

Characterization of a thaumarchaeal symbiont that drives incomplete nitrification in the tropical sponge *Ianthella basta*

Florian U. Moeller¹, Nicole S. Webster^{2,3}, Craig W. Herbold¹, Faris Behnam¹, Daryl Domman¹,
5 Mads Albertsen⁴, Maria Mooshammer¹, Stephanie Markert^{5,8}, Dmitrij Turaev⁶, Dörte Becher⁷,
Thomas Rattai⁶, Thomas Schweder^{5,8}, Andreas Richter⁹, Margarete Watzka⁹, Per Halkjaer Nielsen⁴,
and Michael Wagner^{1,*}

¹Division of Microbial Ecology, Department of Microbiology and Ecosystem Science, University of Vienna,
10 Austria.

²Australian Institute of Marine Science, Townsville, Queensland, Australia.

³ Australian Centre for Ecogenomics, School of Chemistry and Molecular Biosciences, University of
15 Queensland, St Lucia, QLD, Australia

⁴Center for Microbial Communities, Department of Chemistry and Bioscience, Aalborg University, 9220
Aalborg, Denmark.

20 ⁵Institute of Marine Biotechnology e.V., Greifswald, Germany

⁶Division of Computational Systems Biology, Department of Microbiology and Ecosystem Science,
University of Vienna, Austria.

25 ⁷Institute of Microbiology, Microbial Proteomics, University of Greifswald, Greifswald, Germany

⁸Institute of Pharmacy, Pharmaceutical Biotechnology, University of Greifswald, Greifswald, Germany

⁹Division of Terrestrial Ecosystem Research, Department of Microbiology and Ecosystem Science,
30 University of Vienna, Austria.

*Corresponding author: Michael Wagner, Department of Microbiology and Ecosystem Science,
Althanstrasse 14, University of Vienna, 1090 Vienna, Austria. Email: wagner@microbial-ecology.net

35 **Originality-Significance Statement**

Many marine sponges harbor symbiotic members of the *Thaumarchaeota*, but there is generally only indirect evidence available about their functional role within these filter-feeding animals. Furthermore, the specific adaptations of thaumarchaeal symbionts to their sponge hosts are incompletely understood. In this study, we thoroughly characterized a thaumarchaeal symbiont
40 residing in the reef sponge *Ianthella basta* and demonstrate by using a combination of molecular tools and isotope techniques, that it is the only ammonia-oxidizer in its host. In contrast to other sponges, *I. basta* does not contain nitrite-oxidizing microbes and thus excretes considerable amounts of nitrite. Furthermore, using metagenomics and metaproteomics we reveal important adaptations of this symbiont, that represents a new genus within the *Thaumarchaeota*,
45 and conclude that it most likely lives as a mixotroph in its sponge host.

Summary

Marine sponges represent one of the few eukaryotic groups that frequently harbor symbiotic members of the *Thaumarchaeota*, which are important chemoautotrophic ammonia-oxidizers in many environments. However, in most studies, direct demonstration of ammonia-oxidation by these archaea within sponges is lacking, and little is known about sponge-specific adaptations of ammonia-oxidizing archaea (AOA). Here, we characterized the thaumarchaeal symbiont of the marine sponge *Ianthella basta* using metaproteogenomics, fluorescence *in situ* hybridization, qPCR and isotope-based functional assays. “*Candidatus Nitrosospongia bastadiensis*” is only distantly related to cultured AOA. It is an abundant symbiont that is solely responsible for nitrite formation from ammonia in *I. basta* that surprisingly does not harbor nitrite-oxidizing microbes. Furthermore, this AOA is equipped with an expanded set of extracellular subtilisin-like proteases, a metalloprotease unique among archaea, as well as a putative branched-chain amino acid ABC transporter. This repertoire is strongly indicative of a mixotrophic lifestyle and is (with slight variations) also found in other sponge-associated, but not in free-living AOA. We predict that this feature as well as an expanded and unique set of secreted serpins (protease inhibitors), a unique array of eukaryotic-like proteins, and a DNA-phosphorothioation system, represent important adaptations of AOA to life within these ancient filter-feeding animals.

Introduction

65 Marine sponges (phylum *Porifera*) are among the most basal metazoan lineages (Simion *et al.*,
2017), with fossil records suggesting their evolutionary emergence more than 600 million years ago
(Love *et al.*, 2009). Sponges form a major part of the marine benthic fauna across the world's
oceans (Bell, 2008) where they mediate critical biogeochemical processes through their filtration of
immense volumes of seawater (Maldonado *et al.*, 2012; de Goeij *et al.*, 2013). Many sponges
70 harbor dense, diverse and species-specific communities of microbes (Hentschel *et al.*, 2012;
Thomas *et al.*, 2016), and these associations are often temporally and geographically stable (Luter
et al., 2010; Schmitt *et al.*, 2012; Astudillo-García *et al.*, 2017). Functional roles that have been
assigned to specific sponge-associated microorganisms include the provision of photosynthates
(Wilkinson, 1983) and fixed N₂ (Wilkinson, 1979) by cyanobacterial symbionts and the production
75 of bioactive secondary metabolites (Wilson *et al.*, 2014; Freeman *et al.*, 2016; Agarwal *et al.*,
2017). Furthermore, sponge symbionts can serve as an endogenous food source as exemplified by
biomass transfer from a sponge-associated sulfate-reducing bacterial community to host cells
(Hoffmann *et al.*, 2005), and the ingestion of symbiotic methanotrophs in a deep-sea carnivorous
sponge (Vacelet *et al.*, 1995). While a vast array of additional putative symbiotic functions have
80 been hypothesized from taxonomic or metagenomic data (Webster and Thomas, 2016), unequivocal
evidence for specific sponge symbiont physiologies is comparably rare.

Nitrogen cycling in sponge holobionts has received considerable attention and sponge-
symbiont-driven nitrification, denitrification, and anaerobic ammonium oxidation has been
described (Bayer *et al.*, 2008; Southwell *et al.*, 2008; Hoffmann *et al.*, 2009; Schläppy *et al.*, 2010;
85 Radax *et al.*, 2012a). Many marine sponges harbor symbionts phylogenetically related to nitrifying
microbes (Steger *et al.*, 2008; Hoffman *et al.*, 2009; Off *et al.*, 2010), indicating that ammonia
oxidation via nitrite to nitrate is a widely distributed process in these animals. More specifically,
molecular signatures of proteobacterial as well as thaumarchaeal ammonia oxidizers are frequently
detected in sponges, although these microbes rarely co-occur in the same host (Bayer *et al.*, 2008;

90 Radax *et al.*, 2012b). Most sponges that contain symbionts related to ammonia-oxidizers also host
bacteria affiliated with known nitrite-oxidizers, particularly members of the genus *Nitrospira*
(Hoffmann *et al.*, 2009; Off *et al.*, 2010; Reveillaud *et al.*, 2014; Moitinho-Silva *et al.*, 2017a).
While many studies simply equate the molecular detection of microbes related to recognized
nitrifiers with the occurrence of canonical nitrification in sponges (Fan *et al.*, 2012; Mohamed *et al.*,
95 2010), this assumption could be misleading due to the functional versatility of both
Thaumarchaeota and *Nitrospira* (Mussmann *et al.*, 2011; Koch *et al.*, 2014; Daims *et al.*, 2015;
Palatinszky *et al.*, 2015).

Thaumarchaeotes are important ammonia oxidizers in many environments (Pester *et al.*,
2011). Marine sponges host a particularly high diversity of thaumarchaeotes, with many of the 16S
100 rRNA gene sequences falling into phylogenetic clusters that exclusively contain sponge-derived
sequences (Simister *et al.*, 2012). The detection of thaumarchaeotes in sponge larvae also suggests
vertical transmission or early environmental acquisition of these symbionts (Sharp *et al.*, 2007;
Steger *et al.*, 2008; Schmitt *et al.*, 2012). The first genomic information of a member of this phylum
was derived from *Ca. Cenarchaeum symbiosum* in the marine sponge *Axinella mexicana* (Hallam *et*
105 *al.*, 2006), and subsequently thaumarchaeal genomes were recovered from the glass sponge
Lophophysema eversa (Tian *et al.*, 2016) and the temperate sponge *Cymbastella concentrica*, with
the latter also shown to transcribe genes for ammonia oxidation (referred to as CCThau in Moitinho
Silva *et al.*, 2017a and in this manuscript). However, little information is available on the genomic
plasticity and mechanisms of host adaptation in sponge thaumarchaeal symbionts. Furthermore,
110 with the exception of a thaumarchaeal symbiont that inhabits the cold-water sponge *Geodia barretti*
(Radax *et al.*, 2012a, b), direct evidence for the catalysis of ammonia oxidation by sponge-
associated thaumarchaeotes is lacking. While transcription and translation of key functional genes
like the *amoA* gene encoding a subunit of the ammonia monooxygenase of thaumarchaeotes has
been detected across multiple sponge species (Liu *et al.*, 2012; Fiore *et al.*, 2015; Moitinho-Silva *et*

115 *al.*, 2017a), experimental validation of their involvement in nitrification is required to confirm this activity *in situ* (Musmann *et al.*, 2011).

To better understand the physiological capability of Thaumarchaeota in sponges, we used a metaproteogenomic approach to characterize the Thaumarchaeota symbiont of the marine sponge *Ianthella basta*. *I. basta* is an abundant and ecologically important reef sponge found throughout the Indo-Pacific (Berquist and Kelly, 1995). In contrast to many microbially diverse sponge species, 16S rRNA gene surveys have shown that *I. basta* harbors only three dominant microbial phylotypes which cluster within the α - and γ -proteobacteria as well as the *Thaumarchaeota* (Luter *et al.*, 2010). This community structure is stable among different host color morphotypes (Freckelton *et al.*, 2012), between individuals sampled from different geographic regions (Luter *et al.*, 2010), across different host health states (Luter *et al.*, 2010) and in sponges exposed to different environmental stressors (Luter *et al.*, 2012). Here we (i) quantify the abundance of the thaumarchaeal symbionts in *I. basta*, (ii) confirm their role as ammonia oxidizers (iii) document the expression of almost 100 thaumarchaeal genes *in situ*, and (iv) reveal sponge-specific adaptations of these archaea including a putative mixotrophic lifestyle. Furthermore, we demonstrate that nitrite oxidation surprisingly does not occur in *I. basta*.

Results and Discussion

Quantification of the thaumarchaeal *I. basta* symbiont using FISH and qPCR. FISH was performed on fixed cryosections from one sponge individual using the general archaeal probe Arch915 that is fully complementary to the 16S rRNA of the single archaeal phylotype known to inhabit *I. basta* (Fig. 1). Quantitative FISH across 10 images revealed that the thaumarchaeal symbiont comprised $24 \pm 1.6\%$ (standard error, SE) of the total bacterial and archaeal cells detected with probes Arch915 and the probe set EUB338-I-III targeting most bacteria. Absolute quantification of the thaumarchaeal symbiont in five sponge individuals (including those used for

140 metagenome sequencing) using specifically designed qPCR primers, revealed an average absolute
abundance of 2.41 ± 0.7 (SE) $\times 10^{10}$ 16S rRNA gene copies per g wet weight of sponge tissue.
Consistent with the relative abundances derived from FISH, the ratio of thaumarchaeal 16S rRNA
sequences to the total 16S rRNA genes derived from qPCR assays targeting the additional α - and γ -
proteobacterial symbionts of *I. basta* (data not shown) was 22 ± 2.3 (SE)%. Collectively, these data
145 demonstrate that the *I. basta* thaumarchaeote is a dominant member of the *I. basta* microbiome,
occurring at densities that exceed the total microbial biomass of many tropical sponge species
(Taylor *et al.*, 2007).

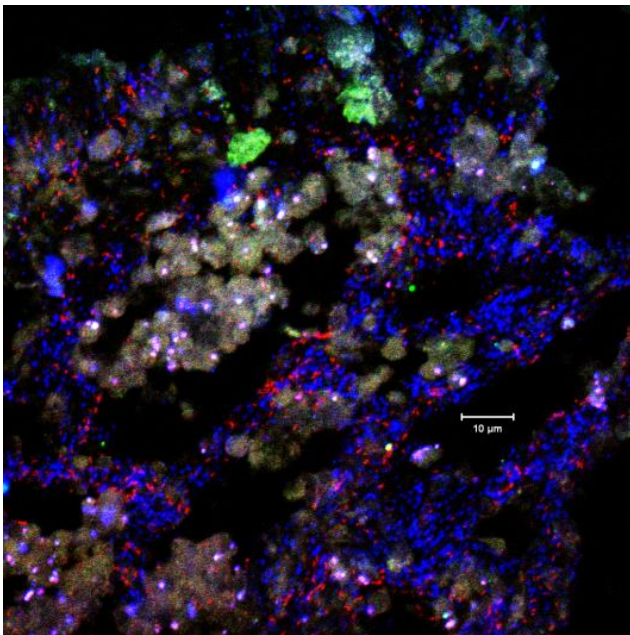


Figure 1. Fluorescence *in situ* hybridization of a 5 μ m cryosection of *Ianthella basta* using double-labeled (Stoecker *et al.*, 2010) probe Arch915 in red and the double-labeled probe EUB338-I-III set in blue. Green and white/lila structures represent autofluorescence. As *I. basta* harbors “*Ca. Nitrosospongia bastadiensis*” as the only archaeon, all red signals represent this AOA.

Metaproteogenomic analyses of the thaumarchaeal symbiont of *I. basta*. An extensive
metagenomic data set consisting of 25.8 Gbp of sequence information was obtained for an *I. basta*
150 individual by Illumina sequencing. A 1.99 Mbp Metagenome-Assembled Genome (MAG)
consisting of 113 contigs was recovered, representing a nearly complete thaumarchaeal genome
(99%) with very low contamination (Supporting Information Table S1). An additional unpublished

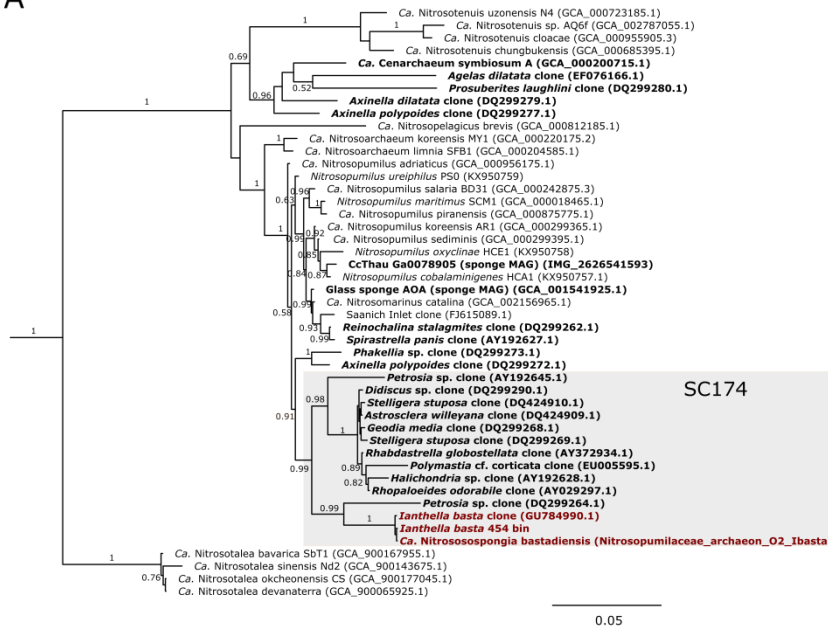
metagenomic data set (250 Mbp generated from pyrosequencing) derived from a different *I. basta* individual was also screened to confirm the presence of genes of interest (Supporting Information Table S2) and support the phylogenetic inferences displayed in Fig. 2. While the shallow metagenome was excluded from detailed analyses, it confirmed that closely related thaumarchaeal symbionts inhabited both sponge individuals. The average nucleotide identity (ANI) between the two thaumarchaeote MAGs was 98.2%, with 82% coverage of the smaller (Illumina data set) MAG, confirming that members of the same thaumarchaeal species (Konstantinidis *et al.*, 2017) reside in both sponge individuals. Both thaumarchaeal symbiont MAGs from *I. basta* possess the highest GC content (64.8%) of any genome-sequenced thaumarchaeote (Fig. 2C) (see Supporting Information for a more detailed discussion of the high GC content).

Metaproteomic analysis performed on the same *I. basta* individual used for Illumina metagenomic sequencing detected a total of 513 proteins, 96 of which were specifically assigned to the *I. basta* thaumarchaeote (representing 5.4% of the genes in the MAG). When combining the normalized spectral abundance factor (NSAF; Florens *et al.*, 2006) values for all samples and analyses, the 96 thaumarchaeal proteins comprised 10.5% of the 513 identified proteins (Supporting Information Table S3). Of the thaumarchaeal proteins, 60.2% (67.4% NSAF) were encoded by gene families shared by all thaumarchaeotes, and 7.1% (3.1% NSAF) were encoded by genes unique to the *I. basta* thaumarchaeote. Of the 96 expressed proteins, 88 were also encoded at a predicted average amino acid identity of $99.4\% \pm 1.4$ (SD) in the second MAG recovered by pyrosequencing. The other 8 protein homologues were also found in the second MAG but at much lower predicted amino acid identities (from 31.4 to 89%).

