization candidates

To see details of the expression patterns (e.g. is a gene of interest expressed in sponge cells or bacteria cells) in WMISH experiments it is necessary to embed and section the samples after the WMISH procedure. An alternative method is the slide *in situ* hybridization where samples get sectioned before the *in-situ* hybridization procedure. This is the preferable method for scarce sample material as individual samples can be used for many different genes of interest in comparison to the whole mount approach. Slide *in situ* experiments were performed on *Vaceletia* sp. with the biomineralization candidate genes. Genes with a role in skeleton formation should be expressed in sponge and/or bacterial cells that are intimately associated with the active biomineralization site. In the case of *Vaceletia* an expression of biomineralization-associated genes would be expected in the outermost part

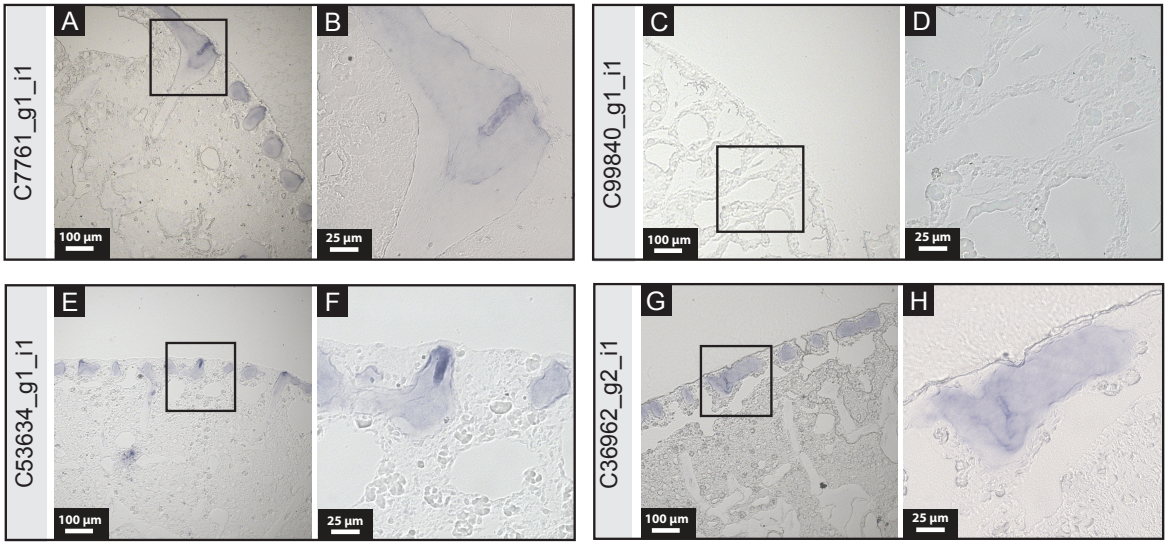

Figure 2. Overview of the *in situ* background signals produced by different biomineralization gene candidates. A,C,E,G All slides show that background staining if occurring were localized outside of the living tissue on the organic framework of the previously calcified pillars of the sponge. **B,D,F,H** Magnifications of the boxed sections.

of the sponge where new chambers are constructed and in the older parts of the sponge where the chambers get subsequently mineralized. Contig C7761 shows significant similarities to Astrosclerins, a family of α-CAs previously identified in the coralline sponge *Astrosclera willeyana*, that is directly involved in the biomineralization process of this sponge [9, 14]. It is by far the most abundant component of *Vaceletia* sp. proteome in both head and stalk regions [9]. Therefore, it displays an ideal candidate for *in situ* hybridization as it should be highly expressed. Expressions of neither contig C7761 nor one of the other tested biomineralization candidate genes identified in the proteome survey were detected in preliminary SISH-experiments. Only background staining or no staining was observed (Fig. 2).

Control experiments on *Vaceletia* sections lacking riboprobe demonstrated non-specific background staining (Fig. 3). Counterstaining with DAPI (binding to the DNA of sponge and bacteria cells) showed that the color reaction occurred in areas without living sponge tissue (Fig 3 B, D, F). The background staining likely reflects the location of previously calcified pillars and was more pronounced in the outermost part of the head (Fig. 3 A). In this region, new chambers are build. In the older parts of the sponge background staining was located on the transition zone from head to stalk. In this area, existing chambers get subsequently mineralized (Fig. 3 E). Here too, counterstaining with DAPI showed that the staining occurred outside of the living tissue (Fig. 3 F). These patterns imply a non-specific color reaction with the insoluble organic framework of the skeleton [9]. In addition, the

organic framework that was quite recently constructed, showed a higher affinity for non-specific staining.