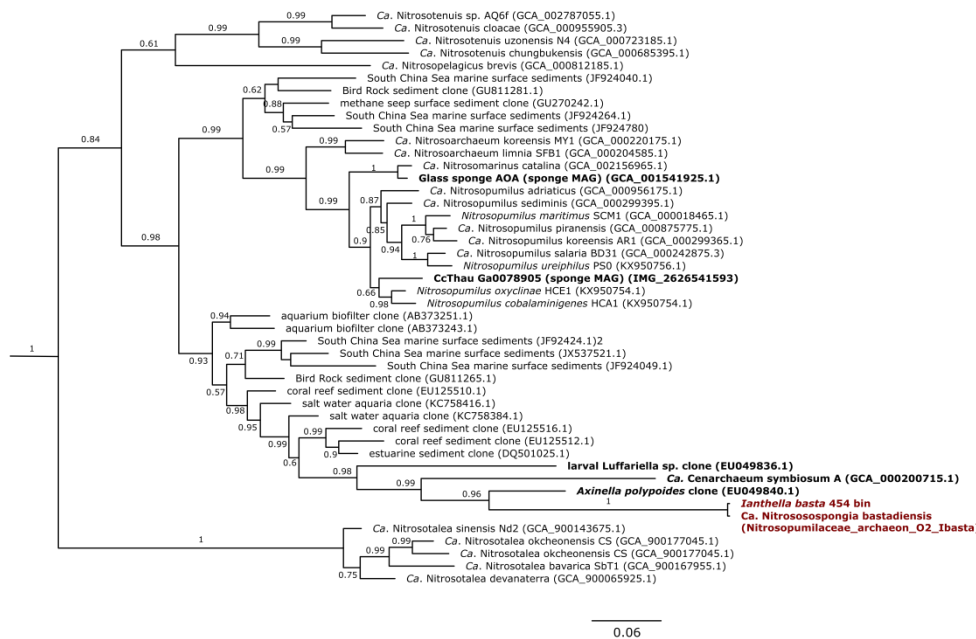
***I. basta* contains a symbiont from a new thaumarchaeal genus.** Of all genome-sequenced thaumarchaeotes, the *I. basta* symbiont has the highest average genomic amino acid identity (AAI) of 58.1% with *Ca. Nitrosopumilus piranensis*. As this represents a new species within a new genus (60-80% AAI is typical for organisms grouped at the genus level; Luo *et al.*, 2014; Konstantinidis

et al., 2017), we propose the name *Ca. Nitrosospongia bastadiensis*. 16S rRNA gene, *amoA*, and
180 concatenated marker gene phylogenies suggest that *Ca. N. bastadiensis* is a member of the family
Ca. Nitrosopumilaceae (Qin *et al.*, 2016) (Fig. 2). Based on the 16S rRNA gene sequence, *Ca. N.*
bastadiensis is a member of the sponge-specific sequence cluster 174 (Simister *et al.*, 2012) which
does not include the three other sponge thaumarchaeal symbionts (*C. symbiosum*, CCThau, and a
putative deep-sea glass sponge thaumarchaeal symbiont) for which genome sequences are available
185 (Hallam *et al.*, 2006; Tian *et al.*, 2016; Moitinho-Silva *et al.*, 2017a). The lack of a close
relationship between *Ca. N. bastadiensis* and the other sponge-associated symbionts is consistent
with the topology of the concatenated single copy conserved marker gene tree (Fig. 2C), while
unexpectedly *amoA* phylogeny supported a clustering of *Ca. N. bastadiensis* with *C. symbiosum*
(Fig. 2B). While the *amoA* genes of *C. symbiosum* and *Ca. N. bastadiensis* showed compositional
190 bias, it was not possible to determine whether this caused their monophyletic grouping in the *amoA*
gene tree. Upon individual addition of *C. symbiosum* and *Ca. N. bastadiensis amoA* gene sequences
to an *amoA* data set, their phylogenetic position in the tree varied dependent on taxa selection (data
not shown). Querying the *Ca. N. bastadiensis* 16S rRNA gene against the Sponge Microbiome
Project (SMP) database (Moitinho-Silva *et al.*, 2017b) and against most other publicly available
195 16S rRNA gene amplicon data sets demonstrated that the habitat of *Ca. N. bastadiensis* is restricted
to a few sponge species (see Supporting Information for more details).

A



B



C

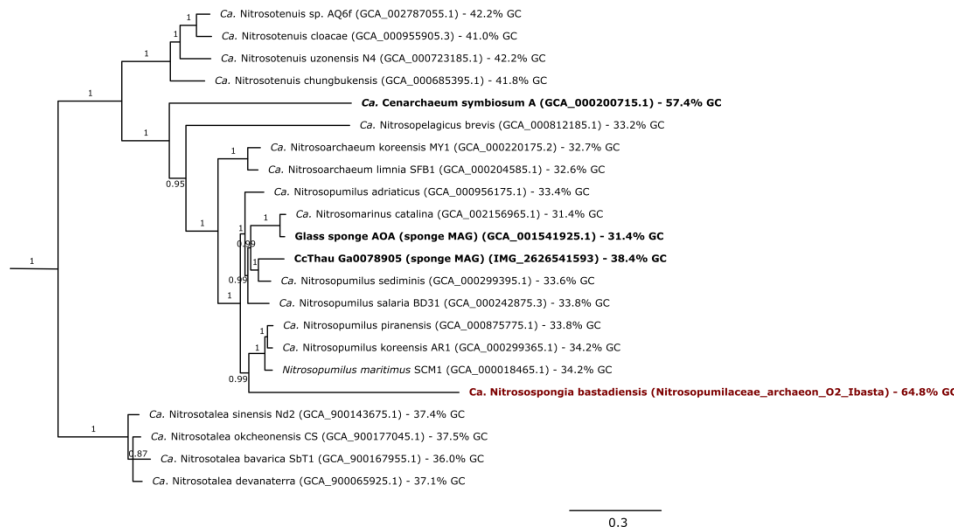


Figure 2. Phylogeny of *Ca. Nitrosospongia bastadiensis*. (A) Bayesian 16S rRNA gene tree. (B) Bayesian *amoA* gene tree. (C) Bayesian phylogenomic tree. For this analysis, *Ca. N. bastadiensis* and *Ca. C. symbiosum amoA* sequences were placed into a reference tree, representing all OTU representatives of the curated database provided in Alves *et al.* (2018), using the Evolutionary Placement Algorithm (EPA; Berger *et al.*, 2011) implemented in RAxML-HPC 8.2.11 (Stamatakis, 2014). Representative sequences clustering with the sponge symbiont *amoA* sequences were then analyzed along with *amoA* sequences from the other genome-sequenced AOA. Note that *Ca. Nitrosotalea okcheonensis* possesses two *amoA* gene copies (Herbold *et al.*, 2017). (C) Bayesian phylogenomic tree based on 34 concatenated universal, single-copy marker genes identified with CheckM (Parks *et al.*, 2015). Bayesian posterior support values >0.5 are indicated for each branch. Outgroups for all trees consisted of all three genome-sequenced members of the *Nitrososphaera* cluster, both members of the *Nitrosocosmicus* clade, and *Ca. Nitrosocaldus icelandicus*. In all trees, sequences obtained from sponges are depicted in bold.

Core metabolism of *Ca. N. bastadiensis*. No major difference in core metabolism was detected between *Ca. N. bastadiensis* and other AOA. Consistent with marine and non-marine thaumarchaeota (Hallam *et al.*, 2006; Tourna *et al.*, 2011; Spang *et al.*, 2012; Park *et al.*, 2014; Bayer *et al.*, 2016), *Ca. N. bastadiensis* possesses a complete urease gene cluster in addition to a gene encoding a urea active transporter (*DUR3*; 86% homologous to *Ca. Nitrosopumilus piranensis*) and is thus capable of taking up (or using internally produced) urea and converting it to ammonia and CO₂. The sponge symbionts *C. symbiosum* and CCThau also encode urea transporters and urease gene clusters. Furthermore, urease subunit gene transcripts from CCThau were found in the metatranscriptomes from *C. concentrica* (Moitinho-Silva *et al.*, 2017a). Thus, it seems likely that urea is not only used by marine AOA in Arctic waters (Alonso-Sáez *et al.*, 2012), but is also a common substrate for group I.1a *Thaumarchaeota* residing in marine sponges and likely represents an important adaptation to life within their hosts that might be excreting urea (Morley *et al.*, 2016). In contrast to *Nitrososphaera gargensis* (Palatinszky *et al.*, 2015), *Ca. N. bastadiensis* does not contain a cyanase for ammonia generation from cyanate, but like all other AOA, it encodes a protein with modest homology to a creatinine-amidohydrolase indicating that it could utilize sponge-derived creatinine and convert it to creatine. While all genome-sequenced AOA symbionts lack a canonical creatinase that would form urea from creatine, AOA including *Ca. N. bastadiensis* do possess a Xaa-Pro aminopeptidase that has been hypothesized as a functional analog (Moitinho-

215 Silva *et al.*, 2017a). All genes encoding the putative subunits of the ammonia monooxygenase
(*amo*) enzyme were also found in the *I. basta* thaumarchaeote (*amoA*, *amoB*, and *amoC*), including
the hypothetical gene *amoX*. *AmoB* and *C* were also detected as proteins (Supporting Information
Table S3). The *amo* gene arrangement [*amoA-amoX-amoC-amoB*] is syntenic to most analyzed
members of *Ca. Nitrosopumilaceae* (Lehtovirta-Morley *et al.*, 2011; Park *et al.*, 2014). Consistent
220 with other AOA, no canonical hydroxylamine dehydrogenase was found but *Ca. N. bastadiensis*
did encode lineage 1 multicopper oxidases (MCO), which have been identified as candidates for
archaeal hydroxylamine dehydrogenases (Kerou *et al.*, 2016). Like most other AOA (with the
exception of *Ca. C. symbiosum*), *Ca. N. bastadiensis* encodes the putative NO-forming nitrite
reductase (*nirK*; found to be highly expressed as protein; Supporting Information Table S3).
225 Interestingly however, the purple cupredoxin Nmar_1307 from *Nitrosopumilus maritimus* that is
capable of oxidizing NO to NO₂⁻ (Hosseinzadeh *et al.*, 2016) is absent in *Ca. N. bastadiensis* and
other sponge-associated AOA (and also many other AOA), but there are several mononuclear
cupredoxins encoded by *Ca. N. bastadiensis* that could have the same function. NO has been
postulated as a key intermediate of AOA (Kozlowski *et al.*, 2016; Carini *et al.*, 2018), a hypothesis
230 that nicely explains the inhibitory effect of the NO scavenger PTIO used in our incubation
experiments described below. Consistent with all other AOA, no canonical NO- or N₂O-reductases
were found, although - like in many other thaumarchaeotes (Liu *et al.*, 2012, Zhalnina *et al.*, 2014;
Santoro *et al.*, 2015) - the putative nitric oxide reductase accessory proteins, NorQ and NorD are
encoded in the genome (with the NorD subunit being expressed).

235 Intriguingly, *Ca. N. bastadiensis* encodes many putative surface-layer (S-layer) proteins
(Supporting Information Fig. S1), with six of them being expressed at combined NSAF values of
6.78% (Supporting Information Table S3). The expanded group of S-layer proteins in *Ca. N.*
bastadiensis (and *Ca. C. symbiosum*) may be involved in molecular sieving, adhesion, surface
recognition or other types of interactions within the extracellular matrix of the sponge environment.
240 The negatively charged surface of the Thaumarchaeota S-layer was recently proposed to help

concentrate the charged solute ammonium into the pseudo-periplasmic space (Li *et al.*, 2018). The theoretical isoelectric points of all *Ca. N. bastadiensis* S-layer proteins (3.62 – 4.92) were in a similar range to those calculated for *N. maritimus* and *Nitrosoarchaeum limnia* SFB1 S-layer proteins (3.4 – 4.08), consistent with the proposed mechanism for charged solute acquisition.

245 As expected, *Ca. N. bastadiensis* encodes all key enzymes of the thaumarchaeal 3-hydroxypropionate/4-hydroxybutyrate pathway (Könneke *et al.*, 2014; Otte *et al.*, 2015) and three proteins catalyzing five steps of this thaumarchaeal CO₂ fixation pathway were also detected within the metaproteome. Regarding nitrogen assimilation, *Ca. N. bastadiensis* encodes and expresses an *amt* transporter of the *amt-2* lineage which has been hypothesized to be a high affinity ammonia
250 transporter (Nakagawa and Stahl, 2013; Offre *et al.*, 2014). The *amt-2* gene of *Ca. N. bastadiensis* is nested within the complex V ATP synthase operon (of which four of the encoded subunits were found to be expressed), indicating that energy generated from ammonia oxidation directly fuels the transport of ammonium for assimilation. Interestingly, in contrast to most other AOA, no putative low affinity *amt* transporter gene was found in *Ca. N. bastadiensis*, a feature shared with the sponge
255 symbionts *Ca. C. symbiosum* (Hallam *et al.*, 2006) and CCThau (Fig. 3; Moitinho-Silva *et al.*, 2017a). This pattern indicates that sponge thaumarchaeotes are either not exposed to high concentrations of ammonia or that they can assimilate nitrogen under most conditions by uptake of organic nitrogen sources and only use the high affinity ammonia transporter if organic nitrogen sources are scarce. Regarding nitrogen assimilation, it is also noteworthy, that the almost complete
260 bin of *Ca. N. bastadiensis* contains only a single gene encoding a member of the nitrogen regulatory protein PII superfamily, whereas, with the exception of *Ca. C. symbiosum*, higher copy numbers of these genes are generally found in other AOA (Fig. 3; Kerou *et al.*, 2016). Furthermore, in contrast to most other AOA, the two key genes coding for subunits of the polyhydroxyalkanoate (PHA) synthase *phaC* and *phaE* are apparently lacking in *Ca. N. bastadiensis*. We also found no
265 indications for use of alternative carbon storage compounds such as starch or glycogen in *Ca. N. bastadiensis*.

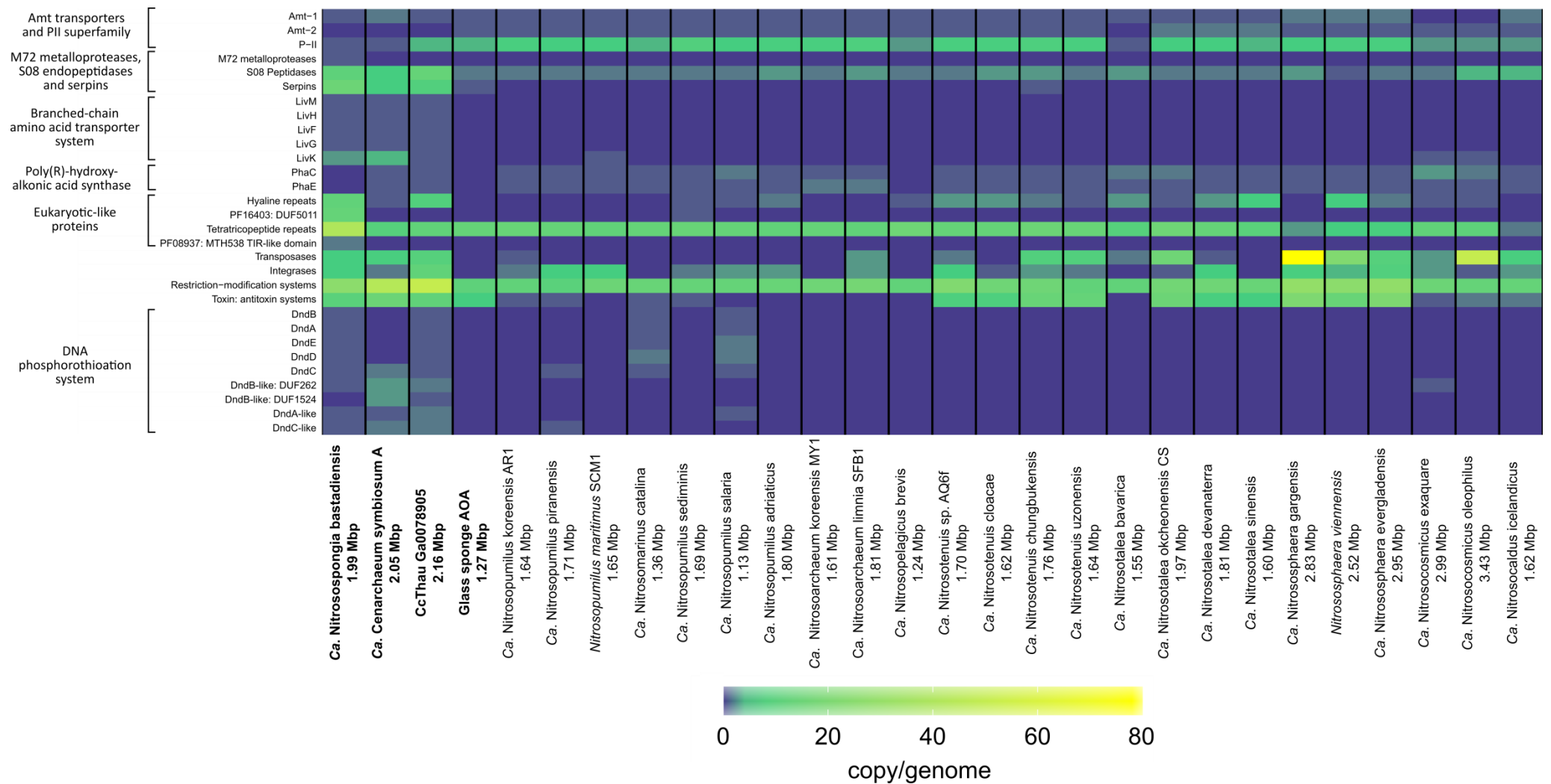


Figure 3. Heat map showing the distribution and gene copy number per genome of selected genes and gene classes among genome-sequenced AOA. The color scale indicates copies per genome and MAG, respectively. Sponge-derived MAGs start on the left and are depicted in bold, followed by members of *Ca. Nitrosopumilaceae*, *Ca. Nitrosotenuaceae*, *Ca. Nitrosotaleae*, the *Nitrososphaerales*, and *Ca. Nitrosocaldales*. Genome sizes are listed next to each species name. An extended version of this Figure is available as Supplemental Information Fig. 5.

In contrast to many free-living AOA (Blainey *et al.*, 2011; Jung *et al.*, 2014; Mosier *et al.*, 2012; Spang *et al.*, 2012; Bayer *et al.*, 2016) and the glass-sponge AOA, *Ca. N. bastadiensis*, CCThau, and *Ca. C. symbiosum* do not encode an archaellum. Lack of chemotaxis and motility
270 traits has been described for other microbial symbionts (Karimi *et al.*, 2018) and likely represents an adaptation to the sponge-associated lifestyle. The absence of an archaellum in *Ca. N. bastadiensis* would also be consistent with vertical transmission of this symbiont, although this remains to be demonstrated.

275 **Shared gene families of sponge AOA.** During genome annotation of *Ca. N. bastadiensis*, we particularly focused on gene families unique to this organism or to *Ca. N. bastadiensis* and other thaumarchaeal sponge symbionts (among all genome-sequenced thaumarchaeotes) as these gene families likely represent adaptations to a sponge-associated lifestyle. While 40% (represented by 681 gene families) of the *Ca. N. bastadiensis* genes were members of core gene families represented
280 in every sequenced thaumarchaeote, 35% (represented by 616 gene families) had no close orthologues among other thaumarchaeotes. Additional pairwise comparisons between *Ca. N. bastadiensis* and all other thaumarchaeote genomes were performed in order to identify shared gene families of species pairs that are absent in all other thaumarchaeotes. Interestingly, the highest number occurred with CCThau (n = 18) and *Ca. C. symbiosum* (n=11) (Supporting Information
285 Fig. S2). Furthermore, 14 gene families were found to be exclusively shared among all three thaumarchaeal sponge symbionts, while only one gene family was found to be shared with the putative deep-sea glass sponge thaumarchaeal symbiont and *C. symbiosum*. In total, 44 gene families were exclusively shared with at least one other thaumarchaeal sponge symbiont representing 2.4% of the *Ca. N. bastadiensis* genes.

290

Ca. N. bastadiensis and other sponge AOA secrete proteases and their inhibitors. Among the genes that were unique to *Ca. N. bastadiensis* or exclusively shared with other sponge AOA were several genes involved in degradation of extracellular protein and inactivation of extracellular (host-derived) proteases. *Ca. N. bastadiensis* possesses a putatively exported metalloprotease of the M72 family (Drapeau, 1980; Passmore *et al.*, 2015) that was also detected in the metaproteome. This large protease (2027 AA) contains hyaline repeat domains suggestive of an additional adhesive property. The M72 family of metalloendopeptidases have so far not been found in any *Archaea* (Trame *et al.*, 2014) and most characterized members of this enzyme family are peptidyl-Asp metalloendopeptidases that hydrolyze bonds on the NH₂-terminal side of aspartic acid and cysteic acid residues (Drapeau, 1980). Furthermore, the *I. basta* thaumarchaeal symbiont contained 11 genes affiliated with four gene families that encode subtilisin-like serine protease domains (S08A family endopeptidases) (Supporting Information Fig. 3A). Although serine endopeptidases of the S08A family are also found in other thaumarchaeotes, with the exception of *C. symbiosum* (4 copies), and CCThau (13 copies), most contain only two copies. Furthermore, it is noteworthy that two of the four S08A family endopeptidase gene families were either exclusively found in *Ca. N. bastadiensis* or shared with CCThau (Fig. 4, Supporting Information Fig. 3A). Interestingly, several of the *Ca. N. bastadiensis* S08A endopeptidases are predicted to be exported without a membrane anchor, while most other thaumarchaeal S08A serine endopeptidases seem to be membrane anchored.

Ca. N. bastadiensis also encodes 15 genes for the I4 family of serpins which represent serine protease inhibitors. Some of these cluster with serpins from *C. symbiosum* (Hallam *et al.*, 2006) and CCThau, although amongst the non-host associated AOA, only *Nitrosotenuis chungkubensis* encodes a serpin (Supporting Information Fig. 3B). Nine of the *Ca. N. bastadiensis* serpins are predicted to be extracellular and are frequently found adjacent to S08A family endopeptidases in the genome (Fig. 4B). Serpins belong to a large family of irreversible inhibitory substrates of proteases, often but not exclusively of the serine class, that has been widely characterized in

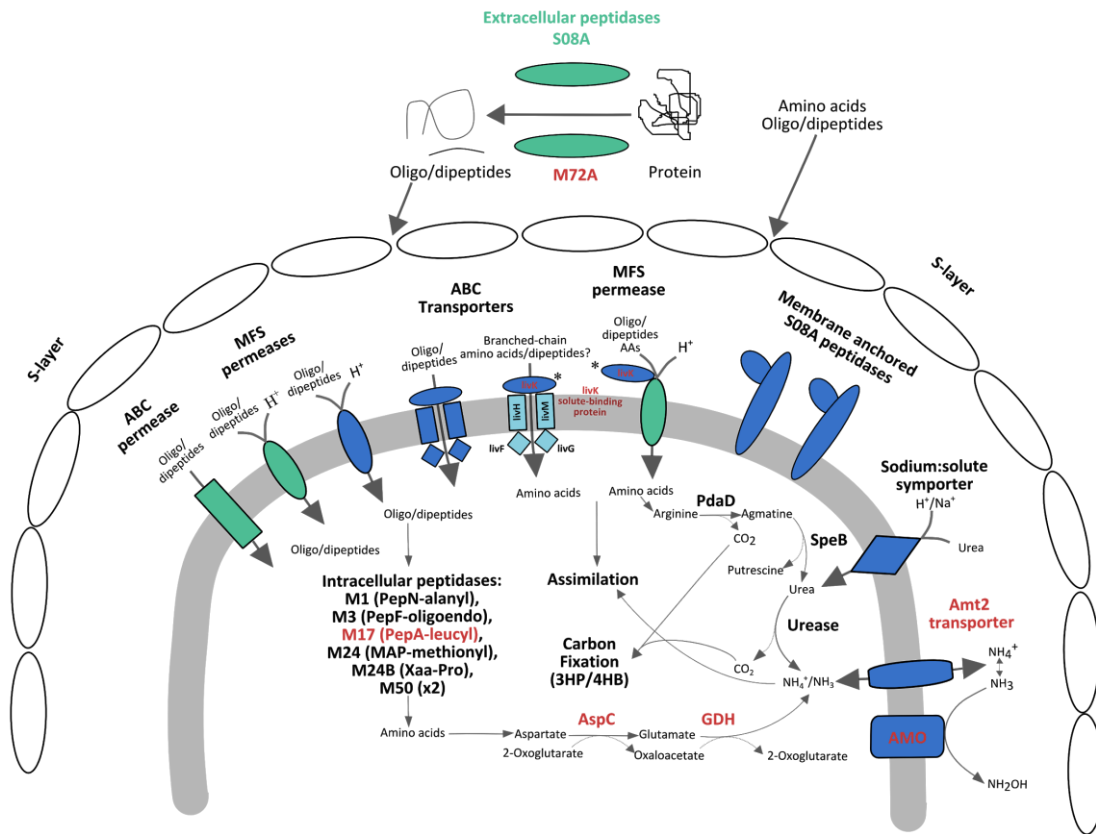
mammals, insects, plants as well as some viruses, and which can mediate host-microbe interactions (Ventura *et al.*, 2012). These findings suggest that AOA living as sponge symbionts use serpins to regulate their own secreted endopeptidases and/or sponge serine proteases found to be highly expressed in other sponges (Riesgo *et al.*, 2014), including within the host of *C. symbiosum* (Zaikova, E., Ph.D. thesis, 2007).

Peptide and amino acid uptake by *Ca. N. bastadiensis* and other sponge AOA. *Ca. N.*

bastadiensis, like many other AOA, is well equipped for uptake of oligo- and dipeptides as well as amino acids by encoding a set of transporters widely distributed in this clade. After uptake, *Ca. N. bastadiensis*, like other AOA, has the genomic repertoire to degrade these peptides and amino acids and release ammonia (Fig. 4). In addition to features common to many AOA, *Ca. N. bastadiensis* also contains the *livFGHMK* operon which in proteobacteria and cyanobacteria encodes a high-affinity branched-chain amino acid transport system (which can also transport other amino acids) belonging to the ATP binding cassette (ABC) superfamily of transporters (Hoshino, 1979; Adams *et al.*, 1990; Hosie *et al.*, 2002; Picossi *et al.*, 2005) (Fig. 4). Among all other thaumarchaeote genomes, only *C. symbiosum*, CCThau, and a thaumarchaeote MAG from the Caspian Sea (Mehrshad *et al.*, 2016) encode this transporter (Supporting Information Fig. S4). The periplasmic subunit LivK from *Ca. N. bastadiensis* was expressed as protein (Supporting Information Table S3) and this subunit was also reported to be expressed in CCThau (Moitinho-Silva *et al.*, 2017a). Adjacent to the expressed *livK* in *Ca. N. bastadiensis* is a gene encoding three consecutive LivK domains (along with the requisite ligand binding sites) which is congruent with *C. symbiosum* containing 4 copies of the LivK solute binding component (Supporting Information Fig. S4A). In *C. symbiosum*, the transporter encoding the *livFGHMK* operon is located next to several extracellular trypsin-like serine protease encoding genes, indicating that the transporter is involved in amino acid uptake in this organism. However, given the substrate promiscuity of this transporter family (Adams *et al.*, 1990; Valladares *et al.*, 2002; Beckers *et al.*, 2004; Picossi *et al.*, 2005) (Supporting

Information Fig. S4), experimental validation will be required before more specific predictions on its function in sponge AOA can be made. Such investigations would be particularly important, as
345 multiple *LivFGHMK* operons are also present in two γ -proteobacterial sponge symbionts (Gauthier *et al.*, 2016), suggesting that amino-acid transport and utilization mediated by this transporter-type could be an important feature of sponge symbionts and may even contribute to sponge-mediated dissolved organic matter transfer to higher trophic levels (de Goeij *et al.*, 2013). Proteases are also important for sponge metabolism and a body of research has focused on sponge proteases and their
350 inhibitors from a biodiscovery context (Arreguín *et al.*, 1993; Wilkesman and Schroeder, 2002; Wilkesman and Schroeder, 2007; Tabares *et al.*, 2011). In these analyses generally, sponge holobiont samples are used for enzyme or inhibitor purification and characterization. Thus, our observation that abundant archaeal sponge symbionts secrete proteases and serpins, demonstrate that in such assays, it will remain unclear whether it is the microbial symbionts, or the host animals
355 that are producing the analyzed biomolecules.

A



B

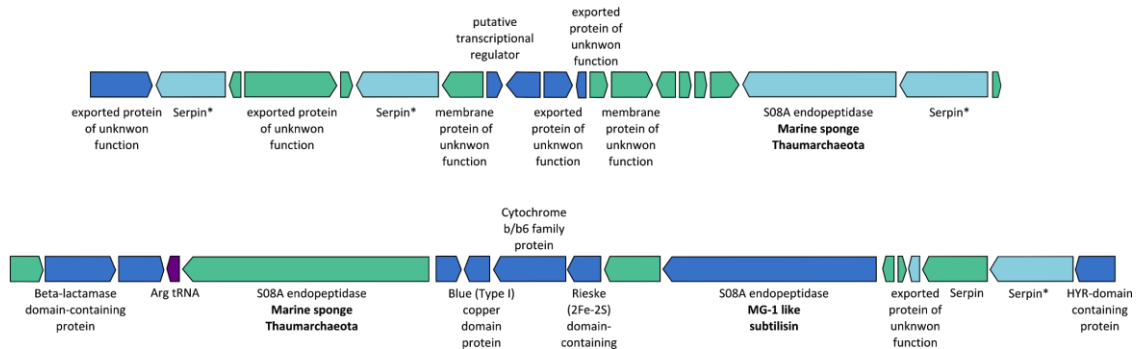


Figure 4. (A) Reconstructed metabolic pathways of genes and their expression (red) detected in *Ca. N. bastadiensis* proposed to be involved in extracellular (and intracellular) protein degradation as well as amino acid transport and assimilation. Predicted proteins (and their respective subunits when relevant) are color coded to denote the degree of homology among all sequenced *Thaumarchaeota*: green - *Ca. N. bastadiensis* unique gene families; light blue - shared exclusively among thaumarchaeal sponge symbionts; dark blue - ubiquitously found in *Thaumarchaeota*. Extracellular proteins derived from the marine environment as well as the sponge mesohyl may be degraded extracellularly or in the thaumarchaeal pseudo-periplasmic space. These resultant oligo/dipeptides and amino acids which can also be derived from the environment can then be transported by a suite of ABC transporters and MFS permeases into the cytoplasm to be further degraded by intracellular peptidases or assimilated. Amino acids such as arginine and aspartate can be further degraded to form $\text{NH}_3/\text{NH}_4^+$ for assimilation or export to the pseudo-periplasmic space

for ammonia oxidation. Arginine can be decarboxylated by arginine decarboxylase (PdaD) to agmatine which can then be degraded to urea by agmatinase (SpeB). Both proteins are ubiquitously distributed among the *Thaumarchaeota*. All but one of the branched-chain amino acid transporter subunits (LivFGHMK operon) are exclusively found among thaumarchaeal sponge symbionts, while the periplasmic solute binding subunit (LivK – found to be expressed and denoted by an asterisk) can be found not only among sponge symbionts but also in *N. maritimus* and members of the genus *Nitrosocosmicus* (see also Fig. 3). (B) Genetic context map depicting examples of co-localized S08A endopeptidases and serine protease inhibitors (serpins). Color coding as in panel (A), except purple denotes the presence of a tRNA, sites at which gene insertions are common. The asterisk next to serpins color coded as sponge-specific (light-blue) denotes that *Ca. N. chungbukensis* is the sole non-sponge symbiont encoding a serpin and belonging to this orthologous group (see Supplemental Information Fig. S3B). AspC, aspartate aminotransferase; GDH, glutamate dehydrogenase.

A mixotrophic lifestyle for *Thaumarchaeota* has been inferred since the earliest reports on planktonic thaumarchaeotes described incorporation of organic carbon into signature lipids (Ingalls *et al.*, 2006) and incorporation of labelled amino acids leucine (Ouverney and Fuhrman, 2000; Herndl *et al.*, 2005) and aspartate (Teira *et al.*, 2006) into single planktonic cells. While initial findings - that pure cultures of *Thaumarchaeota* in group I.1a (Qin *et al.*, 2014) and group I.1b (Tourna *et al.*, 2011) could have increased growth rates in the presence of α -keto acids and ammonia - were subsequently proven to be a result of the H₂O₂ scavenging ability of α -keto acids (Kim *et al.*, 2016; Qin *et al.*, 2017), abundant genes and transcripts of S08A family serine proteases and other metallopeptidases in MG-1 deep-sea thaumarchaea have also been attributed to mixotrophic activity (Li *et al.*, 2015). A deep sea *Thaumarchaeota* has also been found to encode the *livFGHMK* operon (Mershad *et al.*, 2016), which clusters with the two sponge thaumarchaeotes (Supporting Information Fig. S4), highlighting both environments as potential ecological reservoirs of mixotrophic thaumarchaea. Furthermore, growth of thaumarchaeotes in an industrial wastewater treatment plant that was uncoupled to ammonia-oxidation has been described (Mussmann *et al.*, 2011). Interestingly, many members of the Aigarchaeota, a sister group to the Thaumarchaeota (Guy and Ettema, 2011), also encode the full *livFGHMK* operon, and a single cell aigarchaeal-like genome from cold marine sediments encodes additional extracellular proteases, di-/tripeptide

transporters and aminotransferases (Lloyd *et al.*, 2013). It is therefore tempting to speculate that the
375 thaumarchaeal ancestor was a mixotroph or even strict heterotroph (as also indicated by genomic
and experimental data that deep branching clade I1c and d members in the Thaumarchaeota lack
genes required for ammonia oxidation; Beam *et al.*, 2014; Lin *et al.*, 2015; Weber *et al.*, 2015), and
that sponge Thaumarchaeota along with a few other members of this clade retained the capability to
use amino acids due to specialized environmental conditions. Future experiments would need to
380 demonstrate uptake of amino acids by *Ca. N. bastadiensis* to reveal whether it fuels assimilation or
heterotrophic growth and/or is used for sequential intracellular generation and oxidation of
ammonia (de Boer and Laanbroek, 1989; Burton and Prosser, 2001).

Eukaryotic-like proteins (ELPs) in *Ca. N. bastadiensis*. *Ca. N. bastadiensis* encodes a number of
385 genes containing domains postulated to have an evolutionary origin within the eukaryotes and
which are thought to be important for modulating interactions between bacteria and eukaryotic
hosts (Callebaut *et al.*, 2000; Lurie-Weinberger *et al.*, 2010; Patterson *et al.*, 2014). Recent
metagenomic analyses revealed an abundance of genes encoding such ELPs in the bacterial
symbionts of sponges (Fan *et al.*, 2012; Reynolds and Thomas, 2016; Díez-Vives *et al.*, 2017), with
390 many of these being expressed (Díez-Vives *et al.*, 2017). Four types of ELPs are found in *Ca. N.*
bastadiensis: Proteins with tetratricopeptide repeats (TPR), the Toll-interleukin-1 receptor (TIR) -
like domain PF08937 (DUF1863; Cort *et al.*, 2000), immunoglobulin-like (Ig-like) domains
(DUF5011; Shigeno-Nakazawa *et al.*, 2016), and hyaline repeats (HYR; Callebaut *et al.*, 2000) and
an extensive discussion of these ELPs is provided in the Supplementary Information. While the *Ca.*
395 *N. bastadiensis* genome is enriched in proteins containing TPR, TIR, Ig-like, and HYR domains, it
lacks the ankyrin repeat proteins, leucine-rich repeat proteins, protein tyrosine kinases, and
armadillo-repeat proteins frequently detected in other symbiotic and pathogenic microbes (Fan *et*
al., 2012; Jernigan and Bordenstein, 2015).