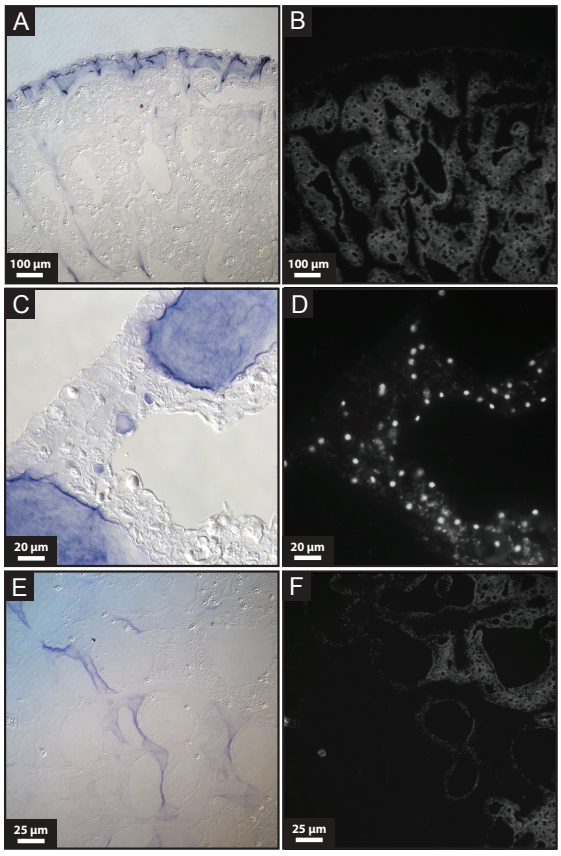

Figure 3. Control experiments lacking riboprobe on sagittal sectioned *Vaceletia* sp. individuals show non-specific background staining. A Overview of the head region illustrates non-specific background staining. **B** Counterstaining with DAPI reveals that the background staining is located outside of the sponge tissue. **C** Magnification of the more intense stained outermost head region reveals that the staining is only located within the organic framework of the previously calcified pillars and **E** within the previously calcified stalk region. **D-F** Counterstaining with DAPI shows again that the non-specific color reaction is not located within the sponge tissue. Dapi staining corresponds to sponge nuclei and bacteria nucleoid.

4.4 Conclusion

The successful outcome of ISH experiments depends on a variety of parameters including an appropriate Pro-K treatment to permeabilize tissue and unmask targets mRNA, appropriate prehybridization conditions to remove background and optimize the ISH signal intensity, and appropriate hybridization conditions to ensure specific and consistent ISH signal. For ISH experiments in *Vaceletia* sp. parameters were used which have been proven suitable for initial experiments in our laboratory's experience (e.g. hybridization temperature of 55 °C, probe concentrations of 100 and 500 ng/ml). The obtained results showed the absence of an ISH signal and extensive background staining of the skeletal organic framework. The background staining is most likely due to a nonspecific color reaction with components of the organic framework. Since background seems to be restricted to the residual organic framework of the sponge meaning it only occurs outside of the living tissue (Fig. 2 and 3), it might not interfere with a specific *in situ* hybridization signal occurring within sponge tissue. Nevertheless, the absence of an ISH signal

demonstrates that this technique needs to be optimized for *Vaceletia*. Because sample material was limited and the optimization of this technique for a novel organism usually requires considerate amount of material, ISH-optimization was beyond the scope of this thesis. However, to gain insight into the different functions of the biomineralization genes of *Vaceletia* and to understand how they are expressed and work together to construct such an elaborate skeleton, it would be highly beneficial to establish this method for *Vaceletia* in the future.

4.5 Outlook

Working with non-model organisms like *Vaceletia* is challenging as many standard protocols are not optimized for unusual tissue types. In order to achieve satisfactory and consistent results with standard applications such as RNA extraction, RACE-PCR and *in situ* hybridization it is necessary to ensure that the source material is of the best quality possible as well as available in sufficient amounts. The access to fresh *Vaceletia* sample material is very limited, thus it might be worth to initially identify the most adequate preservation method for fresh samples as this already might have an impact on quantity and quality of extractable RNA that is used in many downstream applications [15]. With enough material at hand, it would be possible to systematically try out different RNA extraction and clean-up methods in order to yield high quality RNA without contaminants for cDNA library construction. This could help to determine what causes the potential failure of cDNA synthesis and/or failure of RACE PCRs. DNase treatment during RNA extraction or afterwards would give information about if the internal gene fragments that were amplified from the RACE cDNA libraries are false positives due to gDNA contamination and will help to determine if cDNA synthesis is successful. With a working RACE-PCR, it would be possible to amplify longer stretches or even full lengths sequences of the biomineralization candidate genes and generate riboprobes of adequate length for *in situ* hybridization experiments. Likewise, ISH experiments need a systematically conducted optimization to adjust the different parameters to produce a satisfactory ISH signal. So far only one hybridization temperature was tested. Already subtle changes of biochemical and cellular properties of different tissue and even between tissue types and developmental stages might interfere with a successful signal detection [17] and require optimization. The tissue of *Vaceletia* contains high amounts of symbiotic bacteria [3, 8] which probably produce extrapolymeric substances [3]. It would be

interesting to see if a meaningful and satisfactory ISH signal could be produced despite the presences of the background staining. If not, it would be necessary to find out what causes the background and to apply different chemical and/or enzymatically treatments in order to overcome this problem.