400 **Mobile and selfish genetic elements in *Ca. N. bastadiensis*.** In contrast to other members of *Ca.*
Nitrosopumilaceae, *Ca. N. bastadiensis* along with the other two sponge AOA, *C. symbiosum* and
CcThau (but not the deep-sea glass sponge AOA) are enriched in transposases, restriction-
modification (RM) systems (including a Type II restriction endonuclease, PF13156, which is
exclusively found in sponge AOA - Supporting Information Fig. S5), toxin-antitoxin (T-A) systems,
405 as well as genes putatively involved in DNA phosphorothioation (Fig. 3, Supporting Information
Fig. S5). However, no differential enrichment of integrases was detected within sponge-associated
thaumarchaeotes. The abundance of transposases and other mobile/selfish genetic elements
(MGEs/SGEs) in sponge AOA is consistent with what has been reported for other sponge-
associated microbes (Fan *et al.*, 2012, Horn *et al.*, 2016), suggesting that evolution of AOA sponge
410 symbionts is also affected by horizontal gene transfer in the concentrated milieu of environmental
bacteria and viruses resulting from sponge feeding and pumping activity.

Among the above-mentioned genetic elements, the complete genetic repertoire for DNA
phosphorothioation (PT) (Wang *et al.*, 2007; You *et al.*, 2007) in *Ca. N. bastadiensis* and CcThau
(*dndA, B, C, D, E*) is particularly noteworthy as among cultured AOA, only *Ca. Nitrosomarinus*
415 *catalina* (Ahlgren *et al.*, 2017) and *Ca. Nitrosopumilus salaria* also encode this DNA modification
system, which acts as a primitive immune system by enabling discrimination between self and non-
self DNA. In addition, *Ca. N. bastadiensis*, CcThau and *C. symbiosum* have genes with low
similarity to *dndA, B* and *C*, respectively. The *dndB*-like genes contain domains (DUF262,
DUF1524) previously identified as components of R-M systems (Miller *et al.*, 2005; Machnika *et*
420 *al.*, 2015) which are known to be enriched in sponge microbiomes (Fan *et al.*, 2012, Horn *et al.*,
2016). *Ca. N. bastadiensis* also encodes most enzymes necessary for the production of archaeosine,
a highly modified tRNA nucleoside (Phillips *et al.*, 2012). Interestingly, amongst all analyzed
Thaumarchaeota, the critical enzyme (aTGT) of this pathway is missing the RNA binding site
(PUA domain) while maintaining the conserved substrate binding pocket found in *Crenarchaeota*
425 (Phillips *et al.*, 2012). Although the PUA domain is dispensable for archaeosine formation (Sabina

and Söll, 2006) it is as of yet unknown whether *Ca. N. bastadiensis* can use this for DNA modification, as recently described for another system variant in a *Salmonella* species (Thiaville *et al.*, 2016).

430 **Ammonia-oxidation is exclusively mediated by *Ca. N. bastadiensis* in *I. basta* and is coupled to carbon fixation.** The presence of *amoA*-encoding and expressing thaumarchaea does not prove that these microbes actually perform ammonia-oxidation in a system (Mussmann *et al.*, 2011). Consequently, nitrification rates of *I. basta* harboring the thaumarchaeal symbiont were experimentally determined to verify ammonia-oxidizing activity and to obtain insights into the mean nitrifying activity per symbiont cell. Interpretation of these data was facilitated by the fact 435 that *I. basta* according to previous 16S rRNA gene based surveys (Webster *et al.*, 2010; Luter *et al.*, 2010; Freckelton *et al.*, 2012; Luter *et al.*, 2012) and our metagenomic and metaproteomic data contains a single AOA symbiont species and does not harbor bacterial ammonia-oxidizers or comammox organisms (Daims *et al.*, 2015). Incubation experiments with freshly collected sponge 440 clones were performed in the presence of different ammonium concentrations ranging from ambient seawater ($0.29 \pm 0.1 \mu\text{M}$) via $25 \mu\text{M}$ to $100 \mu\text{M}$. Gross and net nitrification rates as well as net fluxes of ammonium, nitrite, and nitrate were determined in 24 h laboratory incubation experiments that were repeated several times over 7 days (Supporting Information Fig. S6). Across all treatments, gross and net nitrification rates were similar in magnitude and highly correlated (gross 445 rates = $1.13 \times \text{net rates} + 0.6$; $R^2 = 0.94$; $p < 0.01$), suggesting that nitrate removal does not occur at significant rates, with the exception of sponges in the $100 \mu\text{M NH}_4^+$ treatment (Fig. 5A). Gross nitrification rates of sponges incubated with 25 or $100 \mu\text{M NH}_4^+$ were significantly greater than in ambient seawater ($p < 0.05$; ANOVA followed by Tukey HSD Test) (Fig. 5A). The stimulation of nitrification by increased ammonium availability, which has also been observed for other sponge 450 species (Corredor *et al.*, 1988; Bayer *et al.*, 2007; Bayer *et al.*, 2008; Schläppy *et al.*, 2010), indicates ammonium limitation of the symbiotic nitrifiers under ambient conditions. Almost no net

ammonium release was observed from the *I. basta* holobiont in unamended seawater, indicating balanced ammonium production and consumption (Fig. 5B). Net nitrification rates of the *I. basta* holobiont under ambient experimental conditions were slightly higher but generally comparable to those reported for other sponge species (Supplemental Information Table S4). These data likely reflect nitrification rates under natural field conditions, as NH_4^+ concentrations in the unamended treatment were consistent with those reported for Orpheus Island, where *I. basta* was collected (Jompa and McCook, 2002).

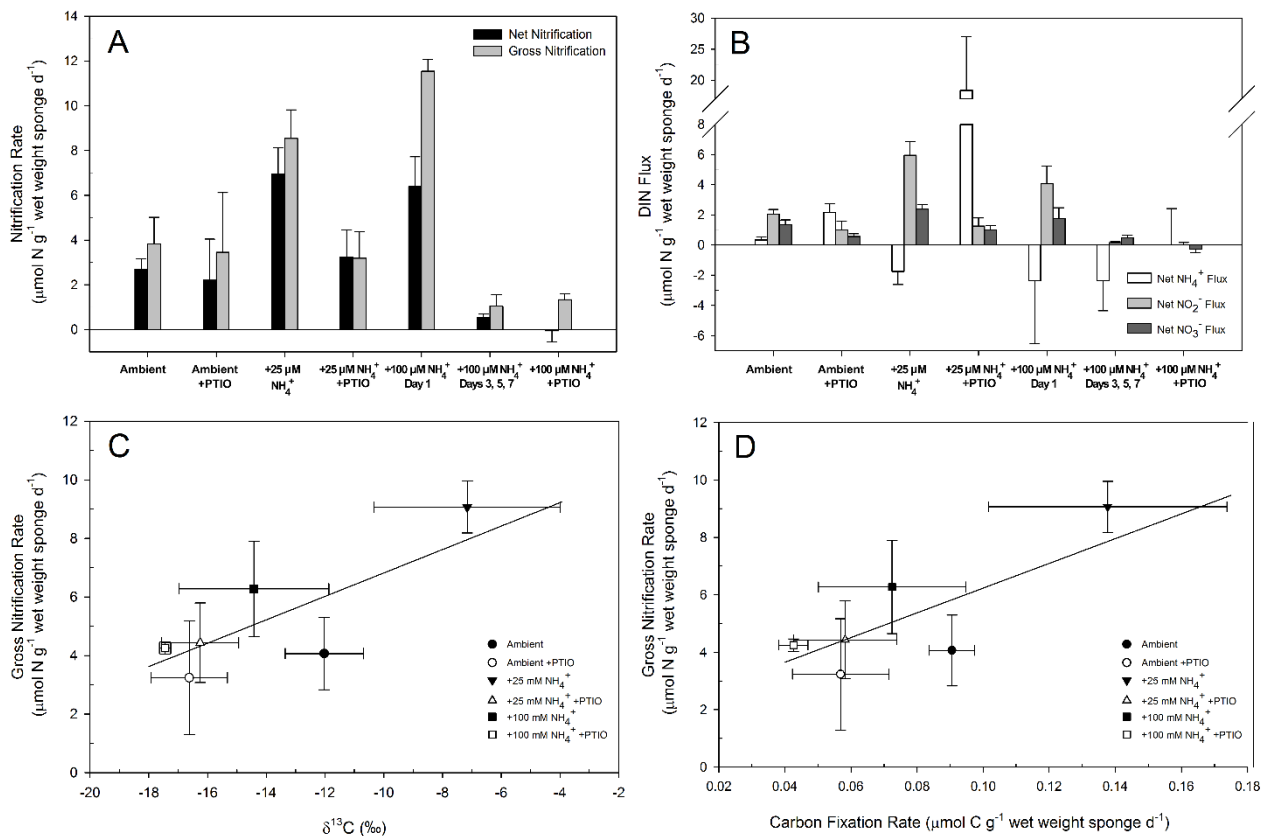


Figure 5. Nitrification activity of the *I. basta* holobiont during 7-day incubation experiments at ambient conditions and with added ammonium (25 or 100 μM) and with/without the AOA inhibitor PTIO. A schematic overview of the corresponding experimental setup is given in Supplementary Information Fig. S6. (A) depicts net nitrification (as calculated from the addition of net NO_2^- and NO_3^- flux) and gross nitrification rates (estimated using the $^{15}\text{NO}_3^-$ isotope pool dilution method). (B) shows net fluxes of all DIN species for the ambient, +25 $\mu\text{M NH}_4^+$ and +100 $\mu\text{M NH}_4^+$ treatments. Rates of the different days in the ambient and +25 $\mu\text{M NH}_4^+$ incubations were averaged as they were not statistically different. (C) shows the relationship between $\delta^{13}\text{C}$ values of sponge

clones (labeled from added ^{13}C -bicarbonate) sampled on day 7 and gross nitrification rates of the *I. basta* holobiont. To yield average nitrification activity over the course of the incubation, we used the weighted mean of the gross nitrification rates (i.e., days are weights). For comparison, the $\delta^{13}\text{C}$ value reflecting the natural abundance was determined from the sponge individual used for Illumina metagenome sequencing and found to be -21.7‰ (not shown). The $\delta^{13}\text{C}$ values derived from the *I. basta* holobiont nitrification experiments displayed a positive and significant correlation with gross nitrification rates ($R = 0.656$, $p < 0.005$). (D) shows the relationship between gross nitrification rates of the *I. basta* holobiont and the carbon fixation rate of sponge clones sampled on day 7. The carbon fixation rates derived from the *I. basta* holobiont nitrification experiments displayed a positive and significant correlation with gross nitrification rates ($R = 0.665$, $p < 0.005$). Error bars in panels C and D reflect the standard error of the sample mean, where $n = 3$ for the carbon fixation rates, for the $\delta^{13}\text{C}$ values, and for all weighted gross nitrification rates in all treatment conditions. PTIO additions started two days after the experiment was commenced, hence fixation of ^{13}C -labeled bicarbonate during the first two days was not influenced by inhibition.

460 Net nitrification rates in sponges subjected to $100\ \mu\text{M}\ \text{NH}_4^+$ for 24 h were significantly lower than gross nitrification rates ($p < 0.05$; two-tailed t-test, Fig. 5A), indicating increased nitrate consumption. Gross and net nitrification rates also became significantly depressed after prolonged exposure (> 3 days) to $100\ \mu\text{M}\ \text{NH}_4^+$, being significantly lower than in the 24 h incubations at $100\ \mu\text{M}\ \text{NH}_4^+$ as well as in the ambient seawater controls ($p < 0.01$, Mann-Whitney U-test) (Fig. 5A).

465 From AOA pure culture studies, there are no indications that $100\ \mu\text{M}\ \text{NH}_4^+$ can be inhibitory for members of this clade, with concentrations of 2 to 20 mM NH_4^+ needed for inhibition (Hatzenpichler *et al.*, 2008; Martens-Habbena *et al.*, 2009; Tourna *et al.*, 2011; Li *et al.*, 2016; Sauder *et al.*, 2017; Sauder *et al.*, 2018). While we cannot exclude that *Ca. N. bastadiensis* is particularly sensitive to ammonium, it is also possible that the inhibition was caused by indirect

470 effects. For example, if the sponge host was adversely affected at $100\ \mu\text{M}\ \text{NH}_4^+$, decay of sponge cells could cause stimulation of heterotrophic microbes and thus oxygen limitation. Typical acute toxicities in freshwater fish and invertebrates (96-hour LC_{50}) are reached at unionized ammonia (NH_3) concentrations between $4.7 - 38\ \mu\text{M}$ (with salinity typically having a slight ameliorating effect and tests on marine invertebrates very rare; Boardman *et al.*, 2004) and susceptibility to

475 ammonia toxicity can be increased by low O_2 concentrations (Camargo and Alonso, 2006). At ambient incubation conditions about 4.5% of total ammonia was present as NH_3 , hence *I. basta* may

have experienced physiological stress and intermittent hypoxic conditions in the 100 μM NH_4^+ treatments.

PTIO, a scavenger of free radical nitric oxide (NO) (Amano and Noda, 1995; Ellis *et al.*, 2001), has been described as a specific inhibitor of AOA (Yan *et al.*, 2012; Martens-Habbena *et al.*, 2015) and was therefore used to demonstrate that *Ca. N. bastadiensis* is responsible for ammonia-oxidation in *I. basta*. As expected, the addition of PTIO to sponge incubations in the ambient and 25 μM NH_4^+ amended treatments, resulted in significantly depressed gross and net nitrification rates and increased ammonium accumulation when compared to the non-PTIO treated sponge incubations ($p < 0.05$; except for gross nitrification rates under ambient conditions, $p = 0.885$; Mann-Whitney U-test; Fig. 5A and Fig 5B). In the case of the 100 μM NH_4^+ amendment, gross and net nitrification was already strongly reduced in the absence of PTIO after 3 days of incubation, so further inhibition by PTIO could not be demonstrated under these conditions (Fig. 5A). Residual ammonia oxidizing activity in some of the experiments in the presence of 75 μM PTIO was unexpected as pure AOA cultures are known to be fully or almost completely inhibited by this concentration of PTIO (Shen *et al.*, 2013; Martens-Habbena *et al.*, 2015). NO production by *I. basta* (many sponges express NO synthases, Riesgo *et al.*, 2014) may have contributed to PTIO inactivation thereby lowering its inhibitory effect on the AOA. To our knowledge, PTIO has not previously been used in sponge microbiome research (but is commonly applied to eukaryotic tissues including sponges; Müller *et al.*, 2006; Ueda *et al.*, 2016) and further optimization of the concentration is recommended for future applications in these animals.

In 29 of the 33 sponge incubations (incubations with added PTIO or inhibited by addition of 100 μM NH_4^+ were excluded) (Fig. 5B), accumulation rates of NO_2^- exceeded those of NO_3^- by an average factor of 2.8 and this difference was found to be significant in the ambient and 25 μM NH_4^+ treatments (both $p < 0.01$, Mann-Whitney U-test). In contrast, no significant difference between NO_2^- and NO_3^- production was observed in the PTIO-amended incubations nor in those experiments

where prolonged exposure to 100 μM NH_4^+ negatively affected net nitrification rates (all: $p > 0.05$, Mann-Whitney U-test).

All nitrification experiments were performed in intermittently closed aquaria (Supporting Information Fig. S6) to prevent loss of labeled CO_2 but to provide sufficient oxygen exchange to maintain sponge health. Additional control experiments were performed at ambient and 100 μM NH_4^+ , to compare DIN fluxes and inferred net nitrification rates using this setup, to incubations performed in constantly open containers. Interestingly, slightly but significantly higher NO_2^- fluxes and net nitrification rates were observed with ambient seawater in the intermittently closed aquaria (both $p < 0.01$, two-tailed t-test; Supporting Information Fig. S7), indicating that ammonia-oxidizers in the closed system either benefited from the minimized loss of the added bicarbonate (via CO_2 off-gassing) or from potentially increased ammonia production from stressed sponge clones.

To determine the contribution of the seawater microbial community to net nitrification, rates were also determined from aquaria seawater without sponges, revealing minimal nitrification with net rates of only $0.06 \pm 0.3 \mu\text{M N d}^{-1}$ ($n = 12$).

To calculate cell specific ammonia oxidation rates for *Ca. N. bastadiensis*, its 16S rRNA genes were quantified at the end of the 7 day sponge incubation under ambient and +25 μM NH_4^+ treatments. Sponges subjected to ambient and +25 μM NH_4^+ treatments contained thaumarchaeal symbiont gene copy numbers of $8.58 \pm 4.9 \times 10^{10}$ and $1.64 \pm 0.12 \times 10^{10}$ per g wet weight (SE, $n = 3$, for both), respectively. By dividing the net nitrification rate on day 7 by the 16S rRNA gene copy number of *Ca. N. bastadiensis* on this day, average cell specific rates were estimated to be 0.11 ± 0.08 and $0.66 \pm 0.09 \text{ fmol NH}_4^+$ oxidized per cell per day (SE, $n = 3$, for both) for the ambient and +25 μM NH_4^+ incubations, respectively. Inferred cell specific ammonia oxidation rates for *Ca. N. bastadiensis* were ~20-120 times lower than what has been reported for *N. maritimus* (Martens-Habbena *et al.*, 2009) and ~6-65 times lower than coastal marine seawater (Wuchter *et al.*, 2006),

while the highest ammonia oxidation rate of *Ca. N. bastadiensis* was comparable to the lowest rates detected in the sponge *Phakellia ventilabrum* (Radax *et al.*, 2012a).

Furthermore, as ^{13}C -bicarbonate was added during all experiments (Supporting Information Fig. S6), we were able to measure the $\delta^{13}\text{C}$ values of sponge tissues and calculate inorganic carbon fixation rates to compare with the gross nitrification rates. PTIO was found to reduce the $\delta^{13}\text{C}$ value in both the ambient and +25 μM NH_4^+ treatments (Fig. 5C; both: $p < 0.05$, one-tailed t-test), while no significant differences were observed with the +100 μM NH_4^+ treatments. In addition, gross nitrification rates and carbon fixation were significantly positively correlated (Fig. 5D). These data suggest that *Ca. N. bastadiensis* contributes significantly to carbon fixation by the sponge holobiont. This finding is consistent with (i) the detection of key genes for CO_2 fixation in the *I. basta* thaumarchaeote metagenome bin, (ii) the detection of some of the respective proteins in the metaproteome, (iii) and the absence of autotrophic CO_2 fixation pathways in the metagenomic bins of the alpha- and gamma-symbionts (data not shown). In addition, the slope of the positive linear relationship between the gross nitrification rate and the carbon fixation rate (Fig. 5D ; 43.0), reflects the gross nitrification: carbon fixation ratio (N:C ratio), which is most likely over-estimated since the experimental setup of the intermittently closed aquaria still allowed for the loss of added ^{13}C -bicarbonate. In addition, CO_2 production derived from respiratory activities within the *I. basta* holobiont were not accounted for, which may also lead to an underestimation of carbon fixation rates. Despite these possible biases the N:C ratio determined for *I. basta* is within the range typically found for aquatic environmental samples (2-60)(Andersson *et al.*, 2006). The relatively high N:C ratio along with the low cell-specific ammonia-oxidation rates within the *I. basta* holobiont are consistent with the proteogenomic-derived hypothesis that *Ca. N. bastadiensis* is not sustaining its population using ammonia as its only substrate, but exploits peptides and/or amino acids as additional sources of energy, nitrogen, and carbon by growing as a mixotroph.

Surprisingly, nitrite-oxidizing bacteria (NOB) could not be detected in *I. basta* by amplicon sequencing, metagenomic sequencing, or with FISH. Consistent with the absence of NOB, NO_2^-

accumulated to high concentrations ranging from 8 – 21 μM in ambient treatments containing *I. basta* (Fig. 5B, Supporting Information Fig. S7), making this sponge one of the few natural systems
555 in which greater NO_2^- than NO_3^- concentrations occur (Brezonik and Lee, 1968; Lam *et al.*, 2011; Schaefer and Hollibaugh, 2017 and refs. therein). It will be interesting to explore in future studies whether nitrite production contributes (via nitrite toxicity; Camargo and Alonso, 2006 and references therein) to protection from predators or to the unusually low microbial diversity in *I. basta*.

560 Given the absence of NOB in *I. basta*, the observed production of nitrate is difficult to explain. One possibility is that some NO released by either the host or a host-associated microorganism is detoxified to nitrate by another member of the sponge holobiont. However, we did not detect genes with homology to those encoding the NO detoxifying enzyme, flavohemoglobin-NO-dioxygenase (Hmp), which converts NO together with O_2 to NO_3^- (Gardner,
565 2012 and references therein), in our microbial metagenomic datasets. Still, NO dioxygenases (NODs), could be encoded in the sponge genome since neuroglobin-like sequences are encoded in other sponges (Lechauve *et al.*, 2013) and members of this enzyme family have been shown to have NOD activity *in vitro* (Brunori *et al.*, 2005). Furthermore, a NADH-cytochrome *b5* reductase, which belongs to the same family as Hmp, was found to be highly expressed at all life stages of the
570 sponge *Amphimedon queenslandica* (Conaco *et al.*, 2012). Alternatively, partial nitrite oxidation may be catalyzed by free-living NOB in the seawater, although nitrification rates in the control were negligible.

Conclusions and outlook. In this study we combined proteogenomic and experimental analyses to
575 show that *Ca. N. bastadiensis*, the first characterized but yet uncultured representative of a new genus within the thaumarchaeotes, is responsible for ammonia oxidation in the widespread marine sponge *I. basta*. While *Ca. N. bastadiensis* is equipped with the typical genetic repertoire of free-living AOA, it also exhibits a number of putative adaptations to a host-associated lifestyle. Several

of these adaptive features were unique to *Ca. N. bastadiensis*, whereas others were shared
580 exclusively with the previously described sponge symbionts *Ca. C. symbiosum* (Hallam *et al.*,
2006) and the AOA symbiont of the sponge *C. concentrica* (Moitinho-Silva *et al.*, 2017a). Many of
these putative adaptations to a sponge-associated lifestyle were not encoded in the recently
sequenced genome of an AOA from a glass sponge (Tian *et al.*, 2016), indicating that this archaeon
is not an obligate symbiont or that thriving in a glass sponge requires very different traits. Our
585 results confirm an emerging view that marine sponge microbiomes tend to converge on shared
functional traits, a process shaped by the environmental niche provided by the sponge host and
governed by specific ecological factors such as high dissolved nutrient loads and frequent contact
with resident and transient microorganisms (Liu *et al.*, 2012). Whereas other studies have focused
on this phenomenon by comparing functional convergence across entire sponge microbiomes (Fan
590 *et al.*, 2012; Liu *et al.*, 2012; Horn *et al.*, 2016), we demonstrate that in symbiotic marine sponge
Thaumarchaeota, a similar evolutionary convergence is achieved that stands in contrast to their
strictly chemoautotrophic non-symbiotic free-living relatives and that this process is likely not
solely achieved by gene acquisition but rather by selective gene retention and gene family
expansion.

595 With *I. basta* emerging as a model species for sponge symbiosis research, additional work
should be undertaken to address the conspicuous absence of nitrite-oxidizing bacteria and ascertain
why, in contrast to most tropical sponge species, *I. basta* hosts such a low diversity of microbial
symbionts. Furthermore, hypotheses about the interaction of *Ca. N. bastadiensis* with other
members of the sponge holobiont should be experimentally confirmed. For example, assimilation of
600 amino acids and peptides should be tested using stable isotope probing and the function of unique
serine proteases, serpins and ELPs should be analyzed via heterologous gene expression to reveal
mechanistic insights into the interaction of this archaeon with its host. Finally, the mechanism for
symbiont acquisition should be assessed via screening for *Ca. N. bastadiensis* in gametes and
larvae.

605 **Experimental Procedures**

Sponge collection. Large adult specimens (n=4) of the sponge *Ianthella basta* were collected from Orpheus Island (18°36.878'S, 146°29.990'E), Queensland, Australia, cut into 10 cm x 10 cm explants and transferred to racks on the reef. After a 12 week healing period in the field, sponge clones were collected in two separate sampling trips by scuba diving between September and
610 October 2011 and transported to the indoor temperature-controlled aquarium at the Australian Institute of Marine Science (AIMS), Townsville, where they were acclimated at ambient temperature (25 °C) for 48 h and then randomly assigned to experimental treatments. The two adult *I. basta* specimens used for metaproteogenomic analyses were collected from Orpheus Island in October 2010 and 2011 at Orpheus Island. Upon sample collection, specimens were cut into small
615 strips, immediately snap-frozen in liquid N₂ and subsequently stored at -80 °C until DNA extraction.

Sponge incubations and nitrification rate measurements. To infer nitrification activity of the *I. basta* holobiont, a 7-day incubation experiment was performed using triplicate sponge clones under
620 different ammonium concentrations: ambient ($0.29 \pm 0.1 \mu\text{M}$), 25 and 100 $\mu\text{M NH}_4^+$. A schematic representation of the experimental design is given in Supporting Information Fig. S6. Specifically, sponge clones (1.5 – 13.0 g post-experimental wet weight; mean = 5.1; standard deviation = 2.2) were incubated at 25 °C in the dark in 1.5 L acid-washed glass containers completely filled with 5 μm filtered, seawater or with seawater amended with 25 μM or 100 $\mu\text{M NH}_4\text{Cl}$. Additionally,
625 abiotic control experiments were performed with ambient seawater without sponge clones that were incubated for 24 h under identical conditions. Sponge clones were transferred to a new container with fresh seawater every 24 h during the 7-day incubation, in order to reduce the effects of O₂ depletion and NH₄⁺ accumulation. To assess carbon fixation in concert with nitrification, NaH¹³CO₃⁻ (100 μM ; 99% ¹³C) was added every 12 h to all experimental treatments. To prevent

630 loss of ^{13}C -labeled CO_2 , containers were closed for six hours after the $\text{NaH}^{13}\text{CO}_3^-$ addition and subsequently opened until the next $\text{NaH}^{13}\text{CO}_3^-$ addition to avoid oxygen depletion. The effect of such intermittently closed containers on nitrification activity was assessed in a control experiment by comparing ambient seawater and $100\ \mu\text{M}$ NH_4Cl amended seawater in constantly open containers (Supporting Information Fig. S7). Nitrification activity was determined as gross
635 nitrification rates using ^{15}N isotope pool dilution technique (Inselsbacher *et al.*, 2007), and net fluxes of individual dissolved inorganic N species (DIN), namely NH_4^+ , NO_2^- , NO_3^- . Gross nitrification rates were also measured on day 1, 3 and 5, where at the beginning of each day K^{15}NO_3 was added to a final $^{15}\text{NO}_3^-$ concentration of 1 – 10% of the nitrate pool (Murphy *et al.*, 2003) in the filtered ambient seawater. Gross nitrification rates were calculated from seawater samples collected
640 after a short equilibration following ^{15}N -label addition (0 h) and after 20 h. Net fluxes of NH_4^+ , NO_2^- , and NO_3^- were measured on day 1, 3, 5 and 7 and were calculated as change in concentration over time (i.e., between 6 and 18 h after ^{15}N -label addition to avoid biased net NO_3^- fluxes through a stimulation of consumptive processes from the added $^{15}\text{NO}_3^-$). Net nitrification refers to the change in concentration of $\text{NO}_2^- + \text{NO}_3^-$. Furthermore, we used the AOA-specific inhibitor PTIO (2-
645 phenyl-4,4,5,5-tetramethylimidazoline-1-oxyl 3-oxide; Tokyo Chemical Industry) (Martens-Habbena *et al.*, 2015) in additional parallel incubations, which was added at a concentration of $75\ \mu\text{M}$ at the beginning of day 3 and 7.

Seawater samples (10 mL) were taken from each aquarium and filtered using $0.45\ \mu\text{m}$ Sartorius Minisart cellulose acetate filters (Göttingen, Germany). Duplicate samples for dissolved
650 inorganic nitrogen (NH_4^+ , NO_2^- , NO_3^-) were measured on a Seal AA3 segmented flow analyzer and referenced against OSIL standards and in-house reference samples. For ^{15}N -analysis of $\text{NO}_2^- + \text{NO}_3^-$, sample water was filtered through pre-combusted GF/Fs (Whatman International; treated for 4 h at 450°C), and subsequently through $0.2\ \mu\text{m}$ filters (Sartorius). All samples were immediately frozen at -20°C for later analysis. Prior to shipment to the University of Vienna, samples were thawed at
655 room temperature and the microbial inhibitor phenylmercuric acetate was added (to a final

concentration of 10 μM). Upon arrival in Vienna the samples were promptly stored at -80°C . Nitrite and nitrate were isolated together from seawater by sequential microdiffusion (Sørensen and Jensen, 1991). To remove ammonium from sample water, 100 mg MgO, and an acid trap (acidified cellulose filter disc enclosed in a semi-permeable Teflon membrane) was added to 9 mL of sample
660 and 1.5 mL of 3 M KCl. After 5 days shaking at 35°C , the acid traps were removed, and 50 mg of Devarda's alloy was added along with a new acid trap, and shaken at 35°C for 7 days. Devarda's alloy is a reducing catalyst converting both NO_2^- and NO_3^- to NH_4^+ and the subsequently formed NH_3 was collected in the acid trap. Acid traps were dried over concentrated sulfuric acid and analyzed for ^{15}N by an elemental analyser (EA 1110, CE Instruments, Milan, Italy) coupled to an
665 isotope ratio mass spectrometer (IRMS; Finnigan MAT Delta^{Plus} IRMS with a Finnigan MAT ConFlo III interface). Gross nitrification rates were calculated based on Wanek *et al.* (2010). Gross and net rates are expressed as μmol nitrogen species per g wet weight sponge per day ($\mu\text{mol N g}^{-1} \text{d}^{-1}$).

For the determination of ^{13}C enrichment in whole sponge tissue at the end of the incubation,
670 sponge tissue was freeze-dried, ground to a fine powder, and stored at dry conditions prior to analysis. The $\delta^{13}\text{C}$ values of sponge tissue were determined using an EA-IRMS system as described above. Carbon fixation rates were calculated based on the $\delta^{13}\text{C}$ values treatments amended with 100 μM $\text{NaH}^{13}\text{CO}_3^-$. For the calculations, we used a background (natural abundance) bicarbonate concentration of 1975 μM .

675

DNA extraction from whole sponge tissue and qPCR for symbiont quantification.

Between 80-150 mg of *I. basta* tissue was thawed, rinsed successively (3x) in 1X calcium- and magnesium-free artificial seawater (CMF-ASW) and immediately ground into a paste with a mortar and pestle in liquid N_2 . After resuspension in TE buffer (10 mM Tris-HCL pH 8.0, 1 mM
680 EDTA), DNA was extracted from the suspension using an adapted SDS-based isolation method

(Zhou *et al.*, 1996) and using 1% polyvinylpyrrolidone. DNA was extracted from the two individuals used for metaproteogenomics (see below) as well as three healthy individuals from a previous study (Luter *et al.*, 2010). Additionally, DNA was extracted from a subset of sponge clones that were subjected to nitrification incubations. Quantitative PCR (qPCR) was used to estimate the number of specific thaumarchaeal, as well as α - and γ -proteobacterial symbionts by quantifying the 16S rRNA gene using specific primers designed for each symbiont phylotype. The following primer sets were thus used for the dominant *I. basta* thaumarchaeal, and α -, and γ -proteobacterial symbionts respectively; IBthaum16S_523F, 5'-CCG TAG CCT GCC CTG TTA G-3', IBthaum16S_727R, 5'-GCT TTC ATC CCT CAC CGT-3'; IBalpha16S_1010F, 5'-CGG AGA CGC TTC CTT CG -3', IBalpha16S_1206R, 5'-GCC CAG CCC ATA AAT GC-3'; IBgamma16S_466F, 5'-TAC CCY TGY GTT TTG ACG-3', IBgamma16S_655R, 5'- CCR CTT CTC TCT RCC ATA C-3'. An iCycler real-time PCR system (Bio-Rad, Hercules, CA) was used to measure all samples, in duplicate wells per reaction and reactions were performed in a 25 μ L volume with 1 μ L of DNA template. All symbiont 16S rRNA gene qPCR assays used SYBR®Green reaction mixtures containing 12.5 μ L iQ SYBR®Green Supermix (Bio-Rad), and optimized concentrations of 400 nM primer as well as 0.25 mg ml⁻¹ BSA. Cycling conditions were 95°C for 5 min followed by 40 cycles of 95°C for 40 s, 58°C for 30 s, and 72°C for 40 s. Fluorescence intensities were measured after each cycle, and a final elongation at 72°C was followed by a melting curve analysis from 55-95°C in 10 s increments of 0.5°C.

Standard curves were generated for each primer set using serial dilutions of a standard containing a known number of the target sequences. Standards used the M13 primer set to amplify 16S rRNA gene clones derived from the sponge symbionts. PCR products were visualized on an agarose gel, purified separately using the QIAquick PCR Purification Kit (Qiagen), followed by fluorometric quantification of DNA concentrations using PicoGreen (Molecular Probes, Eugene, OR) and a NanoDrop ND-3300 Fluorospectrometer (NanoDrop). Gene abundance was calculated based on DNA concentration and product size. Dilution series ranging from 10⁶ to 10⁰ copies μ L⁻¹

were used to generate standard curves. Final 16S rRNA gene abundances for each microbial symbiont were then normalized by the wet weight of the sponge tissue used for DNA extraction.