References

1. Lowenstam, H. A., Weiner, S. (1989) **On Biomineralization.** Oxford University Press; 1-324.

2. Knoll, A. H. (2003) **Biomineralization and evolutionary history.** *Rev. Mineral. Geochem.*, 54(1), 329-356, doi:10.2113/0540329.

3. Reitner, J., Wörheide, G., Lange, R., Thiel, V. (1997) **Biomineralization of calcified skeletons in three Pacific coralline demosponges - an approach to the evolution of basal skeletons.** *Cour. Forsch-Inst. Senckenberg* 201, 371-383.

4. Vacelet, J. (1977) **Une nouvelle relique du Secondaire: un représentant actuel des Eponges fossiles Sphinctozoaires.** *Comptes Rendus De L'Academie Des Sciences Paris* (série D) 285, 509-511.

5. Wörheide, G. (2008) **A hypercalcified sponge with soft relatives: *Vaceletia* is a keratose demosponge.** *Mol. Phylogenet. Evol.*, 47(1), 433-438, doi:10.1016/j.ympev.2008.01.021.

6. Reitner, J., Wörheide, G. (2002) **Non-Lithistid Fossil Demospongiae - Origins of their Palaeobiodiversity and Highlights in History of Preservation.** In: *Systema Porifera: A Guide to the Classification of sponges*, edited by Hooper, John, van Soest, R. W. M., New York: Springer; 52-68.

7. Gautret, P., Reitner, J., Marin, F. (1996) **Mineralization events during growth of the coralline sponges *Acanthochaetetes* and *Vaceletia*.** *Bulletin de l'Institut océanographique*, 325-334,

8. Germer, J., Cerveau, N., Jackson, D. J. (2017) **The Holo-Transcriptome of a Calcified Early Branching Metazoan.** *Front Mar Sci*, 4, 81, doi:10.3389/fmars.2017.00081.

9. Germer, J., Mann, K., Wörheide, G., Jackson, D. J. (2015) **The skeleton forming proteome of an early branching metazoan: a molecular survey of the biomineralization components employed by the coralline sponge *Vaceletia* sp.** *PLoS One*, 10(11), e0140100, doi:10.1371/journal.pone.0140100.

10. Jackson, D. J., Wörheide, G. (2014) **Symbiophagy and biomineralization in the „living fossil" *Astrosclera willeyana*.** *Autophagy*, 10(3), 408-415, doi:10.4161/auto.27319.

11. Jackson, D. J., Thiel, V., Wörheide, G. (2010) **An evolutionary fast-track to biocalcification.** *Geobiology*, 8(3), 191-196, doi:10.1111/j.1472-4669.2010.00236.x.

12. Uriz, M. J., Agell, G., Blanquer, A., Turon, X., Casamayor, E. O. (2012) **Endosymbiotic calcifying bacteria: a new cue to the origin of calcification in metazoa?** *Evolution*, 66(10), 2993-2999, doi:10.1111/j.1558-5646.2012.01676.x.

13. Voigt, O., Adamska, M., Adamski, M., Kittelmann, A., Wencker, L., Wörheide, G. (2017) **Spicule formation in calcareous sponges: Coordinated expression of biomineralization genes and spicule-type specific genes.** *Sci. Rep.*, 7, 45658, doi:10.1038/srep45658.

14. Jackson, D. J., Macis, L., Reitner, J., Degnan, B. M., Wörheide, G. (2007) **Sponge paleogenomics reveals an ancient role for carbonic anhydrase in skeletogenesis.** *Science*, 316(5833), 1893-1895, doi:10.1126/science.1141560.

15. Riesgo, A., Pérez-Porro, A. R., Carmona, S., Leys, S. P., Giribet, G. (2012) **Optimization of preservation and storage time of sponge tissues to obtain quality mRNA for next-generation sequencing.** *Mol. Ecol. Resour.*, 12(2), 312-322, doi:10.1111/j.1755-0998.2011.03097.x.

16. Simister, R. L., Schmitt, S., Taylor, M. W. (2011) **Evaluating methods for the preservation and extraction of DNA and RNA for analysis of microbial communities in marine sponges.** *J. Exp. Mar. Biol. Ecol.,* 397(1), 38-43, doi:10.1016/j.jembe.2010.11.004.

17. Jackson, D. J., Herlitze, I., Hohagen, J. (2016) **A Whole Mount *In Situ* Hybridization Method for the Gastropod Mollusc *Lymnaea stagnalis*.** (109), e53968, doi:10.3791/53968.

Chapter 5:

General Discussion

Sponges are among the earliest evolving metazoans and they still represent ecologically important members of benthic communities throughout the world [1]. Over the last decades sponges have become the center of attention for researchers from many different disciplines. Although their simple body plan lacks complex structures such as nerves, muscles and organs, sponges possess a remarkably complex gene repertoire [2–4] that is still far from being fully explored. Additionally, many sponges harbor dense and diverse microbial communities [5] and we are only at the beginning of being able to reveal and understand the functional and ecological roles of these microbial communities, which are, for example, the source of many bioactive compounds with important pharmaceutical and biotechnological significance [6]. Another interesting aspect of the biology of sponges is their ability to secrete a mineralized skeleton. Biomineralization was a major innovation during the evolution of the late Precambrian/early Cambrian for Metazoans [7]. The fossil record shows that sponges are among the first multicellular animals to display this ability [8]. This feature together with their basal branching position make sponges an informative phylum for uncovering the origin of biomineralization in metazoans.