710 **Cryosectioning and FISH for symbiont quantification.** The *I. basta* individual collected for metagenomic sequencing in October 2010 was also assessed using FISH. Briefly, after sample collection, the *I. basta* specimen was cut into tissue strips of approximately 2 mm³, fixed in 4% PFA for 1 h at room temperature and stored in ethyl alcohol (EtOH)-phosphate-buffered saline (PBS) at -20°C. For FISH, PFA-fixed samples of *I. basta* were embedded in Neg-50 (Richard-Allan
715 Scientific), and cut to 5-µm sections (Leica CM3050 S). A double-labeled Arch915 probe in Cy3 (Thermo Fisher Scientific, Waltham, MA, USA) was used for the microscopic visualization and quantification of the thaumarchaeal symbiont of *I. basta*. To calculate the relative abundance of the thaumarchaeal symbiont, equimolar amounts of the double-labeled probes EUB338-I, EUB338-II, and EUB338-III (Fluos and Cy5) (Stoecker *et al.*, 2009) were used for quantification of most
720 bacteria. Hybridizations were prepared with an equimolar mixture of both probes and using 25% formamide in the hybridization buffer, with the stringency of the washing buffer adjusted accordingly. As a negative control, the nonEUB338-I (reverse complementary probe to EUB338-I) was applied on one section per well per slide hybridized (Wallner *et al.*, 1993). All hybridized samples were analyzed with a confocal laser scanning microscope (CLSM) (LSM 510 Meta; Zeiss,
725 Oberkochen, Germany). Archaeal and bacterial cells were counted by eye on 10 randomly selected images derived from multiple tissue sections obtained from a single *I. basta* individual, and the proportion of archaeal cells to total prokaryotic cells was calculated.

Microbial cell enrichment for metaproteogenomics. To separate symbiont cells from host tissue
730 prior to DNA extraction, approximately 15.7 g wet weight of *I. basta* was rinsed successively (3x) in 1X CMF-ASW. Sponge tissue was cut into small pieces (<1 cm³), ground on ice in 1X CMF-ASW with a mortar and pestle, transferred into a glass douncer on ice, and sponge tissue was

dissociated through shear force and vortexing. The supernatant was transferred into multiple Eppendorf microcentrifuge tubes, centrifuged at 39 x g for 15 min at 4°C to pellet larger sponge particles, filtered through a 5 µm Sartorius filter and centrifuged at 10,844 x g for 15 min at 4 °C to pellet microbial cells. Microscopic examination of the pellet revealed an enrichment of microbial cells and an absence of sponge nuclei. For the sponge individual sampled in October 2010, DNA was extracted from this pellet using the procedure referenced above. For the individual sampled in October 2011, the pellets were resuspended in 1X CMF-ASW and the suspension was layered in 1.8 ml amounts onto 6.5 ml cushions of 30% Gastrografin dissolved in 1X CMF-ASW + 0.2 M EDTA, and centrifuged at 40,008 x g for 1 hour at 4°C in an Beckman L-100 XP ultracentrifuge with the SW 41 Ti swing rotor (Beckman Coulter). Following centrifugation, a cell-rich layer above the 30% Gastrografin cushion ('Fraction A') and the cell-rich pellet ('Fraction B') were carefully removed, resuspended in 10 mM Tris, and re-pelleted at 24,400 x g for 15 min at 4°C in a new solution of 10 mM Tris.

The same *I. basta* individual sampled in October 2011 was used for proteomic analyses. However, in order to minimize protein degradation, biomass preparation methods were shortened and modified. The first cell fraction used for proteomic analysis involved rinsing in 1X CMF-ASW and tissue homogenization using a mortar/pestle and glass douncer as described above. However, the tissue homogenization steps occurred in 1X TE buffer with Roche Complete Protease Inhibitor (Roche). Supernatant from the glass douncer was immediately collected into 2 mL Eppendorf tubes and frozen at -80°C before shipment on dry ice to Greifswald, Germany, for protein extraction and downstream analyses. A second cell fraction comprised a 5 µm filtrate from the supernatant in the glass douncer. Finally, a crude sponge homogenate sample was obtained by direct grinding of sponge tissue (without rinsing in 1X CMF-ASW) in liquid N₂ followed by freezing at -80°C in 1X TE with Roche Complete. These three fractions could thus be characterized as a sponge homogenate without sponge skeleton, a sponge homogenate without sponge nuclei, and a direct sponge homogenate, respectively.

760 **DNA extraction, library preparation and sequencing.** DNA was extracted from the individual
sampled in October 2011 using the FastDNA spin kit for soil (MP Biomedicals, Solon, OH, USA)
from the cell-rich layers above the 30% Gastrografin cushion and in the ultracentrifuged pellet
according to the manufacturer's instructions. Sequencing libraries were prepared using the Nextera
kit (Illumina Inc.) according to the manufacturer's instructions and concentrations measured using
765 the QuantIT kit (Molecular Probes, Life Technologies, Naerum, Denmark). The libraries were
paired-end (2x150 bp) sequenced on an Illumina HiSeq2000 using the TruSeq PE Cluster Kit v3-
cBot-HS and TruSeq SBS kit v.3-HS sequencing kit and on an Illumina MiSeq using v3 2x300 bp
kits.

For the individual sampled in October 2010, cell pellets were pooled from the 5 μ m
770 filtration step and DNA was extracted using the modified protocol of Zhou *et al.* (1996) described
above. Metagenomic sequences were then generated at the Ramaciotti Sequencing Centre (Sydney,
Australia) using the GS FLX instrument using Titanium chemistry (Roche) on a 454 half-
sequencing-plate (454 Life Sciences, Branford, CT, USA).

775 **Metagenome assembly and genome binning.** Metagenome reads in fastq format, obtained from
the Illumina sequencing runs, were end-trimmed at a minimum phred score of 15, a minimum
length of 50 bp, allowing no ambiguous nucleotides and Illumina sequencing adaptors removed.
Trimmed reads from each dataset were assembled using Spades version 3.11.0 (Bankevich *et al.*,
2012), using default parameters and genomes were binned using Metabat v 2.12.0 (Kang *et al.*,
780 2015). MAGs from multiple Illumina datasets were dereplicated using dRep (Olm *et al.*, 2017) and
provisionally classified using CheckM (Parks *et al.*, 2015). A single high-quality archaeal MAG
was recovered after dereplication and uploaded to MaGe (Vallenet *et al.*, 2009) for annotation. This

MAG has been submitted to the European Nucleotide Archive with the accession number PRJEB29556.

785 For the 454-pyrosequencing run, artificially amplified reads were dereplicated using CD-
HIT (Li and Godzik, 2006), assembled with MIRA (Chevreux *et al.*, 1999) and binned with a
genome-specific Phymm model (Brady and Salzberg, 2009), which was trained from the contigs
that contained phylogenetic marker genes. The MAG derived from this pyrosequencing run was
similarly GC-rich (64.2%), but considerably larger (6.58 Mb), slightly less complete (98.06%), and
790 considerably more fragmented (2508 scaffolds). This MAG is available on the MaGe platform as
“*Ianthella basta symbiont* thaum” at <http://www.genoscope.cns.fr/agc/microscope/home/index.php>.

Comparative genomics. The annotated archaeal MAG was downloaded from MaGe and compared
to published thaumarchaeotal genomes (Supporting Information Table S5) using genomic average
795 nucleotide identity (gANI), average amino acid identity (AAI), and through construction of
orthologous gene families. For all analyses, annotated genes were supplemented with additional
gene calls predicted by Prodigal (Hyatt *et al.*, 2010). gANI was calculated with MiSI (Varghese *et al.*,
et al., 2015). AAI (Konstantinidis and Tiedje, 2005) was calculated using bidirectional best blastp hits
(Camacho *et al.*, 2009) that aligned over at least 70% of gene length with average identity values
800 weighted according to gene length. Orthologous gene families were constructed using Orthofinder
(Emms and Kelly, 2015). For functional annotation of eukaryotic-like proteins (ELPs) and
mobile/selfish genetic elements, predicted genes from all sequenced thaumarchaeal genomes were
searched against the Protein Family A (Pfam-A) database (v31.0) (Finn *et al.*, 2014) using Hmmer
3 (hmmer.org.) and the gathering threshold option (-cutga). Results were screened for the presence
805 of domains associated with ELPs and mobile/selfish genetic elements.

Phylogenetic analyses. Bayesian trees were constructed using Phylobayes v 4.1c (Lartillot and Philippe, 2004) using the best model identified for each dataset by ModelFinder (Kalyaanmoorthy *et al.*, 2017). Phylogenomic reconstruction was based on a concatenated amino-acid alignment of
810 34 marker genes constructed with CheckM (Parks *et al.*, 2015) with ten independent runs of 11000 generations under the LG4 model. 6000 generations of each independent run were discarded as burn-in and the remaining trees from each run were pooled for calculation of a consensus tree and for determining posterior branch support. For the 16S rRNA gene phylogenies, top representative hits in a blastn query against the Genbank nr database of the full-length sequence from *Ca. N.*
815 *bastadiensis* and *Ca. C. symbiosum*, along with sequences from sequenced thaumarchaeal genomes were aligned with SINA (Pruesse *et al.*, 2012) and analyzed further as described above. For the *amoA* gene tree, the *Ca. N. bastadiensis* and *Ca. C. symbiosum amoA* sequences were placed into a reference tree, representing all OTU representatives of the curated database provided in Alves *et al.* (2018), using the Evolutionary Placement Algorithm (EPA; Berger *et al.*, 2011) implemented in
820 RAxML-HPC 8.2.11 (Stamatakis, 2014). Representative sequences clustering with the *Ca. N. bastadiensis amoA* gene sequence along with *amoA* sequences from the aforementioned genome-sequenced thaumarchaeota were then aligned with MUSCLE (Edgar, 2004) and analyzed. For both 16S rRNA and *amoA* gene phylogenies, ten independent runs of 30000 generations under the GTR model were used. 7500 generations of each independent run were discarded as burn-ins and the
825 remaining trees from each run were pooled for calculation of a consensus tree and for determining posterior branch support.

Protein identification and proteome analyses. 1D PAGE followed by liquid chromatography-based mass spectrometry (1D-PAGE-LC-MS/MS) were used for protein and peptide separation and
830 identification as described previously (Washburn *et al.*, 2001; Otto *et al.*, 2010), with slight modifications. MS spectra and MS/MS spectra were acquired from eluting peptides ionized with electrospray ionization (ESI) and analyzed in a LTQ Orbitrap Velos hybrid mass spectrometer

(Thermo Fisher Scientific, Waltham, MA, USA), as described previously (Verberkmoes *et al.*, 2009; Otto *et al.*, 2010), with minor modifications. Samples of the sponge homogenate without
835 sponge skeleton and crude sponge homogenate processed in liquid N₂ were analyzed in technical
duplicates, whereas the sponge homogenate without sponge nuclei was analyzed only once. All
MS/MS spectra were searched against predicted protein sequence databases composed of the *I.*
basta symbiont-enriched metagenome bins and common laboratory contaminants using the Sorcerer
SEQUEST (v.27, rev. 11) algorithm. The CD-HIT software (Li and Godzik, 2006) was used to
840 remove redundancies from the database due to the potential for strain-level redundancies. Protein
identifications were filtered with Scaffold version 3.5.1 applying the “sequest” filter criteria
described previously (Heinz *et al.* 2012). For protein identification only peptides identified with
high mass accuracy (maximum \pm 10 ppm difference between calculated and observed mass) were
considered and at least two exclusively unique peptides were required to identify a protein. False-
845 discovery rates (FDRs) were estimated with searches against a target-decoy database as described
previously (Peng *et al.*, 2003; Elias and Gygi, 2007). Peptide FDRs were between 2.5% and 3.1%,
and protein FDRs were below 0.4% throughout all samples. For relative quantitation of proteins,
normalized spectral abundance factor were calculated for each sample according to the method of
Florens *et al.* (2006) and averaged for all replicates and samples. The proteomics dataset for this
850 study have been submitted to the PRIDE archive.

Acknowledgements

This work was supported by the Marie Curie Initial Training Network - SYMBIOMICS. In
addition, F.U.M., C.W.H, M.M., and M.W. were supported by the European Research Council
855 Advanced Grant project NITRICARE 294343. M.A. and P.H.N. were supported by research grants
(15510 and 16578) from VILLUM FONDEN. Many thanks to Christian Hentschker for Orbitrap
Velos measurements and database searches.

References

- 860 Adams, M.D., Wagner, L.M., Graddis, T.J., Landick, R., Antonucci, T.K., Gibson, A.L., and
Oxender, D.L. (1990) Nucleotide sequence and genetic characterization reveal six
essential genes for the LIV-I and LS transport systems of *Escherichia coli*. *J Biol Chem*
265: 11436–11443.
- 865 Agarwal, V., Blanton, J.M., Podell, S., Taton, A., Schorn, M.A., Busch, J., *et al.* (2017)
Metagenomic discovery of polybrominated diphenyl ether biosynthesis by marine
sponges. *Nat Chem Biol* **13**: 537–543.
- Ahlgren, N.A., Chen, Y., Needham, D.M., Parada, A.E., Sachdeva, R., Trinh, V., *et al.* (2017)
Genome and epigenome of a novel marine Thaumarchaeota strain suggest viral infection,
phosphorothioation DNA modification and multiple restriction systems: Genome and
epigenome of a new marine thaumarchaeon. *Environ Microbiol* **19**: 2434–2452.
- 870 Alonso-Saez, L., Waller, A.S., Mende, D.R., Bakker, K., Farnelid, H., Yager, P.L., *et al.*
(2012) Role for urea in nitrification by polar marine Archaea. *Proc Natl Acad Sci USA*
109: 17989–17994.
- 875 Alves, R.J.E., Minh, B.Q., Urich, T., von Haeseler, A., and Schleper, C. (2018) Unifying the
global phylogeny and environmental distribution of ammonia-oxidising archaea based on
amoA genes. *Nat Commun* **9**: 1517.
- Amano, F. and Noda, T. (1995) Improved detection of nitric oxide radical (NO.) production in
an activated macrophage culture with a radical scavenger, carboxy PTIO and Griess
reagent. *FEBS Lett* **368**: 425–428.
- 880 Andersson, M., Brion, N., and Middelburg, J. (2006) Comparison of nitrifier activity versus
growth in the Scheldt estuarya turbid, tidal estuary in northern Europe. *Aquat Microb Ecol*
42: 149–158.
- Arreguín, R., Arreguín, B., Soriano-García, M., Hernández-Arana, A., and Rodríguez-Romero,
A. (1993) Isolation and characterization of a protease from the marine sponge
Sphaciospongia vesparia. *FEBS Lett* **320**: 235–238.
- 885 Astudillo-García, C., Bell, J.J., Webster, N.S., Glasl, B., Jompa, J., Montoya, J.M., and Taylor,
M.W. (2017) Evaluating the core microbiota in complex communities: A systematic
investigation. *Environ Microbiol* **19**: 1450–1462.
- 890 Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., *et al.*
(2012) SPAdes: A new genome assembly algorithm and its applications to single-cell
sequencing. *J Comput Biol* **19**: 455–477.

- 895 Bayer, K., Schmitt, S., and Hentschel, U. (2007) Microbial nitrification in Mediterranean sponges: possible involvement of ammonium-oxidizing *Betaproteobacteria*. In *Porifera Research: Biodiversity, Innovation, Sustainability*. Custódio, M., Lôbo-Hajdu, G., Hajdu, E., and Muricy, G. (eds). Série Livros. Rio de Janeiro, Brazil: Museu Nacional, pp. 165–171.
- Bayer, K., Schmitt, S., and Hentschel, U. (2008) Physiology, phylogeny and *in situ* evidence for bacterial and archaeal nitrifiers in the marine sponge *Aplysina aerophoba*. *Environ Microbiol* **10**: 2942–2955.
- 900 Bayer, B., Vojvoda, J., Offre, P., Alves, R.J.E., Elisabeth, N.H., Garcia, J.A., *et al.* (2016) Physiological and genomic characterization of two novel marine thaumarchaeal strains indicates niche differentiation. *ISME J* **10**: 1051–1063.
- Beam, J.P., Jay, Z.J., Kozubal, M.A., and Inskeep, W.P. (2014) Niche specialization of novel Thaumarchaeota to oxic and hypoxic acidic geothermal springs of Yellowstone National Park. *ISME J* **8**: 938–951.
- 905 Beckers, G., Bendt, A.K., Kramer, R., and Burkovski, A. (2004) Molecular identification of the urea uptake system and transcriptional analysis of urea transporter- and urease-encoding genes in *Corynebacterium glutamicum*. *J Bacteriol* **186**: 7645–7652.
- Bell, J.J. (2008) The functional roles of marine sponges. *Estuar Coast Shelf Sci* **79**: 341–353.
- 910 Bergquist PR, Kelly-Borges M. (1995) Systematics and biogeography of the genus *Ianthella* (Demospongiae: Verongida: Ianthellidae) in the south-west Pacific. *The Beagle, Records of the Museums and Art Galleries of the Northern Territory* **12**:151–176.
- Blainey, P.C., Mosier, A.C., Potanina, A., Francis, C.A., and Quake, S.R. (2011) Genome of a low-salinity ammonia-oxidizing archaeon determined by single-cell and metagenomic analysis. *PLoS ONE* **6**: e16626.
- 915 Boardman, G.D., Starbuck, S.M., Hudgins, D.B., Li, X., and Kuhn, D.D. (2004) Toxicity of ammonia to three marine fish and three marine invertebrates. *Environ Toxicol* **19**: 134–142.
- de Boer, W. and Laanbroek, H.J. (1989) Ureolytic nitrification at low pH by *Nitrosospira* spec. *Arch Microbiol* **152**: 178–181.
- 920 Brady, A. and Salzberg, S.L. (2009) Phymm and PhymmBL: metagenomic phylogenetic classification with interpolated Markov models. *Nat Methods* **6**: 673–676.
- Brezonik, P.L. and Lee, G.F. (1968) Denitrification as a nitrogen sink in Lake Mendota, Wisconsin. *Environ Sci Technol* **2**: 120–125.