5.1 Challenges of sponge research

High throughput analyses are becoming faster and cheaper, facilitating the generation of sequence data of non-model organisms such as sponges. These techniques significantly contributed to both our understanding of sponge biology and sponge symbiosis. However, even though genetic information on sponges is increasing, functional studies are still missing. *In silicio* characterization of transcriptomes are based on assumptions inferred from sequence similarities. Sponge transcriptome annotations based on these sequence similarity searches as performed for *Vaceletia* sp. in this thesis (chapter 2, [9]) often contain a significant proportion of sequences that show no similarity to known sequences [4, 9]. This demonstrates the difficulties of trying to infer gene identity and function in

sponges from traditional model organisms that are phylogenetically distant from these [10]; additionally it highlights the lack of curated database entries for sponges. For sequences showing homology to known genes in other phyla, the functional role is inferred but often not experimentally validated. For this reason, sponge research has reached a point where it is crucial to move from *in silico* based studies to meaningful experimental studies to broaden and validate what we think we know about these intriguing animals.

However, it can be quite challenging to work with sponges for several reasons. Like *Vaceletia*, many other sponges occur in remote locations which complicate the access to fresh sample material. Sometimes even standard methods, such as RACE-PCR or *in situ* hybridization experiments (chapter 4) are challenging and require optimization, which is more difficult when sample material is limited. It is still difficult to keep sponges over long periods of time in aquaria, and many sponge symbionts identified through molecular techniques are resistant to cultivation [11, 12]. However, in order to assess reliable and reproducible experiments it would be necessary to develop functional tools that allow the cultivation and manipulation of both the sponges and their microbial community. This would allow a holistic approach to the investigation of the diverse aspects of sponge biology.

Nevertheless, as discussed in the following paragraphs, the results of this thesis provide valuable new insights into the complex gene repertoire of the hypercalcifying demosponge *Vaceletia* sp.. They expand our knowledge on how this sponge interacts with its microbial community, highlight the importance of metabolic interactions between sponges and their microbial endobionts, and give new insights into the biomineralization strategy employed by *Vaceletia*. By comparing the results to other sponge transcriptomic and genomic data it is possible to observe common features or differences, which in turn helps to interpret the results in a broader context.

5.2. Interactions between *Vaceletia* sp. and its microbial community

5.2.1 *Vaceletia*'s immune system and eukaryotic-like proteins from bacteria regulate host-microbe interactions

Many mechanisms of the sponge-microbe symbiosis are still not understood. Recent studies suggest that among other factors, the sponge immune system plays an important

role in mediating microbial recognition [13–15]. *Vaceletia*, and all other investigated sponges to date, have an almost complete Toll-like receptor signaling pathway, suggesting that this pathway plays an important role in the immune response. It is not entirely clear how sponges recognize microbes (see chapter 2, Figure 6 [9]). *Amphimedon queenslandica* possess numerous and diverse Nucleotide-binding domain and Leucine-rich repeat proteins (NLR), indicating an NLR-based immune system in this specific sponge. In *Vaceletia*'s transcriptome both, conventional Toll-like receptors (TLR) and conventional NLRs are missing (chapter 2 [9]), suggesting an alternative pattern recognition mechanism for microbes. To gain deeper insights into how sponges manage their relationship with bacteria, functional assays are now needed to dissect the role that TLR pathway components play. The bacterial subset of *Vaceletia*'s transcriptome is enriched in eukaryotic-like proteins (chapter 2 [9]) and evidence is accumulating that these proteins are enriched in sponge symbionts, playing a role in host-microbe interactions [16]. However, the exact role of eukaryotic-like proteins is unknown and awaits the development of functional assays. Having a genetically traceable sponge model with the possibility of being able to change and manipulate the microbiome would allow the tracking of changes that occur at the molecular level when, for example, harmful bacteria are introduced or the microbial community is removed. This would help to validate the predicted function of the immune system, and other predicted symbiosis factors such as eukaryotic-like proteins, and help to discover unknown components that might mediate microbe recognition. Moreover, it would also help to decipher metabolic interactions that take place between sponges and their microbial community.

5.2.2 The importance of metabolic interactions between *Vaceletia* and its microbial community

Our transcriptome characterization of *Vaceletia* sp. shows evidence for manifold metabolic interaction between the sponge and its endosymbiotic community (chapter 2 [9]). It has been long suggested that a significant amount of short-chain fatty acids from sponges is actually produced by bacteria and that sponges use short-chain FA precursors to biosynthesize long-chain FAs [17, 18]. By studying the underlying molecular mechanisms of different lipid pathway components in *Vaceletia* sp. we could support these statements. Short-chain fatty acid biosynthesis pathway components in *Vaceletia* sp. are bacterially derived and thus produced by the associated bacteria. This suggests that the sponge relies on its bacterial community for the supply of short-chain fatty acids (chapter 2, [9]). Through a comparison to other sponge transcriptomes we could show that this dependency

seems to be more pronounced in sponges with an abundant microbial community (chapter 2 [9]). High microbial abundance sponges are often characterized by abundant mid-chain branched fatty acids (MBFAs). The occurrence of MBFAs has been positively linked to the presences of bacterial polyketide synthase genes and the presence of members of the candidate phylum Poribacteria in bacteria-rich sponges [19]. The presence of bacterial polyketide synthase genes in *Vaceletia* suggests that bacteria, potentially Poribacteria, are the producers of its abundant MBFAs (chapter 2 [9]). A role for MBFAs in establishing and maintaining symbiosis has been proposed, but their function is still unknown [20]. Short-chain fatty acids produced by its bacterial community are used by *Vaceletia* to synthesize endogenous saturated and unsaturated long-chain FAs via an active metazoan elongation and desaturation enzyme system (chapter 2 [9]). These interesting findings raise further questions: do bacterial symbionts present a final FA product for the sponge or does the sponge farm and then harvest its microbial community to obtain the short-chain FAs? If symbionts supply a final FA product, which symbionts are the suppliers? What are the consequences if symbionts as short-chain FA suppliers are removed? Studies exploring the potential metabolic capabilities of sponge symbionts are faced with similar questions. Although examples for potential metabolic interactions between sponges and their microbiomes are accumulating [13, 15] it is still difficult to link specific symbionts unambiguously to a functional role. Likewise, it is difficult to study the influence these metabolic interactions have on the sponge biology. Complementary studies combining sequencing approaches with physiological studies could help to validate the metabolic capabilities of sponge symbionts and the mechanisms of interactions with the sponge host.

5.3 The biomineralization strategy of *Vaceletia* sp.: new insights

5.3.1 Known biomineralization proteins and novel compounds are involved in the biocalcification process

It has been proposed that *Vaceletia* constructs its skeleton in an ancestral way, mineralization starts within a delimited space but is not under direct control of the sponge cells, representing a juncture from biologically induced biomineralization to biologically controlled biomineralization [22]. Five acidic matrix proteins were isolated from the skeleton which were proposed to control the biocalcification process [22]. In general, there is little information about biomineralization proteins in sponges, let alone how these proteins are expressed and control the construction of the complex architecture of sponge

skeletons. By generating a skeletal proteome of *Vaceletia* sp. we identified 40 proteins which most likely represent the majority of components playing an important role in the biomineralization process of *Vaceletia* (chapter 3 [21]). Among these are known biomineralization compounds, such as carbonic anhydrase, spherulin, very acidic proteins, and extracellular matrix proteins as well as novel compounds. These proteins presumably fulfill manifold functions during the mineralization process such as being the template upon mineralization starts, binding and delivering calcium to the mineralization sites, catalyzing the reaction of CO_2 and water to HCO_3^- that reacts with calcium to form calcium carbonate, and sculpturing the mineralized skeleton by guiding the crystal growth but also inhibiting it where it is necessary. The sophisticated architecture of *Vaceletia*'s skeleton is characterized by chambers terraced one upon another, with the ontogenetically youngest chambers located at the top of the animal. Chambers are composed of pillars, radial spines as reinforcement and the ceiling of the chamber which represents the floor for the following generation of chambers (Fig. 1 B,C and I, chapter 3 [21]). Constructing these defined structures requires exquisite biological control over skeletal formation. The results show that the biocalcification process in *Vaceletia* involves a variety of different proteins and is more complex than previously thought Apart from the mineralization site in the uppermost part of the sponge where new chambers are build resulting in the growth of the sponge, a second mineralization site is present. This site is located within the older parts of the sponge where ontogenetically older chambers are subsequently mineralized resulting in the hypercalcified stalk (Fig. 1 M, chapter 3 [21]). We show that chambers are mineralized in layers whereby not all chambers are mineralized entirely (Fig. 1 N and O, chapter 3 [21]). Approximately 50% of the 40 major proteins differ in their abundances within the head and the stalk region; five proteins were only detected in the head, suggesting that different mineralization processes are taking place in the head and the stalk region of *Vaceletia*.

5.3.2 No significant proteinaceous contribution of microbes to the skeletal formation

An earlier study on the biomineralization process of *Vaceletia* proposed that bacteria might be involved in the calcification process of the sponge [22]. This hypothesis was based on macromolecular analysis that suggested the presence of exopolymeric substances produced by bacteria are used as a template for calcification [22]. The involvement of symbiotic bacteria in the biomineralization process has been reported for several sponge species [23–25], and has implications on how bacteria might have supported the early evolution of

skeletons [24, 26]. Even though *Vaceletia* host abundant and diverse microbes, our proteome analysis of the skeleton of *Vaceletia* sp. contains no evidence that proteins produced by bacteria are directly involved in the biomineralization process of this sponge (chapter 3 [21]). The presence of spherulin in the skeleton of *Vaceletia* (showing similarities to bacterial sugar transporters) supports the hypothesis that this gene was delivered horizontally into the genome of a common ancestor of *Astrosclera willeyana, Vaceletia sp., Amphimedon queenslandica, Chondrilla nucula, Spongilla lacustris* and the hexactinellid *Aphrocallistes vastus* and was subsequently co-opted to a biomineralization role in *A. willeyana* and *Vaceletia* sp. (chapter 3 [21]). These finding do not exclude the possibility of microbes contributing via other metabolic pathways to the biomineralization process in *Vaceletia* sp..