- 925 Brunori, M., Giuffre, A., Nienhaus, K., Nienhaus, G.U., Scandurra, F.M., and Vallone, B.
(2005) Neuroglobin, nitric oxide, and oxygen: Functional pathways and conformational
changes. *Proc Natl Acad Sci USA* **102**: 8483–8488.
- Burton, S.A.Q. and Prosser, J.I. (2001) Autotrophic ammonia oxidation at low pH through urea
hydrolysis. *Appl Environ Microbiol* **67**: 2952–2957.
- 930 Callebaut, I., Mornon, J.-P., Gilgès, D., and Vigon, I. (2000) HYR, an extracellular module
involved in cellular adhesion and related to the immunoglobulin-like fold. *Protein Sci* **9**:
1382–1390.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden,
T.L. (2009) BLAST+: architecture and applications. *BMC Bioinformatics* **10**: 421.
- 935 Camargo, J.A. and Alonso, Á. (2006) Ecological and toxicological effects of inorganic nitrogen
pollution in aquatic ecosystems: A global assessment. *Environ Int* **32**: 831–849.
- Carini, P., Dupont, C.L., and Santoro, A.E. (2018) Patterns of thaumarchaeal gene expression
in culture and diverse marine environments. *Environ Microbiol* **20**: 2112–2124.
- 940 Chevreux, B., Pfisterer, T., Drescher, B. *et al.* (1999) Genome Sequence Assembly Using
Trace Signals and Additional Sequence Information. In: *Computer Science and Biology:
Proceedings of the German Conference on Bioinformatics (GCB) 99*. pp. 45–46.
- Conaco, C., Neveu, P., Zhou, H., Arcila, M., Degnan, S.M., Degnan, B.M., and Kosik, K.S.
(2012) Transcriptome profiling of the demosponge *Amphimedon queenslandica* reveals
genome-wide events that accompany major life cycle transitions. *BMC Genomics* **13**: 209.
- 945 Corredor, J.E., Wilkinson, C.R., Vicente, V.P., Morell, J.M., and Otero, E. (1988) Nitrate
release by Caribbean reef sponges. *Limnol Oceanogr* **33**: 114–120.
- Cort, J.R., Yee, A., Edwards, A.M., Arrowsmith, C.H., and Kennedy, M.A. (2000) Structure-
based functional classification of hypothetical protein MTH538 from *Methanobacterium
thermoautotrophicum*. *J Mol Biol* **302**: 189–203.
- 950 Daims, H., Ramsing, N.B., Schleifer, K.-H., and Wagner, M. (2001) Cultivation-independent,
semiautomatic determination of absolute bacterial cell numbers in environmental samples
by fluorescence in situ hybridization. *Appl Environ Microbiol* **67**: 5810–5818.
- Daims, H., Lebedeva, E.V., Pjevac, P., Han, P., Herbold, C., Albertsen, M., *et al.* (2015)
Complete nitrification by *Nitrospira* bacteria. *Nature* **528**: 504–509.
- 955 Díez-Vives, C., Moitinho-Silva, L., Nielsen, S., Reynolds, D., and Thomas, T. (2017)
Expression of eukaryotic-like protein in the microbiome of sponges. *Mol Ecol* **26**: 1432–
1451.

- Drapeau, G.R. (1980) Substrate specificity of a proteolytic enzyme isolated from a mutant of *Pseudomonas fragi*. *J Biol Chem* **255**: 839–840.
- Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high
960 throughput. *Nucleic Acids Res* **32**: 17921797.
- Elias, J.E. and Gygi, S.P. (2007) Target-decoy search strategy for increased confidence in large-scale protein identifications by mass spectrometry. *Nat Methods* **4**: 207–214.
- Ellis, A., Lu, H., Li, C.G., and Rand, M.J. (2001) Effects of agents that inactivate free radical NO (NO center dot) on nitroxyl anion-mediated relaxations, and on the detection of NO
965 center dot released from the nitroxyl anion donor Angeli's salt. *Brit J Pharmacol* **134**: 521–528.
- Emms, D.M. and Kelly, S. (2015) OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol* **16**: 157.
- Essuman, K., Summers, D.W., Sasaki, Y., Mao, X., Yim, A.K.Y., DiAntonio, A., and
970 Milbrandt, J. (2018) TIR domain proteins are an ancient family of NAD(+)-consuming enzymes. *Curr Biol* **28**: 421-430.
- Fan, L., Reynolds, D., Liu, M., Stark, M., Kjelleberg, S., Webster, N.S., and Thomas, T. (2012) Functional equivalence and evolutionary convergence in complex communities of microbial sponge symbionts. *Proc Natl Acad Sci USA* **109**: E1878–E1887.
- 975 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**: D222–D230.
- Fiore, C.L., Labrie, M., Jarett, J.K., and Lesser, M.P. (2015) Transcriptional activity of the giant barrel sponge, *Xestospongia muta* Holobiont: molecular evidence for metabolic interchange. *Front Microbiol* **6**: 364.
- 980 Florens, L., Carozza, M., Swanson, S., Fournier, M., Coleman, M., Workman, J., and Washburn, M. (2006) Analyzing chromatin remodeling complexes using shotgun proteomics and normalized spectral abundance factors. *Methods* **40**: 303–311.
- Freckelton, M.L., Luter, H.M., Andreakis, N., Webster, N.S., and Motti, C.A. (2012) Qualitative variation in colour morphotypes of *Ianthella basta* (Porifera: Verongida).
985 *Hydrobiologia* **687**: 191–203.
- Freeman, M.F., Vagstad, A.L., and Piel, J. (2016) Polytheonamide biosynthesis showcasing the metabolic potential of sponge-associated uncultivated 'Entotheonella' bacteria. *Curr Opin Chem Biol* **31**: 8–14.
- Gardner, P.R. (2012) Hemoglobin: a nitric-oxide dioxygenase. *Scientifica* ID 683729.

- 990 Gauthier, M.-E.A., Watson, J.R., and Degnan, S.M. (2016) Draft genomes shed light on the dual bacterial symbiosis that dominates the microbiome of the coral reef sponge *Amphimedon queenslandica*. *Front Mar Sci* **3**: 196.1–196.18.
- de Goeij, J.M., van Oevelen, D., Vermeij, M.J.A., Osinga, R., Middelburg, J.J., de Goeij, A.F.P.M., and Admiraal, W. (2013) Surviving in a marine desert: the sponge loop retains resources within coral reefs. *Science* **342**: 108–110.
- 995 Guy, L. and Ettema, T.J.G. (2011) The archaeal ‘TACK’ superphylum and the origin of eukaryotes. *Trends Microbiol* **19**: 580–587.
- Hallam, S.J., Konstantinidis, K.T., Putnam, N., Schleper, C., Watanabe, Y., Sugahara, J., *et al.* (2006) Genomic analysis of the uncultivated marine crenarchaeote *Cenarchaeum symbiosum*. *Proc Natl Acad Sci USA* **103**: 18296–18301.
- 1000 Hatzenpichler, R., Lebedeva, E.V., Spieck, E., Stoecker, K., Richter, A., Daims, H., and Wagner, M. (2008) A moderately thermophilic ammonia-oxidizing crenarchaeote from a hot spring. *Proc Natl Acad Sci USA* **105**: 2134–2139.
- Heinz, E., Williams, T.A., Nakjang, S., Noël, C.J., Swan, D.C., Goldberg, A.V., *et al.* (2012) The genome of the obligate intracellular parasite *Trachipleistophora hominis*: New insights into microsporidian genome dynamics and reductive evolution. *PLoS Pathog* **8**: e1002979.
- 1005 Hentschel, U., Piel, J., Degnan, S.M., and Taylor, M.W. (2012) Genomic insights into the marine sponge microbiome. *Nat Rev Microbiol* **10**: 641–654.
- 1010 Herndl, G.J., Reinthaler, T., Teira, E., van Aken, H., Veth, C., Pernthaler, A., and Pernthaler, J. (2005) Contribution of *Archaea* to total prokaryotic production in the deep Atlantic Ocean. *Appl Environ Microbiol* **71**: 2303–2309.
- Hoffmann, F., Larsen, O., Thiel, V., Rapp, H.T., Pape, T., Michaelis, W., and Reitner, J. (2005) An anaerobic world in sponges. *Geomicrobiol J* **22**: 1–10.
- 1015 Hoffmann, F., Radax, R., Woebken, D., Holtappels, M., Lavik, G., Rapp, H.T., *et al.* (2009) Complex nitrogen cycling in the sponge *Geodia barretti*. *Environ Microbiol* **11**: 2228–2243.
- Horn, H., Slaby, B.M., Jahn, M.T., Bayer, K., Moitinho-Silva, L., Förster, F., *et al.* (2016) An enrichment of CRISPR and other defense-related features in marine sponge-associated microbial metagenomes. *Front Microbiol* **7**: 1751.
- 1020 Hoshino, T. (1979) Transport systems for branched-chain amino acids in *Pseudomonas aeruginosa*. *J Bacteriol* **139**: 705–712.
- Hosie, A.H.F., Allaway, D., Galloway, C.S., Dunsby, H.A., and Poole, P.S. (2002) *Rhizobium leguminosarum* has a second general amino acid permease with unusually broad substrate

- 1025 specificity and high similarity to branched-chain amino acid transporters (Bra/LIV) of the ABC family. *J Bacteriol* **184**: 4071–4080.
- Hosseinzadeh, P., Tian, S., Marshall, N.M., Hemp, J., Mullen, T., Nilges, M.J., *et al.* (2016) A purple cupredoxin from *Nitrosopumilus maritimus* containing a mononuclear type 1 copper center with an open binding site. *J Am Chem Soc* **138**: 6324–6327.
- 1030 Hyatt, D., Chen, G.-L., LoCascio, P.F., Land, M.L., Larimer, F.W., and Hauser, L.J. (2010) Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* **11**: 119.
- Ingalls, A.E., Shah, S.R., Hansman, R.L., Aluwihare, L.I., Santos, G.M., Druffel, E.R.M., and Pearson, A. (2006) Quantifying archaeal community autotrophy in the mesopelagic ocean using natural radiocarbon. *Proc Natl Acad Sci USA* **103**: 6442–6447.
- 1035 Inselsbacher, E., Cambui, C.A., Richter, A., Stange, C.F., Mercier, H., and Wanek, W. (2007) Microbial activities and foliar uptake of nitrogen in the epiphytic bromeliad *Vriesea gigantea*. *New Phytol* **175**: 311–320.
- Jernigan, K.K. and Bordenstein, S.R. (2015) Tandem-repeat protein domains across the tree of life. *PeerJ* **3**: e732.
- 1040 Jompa, J. and McCook, L.J. (2002) The effects of nutrients and herbivory on competition between a hard coral (*Porites cylindrica*) and a brown alga (*Lobophora variegata*). *Limnol Oceanogr* **47**: 527–534.
- Jung, M.-Y., Park, S.-J., Kim, S.-J., Kim, J.-G., Sinnighe Damsté, J.S., Jeon, C.O., and Rhee, S.-K. (2014) A mesophilic, autotrophic, ammonia-oxidizing archaeon of thaumarchaeal group I.1a cultivated from a deep oligotrophic soil horizon. *Appl Environ Microbiol* **80**: 3645–3655.
- 1045 Kalyaanamoorthy, S., Minh, B.Q., Wong, T.K.F., von Haeseler, A., and Jermiin, L.S. (2017) ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* **14**: 587–589.
- 1050 Kang, D.D., Froula, J., Egan, R., and Wang, Z. (2015) MetaBAT, an efficient tool for accurately reconstructing single genomes from complex microbial communities. *PeerJ* **3**: e1165.
- Karimi, E., Slaby, B.M., Soares, A.R., Blom, J., Hentschel, U., and Costa, R. (2018) Metagenomic binning reveals versatile nutrient cycling and distinct adaptive features in alphaproteobacterial symbionts of marine sponges. *FEMS Microbiol Ecol* **94**: 6.
- 1055 Kerou, M., Offre, P., Valledor, L., Abby, S.S., Melcher, M., Nagler, M., *et al.* (2016) Proteomics and comparative genomics of *Nitrososphaera viennensis* reveal the core

- genome and adaptations of archaeal ammonia oxidizers. *Proc Natl Acad Sci USA* **113**:
1060 E7937–E7946.
- Kim, J.-G., Park, S.-J., Sinninghe Damsté, J.S., Schouten, S., Rijpstra, W.I.C., Jung, M.-Y., *et al.* (2016) Hydrogen peroxide detoxification is a key mechanism for growth of ammonia-oxidizing archaea. *Proc Natl Acad Sci USA* **113**: 7888–7893.
- King, N., Westbrook, M.J., Young, S.L., Kuo, A., Abedin, M., Chapman, J., *et al.* (2008) The
1065 genome of the choanoflagellate *Monosiga brevicollis* and the origin of metazoans. *Nature*
451: 783–788.
- Koch, H., Galushko, A., Albertsen, M., Schintlmeister, A., Gruber-Dorninger, C., Lucker, S., *et al.* (2014) Growth of nitrite-oxidizing bacteria by aerobic hydrogen oxidation. *Science*
345: 1052–1054.
- 1070 Könneke, M., Schubert, D.M., Brown, P.C., Hugler, M., Standfest, S., Schwander, T., *et al.*
(2014) Ammonia-oxidizing archaea use the most energy-efficient aerobic pathway for
CO₂ fixation. *Proc Natl Acad Sci USA* **111**: 8239–8244.
- Konstantinidis, K.T. and Tiedje, J.M. (2005) Towards a genome-based taxonomy for
prokaryotes. *J Bacteriol* **187**: 6258–6264.
- 1075 Konstantinidis, K.T., Rosselló-Móra, R., and Amann, R. (2017) Uncultivated microbes in need
of their own taxonomy. *ISME J* **11**: 2399–2406.
- Kozłowski, J.A., Stieglmeier, M., Schleper, C., Klotz, M.G., and Stein, L.Y. (2016) Pathways
and key intermediates required for obligate aerobic ammonia-dependent chemolithotrophy
in bacteria and Thaumarchaeota. *ISME J* **10**: 1836–1845.
- 1080 Lam, P., Jensen, M.M., Kock, A., Lettmann, K.A., Plancherel, Y., Lavik, G., *et al.* (2011)
Origin and fate of the secondary nitrite maximum in the Arabian Sea. *Biogeosciences* **8**:
1565–1577.
- Lartillot, N. and Philippe, H. (2004) A Bayesian mixture model for across-site heterogeneities
in the amino-acid replacement process. *Mol BiolEvol* **21**: 1095–1109.
- 1085 Lechauve, C., Jager, M., Laguerre, L., Kiger, L., Correc, G., Leroux, C., *et al.* (2013)
Neuroglobins, pivotal proteins associated with emerging neural systems and precursors of
metazoan globin diversity. *J Biol Chem* **288**: 6957–6967.
- Lehtovirta-Morley, L.E., Stoecker, K., Vilcinskas, A., Prosser, J.I., and Nicol, G.W. (2011)
Cultivation of an obligate acidophilic ammonia oxidizer from a nitrifying acid soil. *Proc*
1090 *Natl Acad Sci USA* **108**: 15892–15897.
- Li, M., Baker, B.J., Anantharaman, K., Jain, S., Breier, J.A., and Dick, G.J. (2015) Genomic
and transcriptomic evidence for scavenging of diverse organic compounds by widespread
deep-sea archaea. *Nat Commun* **6**: 8933.