5.3.2 The deep evolutionary history of some biomineralization components

Sponges in general are an informative phylum when it comes to understanding what genetic repertoire might have been present in the last common ancestor of all metazoans (LCAM) and whether this repertoire might have contributed to the evolution of the ability to biocalcify. By comparing *Vaceletia*'s skeletal proteome to other metazoan biomineralization proteomes we show that *Vaceletia* employs known biomineralization compounds such as carbonic anhydrase, extracellular matrix proteins and very acidic proteins. The presence of deeply conserved biomineralization components suggests that the LCAM indeed contributed some components to its descendants that supported the evolution of biocalcification. However, the diversity of skeletal proteins used by different organisms to construct a skeleton is astonishing [27] suggesting that lineage specific evolution is also an important factor that contributed to the evolution of biocalcification in disparate animal lineages.

5.3.4 Inferring gene function using *in situ* hybridization: potential and limitations

An important approach towards addressing the characterization of biomineralization components are *in situ* hybridization experiments. With this method, it is possible to visualize the spatial and temporal mRNA expression patterns throughout tissues and to infer a functional role to a given gene product. Recent studies on the formation of calcitic spicules of calcareous sponges exemplify how *in situ* hybridization can help to decipher different functions of biomineralization genes and highlight the importance of temporal regulation of gene expression [28, 29]. Calcareous sponges have different spicule types that differ in the number of their rays, and are produced by specialized cells called

sclerocytes. It was shown that some biomineralization genes "provide a common genetic ground pattern for the formation of all spicule types" [29] in the calcareous sponge *Sycon ciliatum* while others are spicule-type specific and determine the different shapes of the spicules [29]. In *Vaceletia* sp. it is not known which cells are responsible for the production of the skeleton but they are most probably intimately associated with the biomineralization site. With the aim of visualizing the spatial expression patterns of the biomineralization gene candidates isolated in the proteomic study, I conducted *in situ* hybridization experiments on *Vaceletia* sp.. Analogous to the study on calcareous sponges, these expression patterns could be used to identify cells that are involved in skeletogenesis, to infer different functions of the biomineralization genes and to show whether genes are expressed at different stages during *Vaceletia*'s skeletal construction. However, the results of the *in situ* hybridization experiments showed that this technique is challenging and that the procedure needs to be further optimized for *Vaceletia* before it can be used to generate meaningful results (chapter 4). Although *in situ* hybridization techniques are important to draw conclusions about the putative function of biomineralizing genes, one should keep in mind that these assumptions should be experimentally validated. To date, the possibility of doing that in sponges is severely limited as *in vivo* gene function assays, and often access to live animals, are lacking. Hence, the field of sponge biomineralization would also benefit from the establishment of new functional tools for these animals. A promising technique is the targeted genome editing method CRISPR/Cas9 (clustered regularly interspaced short palindromic repeats) [30]. This method could allow the editing of sponge genomes and the study of gene function *in vivo*. CRISPR/Cas9 would allow the knock out of entire biomineralization genes or domains that are associated with biomineralization, and monitor the effect on the biomineralization process and the resulting biomineral, and thus provide information about their function and their interplay in constructing a skeleton. However, it is still a long way to the establishment of CRISPR/Cas9 method in sponges. Nevertheless, this method is a promising tool for studying the mechanisms of biomineralization, the mechanisms of symbiosis, and the evolution of complex metazoan traits [10].

Conclusion

The data presented in this thesis provides new and valuable insights into the sponge-holobiont and the biomineralization strategy of *Vaceletia* sp.. It clearly shows that *Vaceletia* sp. possesses the genetic requirements necessary to interact with a diverse and abundant microbiome, as well as to coordinate the deposition of its intricate and elaborate skeleton. The skeletal proteome reported here most likely contains the majority of proteins playing a key role in the biocalcification process of *Vaceletia* sp.. Approximately 50 % of these proteins differ in their abundance in the head and stalk region, suggesting that different mineralization mechanisms take place in the head and the stalk. The microbial community of *Vaceletia* sp. apparently plays a minor role in skeletal formation: no proteins synthesized by symbiotic bacteria are directly involved in the biomineralization process. However, regarding metabolic processes, microbes play a crucial role by providing *Vaceletia* with short-chain and mid-chain branched fatty acids. Although sponges are not easy to work with, they are an informative phylum because of their early branching position on the animal tree of life and their ancient relationship with microbes. This study contributes to our understanding of the molecular complexity of sponges and paves the way to be able to answer important questions regarding, for example, the evolution of animal-microbe relationships and how animals evolve their ability to biomineralize. The development and establishment of novel methods for sponges will help to effectively advance our knowledge on sponge biology and symbiosis. Results like the ones reported in this thesis will provide target areas for future research.

References

1. Bell, J. J. (2008) **The functional roles of marine sponges.** *Estuar. Coast. Shelf. Sci.,* 79(3), 341-353, doi:10.1016/j.ecss.2008.05.002.

2. Nichols, S. A., Roberts, B. W., Richter, D. J., Fairclough, S. R., King, N. (2012) **Origin of metazoan cadherin diversity and the antiquity of the classical cadherin/β-catenin complex.** *PNAS*, (109), 13046-13051, doi:10.1073/pnas.1120685109.

3. Srivastava, M., Simakov, O., Chapman, J., Fahey, B., Gauthier, M. E. A., Mitros, T. et al. (2010) **The *Amphimedon queenslandica* genome and the evolution of animal complexity.** *Nature*, 466(7307), 720-726, doi:10.1038/nature09201.

4. Riesgo, A., Farrar, N., Windsor, P. J., Giribet, G., Leys, S. P. (2014) **The analysis of eight transcriptomes from all poriferan classes reveals surprising genetic complexity in sponges.** *Mol. Biol. Evol.*, 31(5), 1102-1120, doi:10.1093/molbev/msu057.

5. Taylor, M. W., Radax, R., Steger, D., Wagner, M. (2007) **Sponge-Associated microorganisms: evolution, ecology, and biotechnological potential.** *Microbiol. Mol. Biol. Rev.*, 71(2), 295-347, doi:10.1128/MMBR.00040-06.

6. Bibi, F., Faheem, M., Azhar, E. I., Yasir, M., Alvi, S. A., Kamal, M. A. et al. (2016) **Bacteria from marine sponges: A source of new drugs.** *Current Drug Metabolism*, 17,

7. Knoll, A. H. (2003) **Biomineralization and evolutionary history.** *Rev. Mineral. Geochem.*, 54(1), 329-356, doi:10.2113/0540329.

8. Debrenne, F., Zhuravlev, A. Y., Kruse, P. D. (2002) **Class Archaeocyatha Bornemann, 1884.** In: *Systema Porifera*, edited by Hooper, J. N. A., Van Soest, R. W. M., Willenz, P., Boston, MA: Springer US; 1539-1699.

9. Germer, J., Cerveau, N., Jackson, D. J. (2017) **The Holo-Transcriptome of a Calcified Early Branching Metazoan.** *Front Mar Sci*, 4, 81, doi:10.3389/fmars.2017.00081.

10. Pita, L., Fraune, S., Hentschel, U. (2016) **Emerging Sponge Models of Animal-Microbe Symbioses.** *Front. Microbiol.*, 7, 2102, doi:10.3389/fmicb.2016.02102.

11. Schippers, K. J., Sipkema, D., Osinga, R., Smidt, H., Pomponi, S. A., Martens, D. E. et al. (2012) **Chapter six - Cultivation of Sponges, Sponge Cells and Symbionts: Achievements and Future Prospects.** In: *Volume 62*, edited by Mikel A. Becerro, M. J. U., Manuel Maldonado and Xavier Turon, Academic Press; 273-337.

12. Hardoim, C. C. P., Cardinale, M., Cúgcio, A. C. B., Esteves, A. I. S., Berg, G., Xavier, J. R. et al. (2014) **Effects of sample handling and cultivation bias on the specificity of bacterial communities in keratose marine sponges.** *Front. Microbiol.*, 5, 611, doi:10.3389/fmicb.2014.00611.

13. Hentschel, U., Piel, J., Degnan, S. M., Taylor, M. W. (2012) **Genomic insights into the marine sponge microbiome.** *Nat. Rev. Microbiol.*, 10(9), 641-654, doi:10.1038/nrmicro2839.

14. Yuen, B., Bayes, J. M., Degnan, S. M. (2014) **The characterization of sponge NLRs provides insight into the origin and evolution of this innate immune gene family in animals.** *Mol. Biol. Evol.*, 31(1), 106-120, doi:10.1093/molbev/mst174.

15. Webster, N. S., Thomas, T. (2016) **The Sponge Hologenome.** *mBio*, 7(2), e00135-16, doi:10.1128/mBio.00135-16.

16. Fan, L., Reynolds, D., Liu, M., Stark, M., Kjelleberg, S., Webster, N. S. et al. (2012) **Functional equivalence and evolutionary convergence in complex communities of microbial sponge symbionts.** *Proc. Natl. Acad. Sci. USA*, 109(27), E1878-E1887, doi:10.1073/pnas.1203287109.

17. Gillan, F. T., Stoilov, I. L., Thompson, J. E., Hogg, R. W., Wilkinson, C. R., Djerassi, C. (1988) **Fatty acids as biological markers for bacterial symbionts in sponges.** *Lipids*, 23(12), 1139-1145, doi:10.1007/BF02535280.

18. Morales, R. W., Litchfield, C. (1977) **Incorporation of 1-14C-Acetate into C26 fatty acids of the marine sponge *Microciona prolifera*.** *Lipids*, 12(7), 570-576, doi:10.1007/BF02533383.

19. Hochmuth, T., Niederkrüger, H., Gernert, C., Siegl, A., Taudien, S., Platzer, M. et al. (2010) **Linking chemical and microbial diversity in marine sponges: possible role for Poribacteria as producers of methyl-branched fatty acids.** *ChemBioChem*, 11(18), 2572-2578, doi:10.1002/cbic.201000510.

20. Fieseler, L., Hentschel, U., Grozdanov, L., Schirmer, A., Wen, G., Platzer, M. et al. (2007) **widespread occurrence and genomic context of unusually small polyketide synthase genes in microbial consortia associated with marine sponges.** *Appl. Environ. Microbiol.*, 73(7), 2144-2155, doi:10.1128/AEM.02260-06.

21. Germer, J., Mann, K., Wörheide, G., Jackson, D. J. (2015) **The skeleton forming proteome of an early branching metazoan: a molecular survey of the biomineralization components employed by the coralline sponge *Vaceletia* sp.** *PLoS One*, 10(11), e0140100, doi:10.1371/journal.pone.0140100.

22. Reitner, J., Wörheide, G., Lange, R., Thiel, V. (1997) **Biomineralization of calcified skeletons in three Pacific coralline demosponges - an approach to the evolution of basal skeletons.** *Cour. Forsch-Inst. Senckenberg* 201, 371-383.

23. Jackson, D. J., Thiel, V., Wörheide, G. (2010) **An evolutionary fast-track to biocalcification.** *Geobiology*, 8(3), 191-196, doi:10.1111/j.1472-4669.2010.00236.x.

24. Garate, L., Sureda, J., Agell, G., Uriz, M. J. (2017) **Endosymbiotic calcifying bacteria across sponge species and oceans.** *Sci. Rep.*, 7, 43674-43674,

25. Uriz, M. J., Agell, G., Blanquer, A., Turon, X., Casamayor, E. O. (2012) **Endosymbiotic calcifying bacteria: a new cue to the origin of calcification in metazoa?** *Evolution*, 66(10), 2993-2999, doi:10.1111/j.1558-5646.2012.01676.x.

26. Jackson, D. J., Macis, L., Reitner, J., Wörheide, G. (2011) **A horizontal gene transfer supported the evolution of an early metazoan biomineralization strategy.** *BMC Evol. Biol.*, 11, 238, doi:10.1186/1471-2148-11-238.

27. Marin, F., Bundeleva, I., Takeuchi, T., Immel, F., Medakovic, D. (2016) **Organic matrices in metazoan calcium carbonate skeletons: Composition, functions, evolution.** *J. Struct. Biol.*, 196(2), 98-106, doi:10.1016/j.jsb.2016.04.006.

28. Voigt, O., Adamski, M., Sluzek, K., Adamska, M. (2014) **Calcareous sponge genomes reveal complex evolution of α-carbonic anhydrases and two key biomineralization enzymes.** *BMC Evol. Biol.*, 14, 230, doi:10.1186/s12862-014-0230-z.

29. Voigt, O., Adamska, M., Adamski, M., Kittelmann, A., Wencker, L., Wörheide, G. (2017) **Spicule formation in calcareous sponges: Coordinated expression of biomineralization genes and spicule-type specific genes.** *Sci. Rep.*, 7, 45658, doi:10.1038/srep45658.

30. Gaj, T., Gersbach, C. A., Barbas, C. F., III. **ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering.** *Trends Biotech.*, 31(7), 397-405, doi:10.1016/j.tibtech.2013.04.004.

Acknowledgements

I cordially thank my supervisor Daniel Jackson for giving me the opportunity to work in such an exciting research field, for sharing his knowledge with me and for the critical and constructive discussions. I would like to thank Joachim Reitner for giving me the opportunity to work in his lab and for his helpful advice. I thank Oliver Voigt, Nico Posnien, Alexander Schmidt and Gernot Arp for being members of my thesis committee. I gratefully acknowledge Luciana Macis, Wolfgang Dröse, Dorothea Hause-Reitner and Cornelia Conradt for competent technical support in the laboratory, histological experiments and scanning and transmission electron microscopy experiments.

I am indebted to all my colleges and friends that contributed to the success of this study. Sincere thanks are given to all my present and former members of the "Evolution of the Metazoa" group Jennifer Hohagen, Ines Herlitze, Luciana Macis, Nicolas Cerveau, Annette Geißler, Christine Berndmeyer and Susanne Affenzeller and to all my present and former colleagues of the Geobiology Christina Beimforde, Eva Sadowski, Leyla Seyfullah, Tim Leefmann and Jan Bauermeister for creating a pleasant working atmosphere, for constant support, scientific discussions and exchange of ideas. Thank you for the pleasant lunch and coffee breaks and for the refreshing walks through the nature of the north campus. Many of you have become true friends and I am very grateful to have you in my life.

Finally, for steady support and encouragement in every aspect I want to thank my family and friends, especially my parents, Sarah Richert and Fabian Willert. I am grateful for your unconditionally love, for listening, for cheering me up when it was necessary and for reminding me what is important in life. Without you I would not be where I am today.