- 1095 Li, Y., Ding, K., Wen, X., Zhang, B., Shen, B., and Yang, Y. (2016) A novel ammonia-oxidizing archaeon from wastewater treatment plant: Its enrichment, physiological and genomic characteristics. *Sci Rep-UK* **6**: 23747.
- Li, P.-N., Herrmann, J., Tolar, B.B., Poitevin, F., Ramdasi, R., Bargar, J.R., *et al.* (2018) Nutrient transport suggests an evolutionary basis for charged archaeal surface layer proteins. *ISME J* **12**: 2389–2402.
- 1100 Lin, X., Handley, K.M., Gilbert, J.A., and Kostka, J.E. (2015) Metabolic potential of fatty acid oxidation and anaerobic respiration by abundant members of Thaumarchaeota and Thermoplasmata in deep anoxic peat. *ISME J* **9**: 2740–2744.
- Liu, M., Fan, L., Zhong, L., Kjelleberg, S., and Thomas, T. (2012) Metaproteogenomic analysis of a community of sponge symbionts. *ISME J* **6**: 1515–1525.
- 1105 Lloyd, K.G., Schreiber, L., Petersen, D.G., Kjeldsen, K.U., Lever, M.A., Steen, A.D., *et al.* (2013) Predominant archaea in marine sediments degrade detrital proteins. *Nature* **496**: 215–218.
- Love, G.D., Grosjean, E., Stalvies, C., Fike, D.A., Grotzinger, J.P., Bradley, A.S., *et al.* (2009) Fossil steroids record the appearance of Demospongiae during the Cryogenian period. *Nature* **457**: 718–721.
- 1110 Luo, C., Rodriguez-R, L.M., and Konstantinidis, K.T. (2014) MyTaxa: an advanced taxonomic classifier for genomic and metagenomic sequences. *Nucleic Acids Res* **42**: e73–e73.
- Lurie-Weinberger, M.N., Gomez-Valero, L., Merault, N., Glöckner, G., Buchrieser, C., and Gophna, U. (2010) The origins of eukaryotic-like proteins in *Legionella pneumophila*. *Int J Med Microbiol* **300**: 470–481.
- 1115 Luter, H.M., Whalan, S., and Webster, N.S. (2010) Exploring the role of microorganisms in the disease-like syndrome affecting the sponge *Ianthella basta*. *Appl Environ Microbiol* **76**: 5736–5744.
- Luter, H.M., Whalan, S., and Webster, N.S. (2012) Thermal and sedimentation stress are unlikely causes of brown spot syndrome in the coral reef sponge, *Ianthella basta*. *PLoS ONE* **7**: e39779.
- 1120 Maldonado, M., Ribes, M., and van Duyl, F.C. (2012) Nutrient fluxes through sponges. In: *Adv Mar Biol*. Elsevier, pp. 113–182.
- Martens-Habbena, W., Berube, P.M., Urakawa, H., de la Torre, J.R., and Stahl, D.A. (2009) Ammonia oxidation kinetics determine niche separation of nitrifying Archaea and Bacteria. *Nature* **461**: 976–979.
- 1125 Martens-Habbena, W., Qin, W., Horak, R.E.A., Urakawa, H., Schauer, A.J., Moffett, J.W., *et al.* (2015) The production of nitric oxide by marine ammonia-oxidizing archaea and

- 1130 inhibition of archaeal ammonia oxidation by a nitric oxide scavenger: Marine AOA
produce NO and are inhibited by NO scavenger. *Environ Microbiol* **17**: 2261–2274.
- Mehrshad, M., Amoozegar, M.A., Ghai, R., Shahzadeh Fazeli, S.A., and Rodriguez-Valera, F.
(2016) Genome reconstruction from metagenomic data sets reveals novel microbes in the
brackish waters of the Caspian Sea. *Appl Environ Microbiol* **82**: 1599–1612.
- 1135 Miller, W.T. (2012) Tyrosine kinase signaling and the emergence of multicellularity. *BBA-Mol
Cell Res* **1823**: 1053–1057.
- Mohamed, N.M., Saito, K., Tal, Y., and Hill, R.T. (2010) Diversity of aerobic and anaerobic
ammonia-oxidizing bacteria in marine sponges. *ISME J* **4**: 38–48.
- Moitinho-Silva, L., Díez-Vives, C., Batani, G., Esteves, A.I., Jahn, M.T., and Thomas, T.
(2017) Integrated metabolism in sponge–microbe symbiosis revealed by genome-centered
1140 metatranscriptomics. *ISME J* **11**: 1651–1666.
- Moitinho-Silva, L., Nielsen, S., Amir, A., Gonzalez, A., Ackermann, G.L., Cerrano, C., *et al.*
(2017) The sponge microbiome project. *GigaScience* **6**: 1–7.
- Morley, S.A., Berman, J., Barnes, D.K.A., de Juan Carbonell, C., Downey, R.V., and Peck,
L.S. (2016) Extreme phenotypic plasticity in metabolic physiology of Antarctic
1145 demosponges. *Front Ecol Evol* **3**: 157.
- Mosier, A.C., Allen, E.E., Kim, M., Ferriera, S., and Francis, C.A. (2012) Genome sequence of
‘*Candidatus Nitrosopumilus salaria*’ BD31, an ammonia-oxidizing archaeon from the San
Francisco Bay Estuary. *J Bacteriol* **194**: 2121–2122.
- 1150 Müller, W.E.G., Ushijima, H., Batel, R., Krasko, A., Borejko, A., Müller, I.M., and Schröder,
H.-C. (2006) Novel mechanism for the radiation-induced bystander effect: nitric oxide and
ethylene determine the response in sponge cells. *Mutat Res* **597**: 62–72.
- Murphy, D., Recous, S., Stockdale, E., Fillery, I.R., Jensen, L., Hatch, D., and Goulding, K.W.
(2003) Gross nitrogen fluxes in soil : theory, measurement and application of ¹⁵N pool
dilution techniques. In: *Adv Agron*. Elsevier, pp. 69–118.
- 1155 Mußmann, M., Brito, I., Pitcher, A., Sinninghe Damste, J.S., Hatzepichler, R., Richter, A., *et
al.* (2011) Thaumarchaeotes abundant in refinery nitrifying sludges express *amoA* but are
not obligate autotrophic ammonia oxidizers. *Proc Natl Acad Sci USA* **108**: 16771–16776.
- Nakagawa, T. and Stahl, D.A. (2013) Transcriptional response of the archaeal ammonia
oxidizer *Nitrosopumilus maritimus* to low and environmentally relevant ammonia
1160 concentrations. *Appl Environ Microbiol* **79**: 6911–6916.
- Off, S., Alawi, M., and Spieck, E. (2010) Enrichment and physiological characterization of a
novel Nitrospira-like bacterium obtained from a marine sponge *Appl Environ Microbiol*
76: 4640–4646.

- 1165 Offre, P., Kerou, M., Spang, A., and Schleper, C. (2014) Variability of the transporter gene complement in ammonia-oxidizing archaea. *Trends Microbiol* **22**: 665–675.
- Olm, M.R., Brown, C.T., Brooks, B., and Banfield, J.F. (2017) dRep: a tool for fast and accurate genomic comparisons that enables improved genome recovery from metagenomes through de-replication. *The ISME Journal* **11**: 2864–2868.
- 1170 Otte, J., Mall, A., Schubert, D.M., Könneke, M., and Berg, I.A. (2015) Malonic semialdehyde reductase from the archaeon *Nitrosopumilus maritimus* is involved in the autotrophic 3-hydroxypropionate/4-hydroxybutyrate cycle. *Appl Environ Microbiol* **81**: 1700–1707.
- Otto, A., Bernhardt, J., Meyer, H., Schaffer, M., Herbst, F.-A., Siebourg, J., *et al.* (2010) Systems-wide temporal proteomic profiling in glucose-starved *Bacillus subtilis*. *Nat Commun* **1**: 137.
- 1175 Ouverney, C.C. and Fuhrman, J.A. (2000) Marine planktonic archaea take up amino acids. *Appl Environ Microbiol* **66**: 4829–4833.
- Palatinszky, M., Herbold, C., Jehmlich, N., Pogoda, M., Han, P., von Bergen, M., *et al.* (2015) Cyanate as an energy source for nitrifiers. *Nature* **524**: 105–108.
- 1180 Park, S.-J., Ghai, R., Martín-Cuadrado, A.-B., Rodríguez-Valera, F., Chung, W.-H., Kwon, K., *et al.* (2014) Genomes of two new ammonia-oxidizing archaea enriched from deep marine sediments. *PLoS ONE* **9**: e96449.
- Parks, D.H., Imelfort, M., Skennerton, C.T., Hugenholtz, P., and Tyson, G.W. (2015) CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* **25**: 1043–1055.
- 1185 Passmore, I.J., Nishikawa, K., Lilley, K.S., Bowden, S.D., Chung, J.C.S., and Welch, M. (2015) Mep72, a metzincin protease that is preferentially secreted by biofilms of *Pseudomonas aeruginosa*. *J Bacteriol* **197**: 762–773.
- 1190 Patterson, N.J., Günther, J., Gibson, A.J., Offord, V., Coffey, T.J., Splitter, G., *et al.* (2014) Two TIR-like domain containing proteins in a newly emerging zoonotic *Staphylococcus aureus* strain sequence type 398 are potential virulence factors by impacting on the host innate immune response. *Front Microbiol* **5**: 662.
- Peng, J., Elias, J.E., Thoreen, C.C., Licklider, L.J., and Gygi, S.P. (2003) Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J Proteome Res* **2**: 43–50.
- 1195 Pester, M., Schleper, C., and Wagner, M. (2011) The Thaumarchaeota: an emerging view of their phylogeny and ecophysiology. *Curr Opin Microbiol* **14**: 300–306.

- Phillips, G., Swairjo, M.A., Gaston, K.W., Bailly, M., Limbach, P.A., Iwata-Reuyl, D., and de Crécy-Lagard, V. (2012) Diversity of archaeosine synthesis in crenarchaeota. *ACS Chem Biol* **7**: 300–305.
- 1200 Picossi, S., Montesinos, M.L., Pernil, R., Lichtlé, C., Herrero, A., and Flores, E. (2005) ABC-type neutral amino acid permease N-I is required for optimal diazotrophic growth and is repressed in the heterocysts of *Anabaena* sp. strain PCC 7120. *Mol Microbiol* **57**: 1582–1592.
- 1205 Pruesse, E., Peplies, J., and Glöckner, F.O. (2012) SINA: Accurate high-throughput multiple sequence alignment of ribosomal RNA genes. *Bioinformatics* **28**: 1823–1829.
- Qin, W., Martens-Habbena, W., Kobelt, J.N., and Stahl, D.A. (2016) *Candidatus* Nitrosopumilaceae. In: Whitman, W.B., Rainey, F., Kämpfer, P., Trujillo, M., Chun, J., DeVos, P., et al. (eds), *Bergey's Manual of Systematics of Archaea and Bacteria*. John Wiley & Sons, Ltd, Chichester, UK, pp. 1–2.
- 1210 Qin, W., Meinhardt, K.A., Moffett, J.W., Devol, A.H., Virginia Armbrust, E., Ingalls, A.E., and Stahl, D.A. (2017) Influence of oxygen availability on the activities of ammonia-oxidizing archaea: Influence of oxygen availability. *Environ Microbiol Rep* **9**: 250–256.
- 1215 Qin, W., Amin, S.A., Martens-Habbena, W., Walker, C.B., Urakawa, H., Devol, A.H., et al. (2014) Marine ammonia-oxidizing archaeal isolates display obligate mixotrophy and wide ecotypic variation. *Proc Natl Acad Sci USA* **111**: 12504–12509.
- Radax, R., Hoffmann, F., Rapp, H.T., Leininger, S., and Schleper, C. (2012) Ammonia-oxidizing archaea as main drivers of nitrification in cold-water sponges: Archaeal nitrification in cold-water sponges. *Environ Microbiol* **14**: 909–923.
- 1220 Radax, R., Rattei, T., Lanzen, A., Bayer, C., Rapp, H.T., Urich, T., and Schleper, C. (2012) Metatranscriptomics of the marine sponge *Geodia barretti*: tackling phylogeny and function of its microbial community. *Environ Microbiol* **14**: 1308–1324.
- Reveillaud, J., Maignien, L., Eren, A.M., Huber, J.A., Apprill, A., Sogin, M.L., and Vanreusel, A. (2014) Host-specificity among abundant and rare taxa in the sponge microbiome. *ISME J* **8**: 1198–1209.
- 1225 Reynolds, D. and Thomas, T. (2016) Evolution and function of eukaryotic-like proteins from sponge symbionts. *Mol Ecol* **25**: 5242–5253.
- Riesgo, A., Farrar, N., Windsor, P.J., Giribet, G., and Leys, S.P. (2014) The analysis of eight transcriptomes from all poriferan classes reveals surprising genetic complexity in sponges. *Mol Biol Evol* **31**: 1102–1120.

- 1230 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**: 1242592–1242592.
- Sabina, J. and Söll, D. (2006) The RNA-binding PUA domain of archaeal tRNA-guanine
transglycosylase is not required for archaeosine formation. *J Biol Chem* **281**: 6993–7001.
- 1235 Santoro, A.E., Dupont, C.L., Richter, R.A., Craig, M.T., Carini, P., McIlvin, M.R., *et al.* (2015)
Genomic and proteomic characterization of “*Candidatus Nitrosopelagicus brevis*”: An
ammonia-oxidizing archaeon from the open ocean. *Proc Natl Acad Sci USA* **112**: 1173–
1178.
- Sauder, L.A., Engel, K., Lo, C.-C., Chain, P., and Neufeld, J.D. (2018) “*Candidatus*
1240 *Nitrosotenuis aquarius*,” an ammonia- oxidizing archaeon from a freshwater aquarium
biofilter. *Appl Environ Microbiol* **84**: e01430–e01418.
- Sauder, L.A., Albertsen, M., Engel, K., Schwarz, J., Nielsen, P.H., Wagner, M., and Neufeld,
J.D. (2017) Cultivation and characterization of *Candidatus Nitrosocosmicus exaquare*, an
ammonia-oxidizing archaeon from a municipal wastewater treatment system. *ISME J* **11**:
1245 1142–1157.
- Schaefer, S.C. and Hollibaugh, J.T. (2017) Temperature decouples ammonium and nitrite
oxidation in coastal waters. *Environ Sci Technol* **51**: 3157–3164.
- Schläppy, M.-L., Schöttner, S.I., Lavik, G., Kuypers, M.M.M., de Beer, D., and Hoffmann, F.
(2010) Evidence of nitrification and denitrification in high and low microbial abundance
1250 sponges. *Mar Biol* **157**: 593–602.
- Schmitt, S., Tsai, P., Bell, J., Fromont, J., Ilan, M., Lindquist, N., *et al.* (2012) Assessing the
complex sponge microbiota: core, variable and species-specific bacterial communities in
marine sponges. *ISME J* **6**: 564–576.
- Sharp, K.H., Eam, B., Faulkner, D.J., and Haygood, M.G. (2007) Vertical transmission of
1255 diverse microbes in the tropical sponge *Corticium* sp. *Appl Environ Microbiol* **73**: 622–
629.
- Shen, T., Stieglmeier, M., Dai, J., Urich, T., and Schleper, C. (2013) Responses of the
terrestrial ammonia-oxidizing archaeon *Ca. Nitrososphaera viennensis* and the ammonia-
oxidizing bacterium *Nitrosospira multiformis* to nitrification inhibitors. *FEMS Microbiol*
1260 *Let* **344**: 121–129.
- Shigeno-Nakazawa, Y., Kasai, T., Ki, S., Kostyanovskaya, E., Pawlak, J., Yamagishi, J., *et al.*
(2016) A pre-metazoan origin of the CRK gene family and co-opted signaling network. *Sci*
Rep-UK **6**: 3439.

- 1265 Simion, P., Philippe, H., Baurain, D., Jager, M., Richter, D.J., Di Franco, A., *et al.* (2017) A large and consistent phylogenomic dataset supports sponges as the sister group to all other animals. *Curr Biol* **27**: 958–967.
- Simister, R.L., Deines, P., Botté, E.S., Webster, N.S., and Taylor, M.W. (2012) Sponge-specific clusters revisited: a comprehensive phylogeny of sponge-associated microorganisms. *Environ Microbiol* **14**: 517–524.
- 1270 Sørensen, P. and Jensen, E.S. (1991) Sequential diffusion of ammonium and nitrate from soil extracts to a polytetrafluoroethylene trap for ¹⁵N determination. *Anal Chim Acta* **252**: 201–203.
- Southwell, M.W., Popp, B.N., and Martens, C.S. (2008) Nitrification controls on fluxes and isotopic composition of nitrate from Florida Keys sponges. *Mar Chem* **108**: 96–108.
- 1275 Spang, A., Poehlein, A., Offre, P., Zumbärgel, S., Haider, S., Rychlik, N., *et al.* (2012) The genome of the ammonia-oxidizing *Candidatus Nitrososphaera gargensis*: insights into metabolic versatility and environmental adaptations. *Environ Microbiol* **14**: 3122–3145.
- Stamatakis, A. (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**: 1312–1313.
- 1280 Steger, D., Ettinger-Epstein, P., Whalan, S., Hentschel, U., de Nys, R., Wagner, M., and Taylor, M.W. (2008) Diversity and mode of transmission of ammonia-oxidizing archaea in marine sponges. *Environ Microbiol* **10**: 1087–1094.
- Stoecker, K., Dorninger, C., Daims, H., and Wagner, M. (2010) Double labeling of oligonucleotide probes for fluorescence in situ hybridization (DOPE-FISH) improves signal intensity and increases rRNA accessibility. *Appl Environ Microbiol* **76**: 922–926.
- 1285 Tabares, P., Pimentel-Elardo, S.M., Schirmeister, T., Hünig, T., and Hentschel, U. (2011) Anti-protease and immunomodulatory activities of bacteria associated with Caribbean sponges. *Mar Biotechnol* **13**: 883–892.
- Taylor, M.W., Radax, R., Steger, D., and Wagner, M. (2007) Sponge-associated microorganisms: evolution, ecology, and biotechnological potential. *Microbiol Mol Biol Rev* **1**: 295–347.
- 1290 Teira, E., van Aken, H., Veth, C., and Herndl, G.J. (2006) Archaeal uptake of enantiomeric amino acids in the meso- and bathypelagic waters of the North Atlantic. *Limnol Oceanogr* **51**: 60–69.
- 1295 Thiaville, J.J., Kellner, S.M., Yuan, Y., Hutinet, G., Thiaville, P.C., Jumpathong, W., *et al.* (2016) Novel genomic island modifies DNA with 7-deazaguanine derivatives. *Proc Natl Acad Sci USA* **113**: E1452–E1459.

- 1300 Thomas, T., Moitinho-Silva, L., Lurgi, M., Björk, J.R., Easson, C., Astudillo-García, C., *et al.*
(2016) Diversity, structure and convergent evolution of the global sponge microbiome. *Nat Commun* **7**: 11870.
- Tian, R.-M., Sun, J., Cai, L., Zhang, W.-P., Zhou, G.-W., Qiu, J.-W., and Qian, P.-Y. (2016)
The deep-sea glass sponge *Lophophysema eversa* harbours potential symbionts responsible
for the nutrient conversions of carbon, nitrogen and sulfur. *Environ Microbiol* **18**: 2481–
2494.
- 1305 Tourna, M., Stieglmeier, M., Spang, A., Könneke, M., Schintlmeister, A., Urich, T., *et al.*
(2011) *Nitrososphaera viennensis*, an ammonia oxidizing archaeon from soil. *Proc Natl Acad Sci USA* **108**: 8420–8425.
- Trame, C.B., Chang, Y., Axelrod, H.L., Eberhardt, R.Y., Coggill, P., Punta, M., and Rawlings,
N.D. (2014) New mini- zincin structures provide a minimal scaffold for members of this
1310 metallopeptidase superfamily. *BMC Bioinformatics* **15**: 1.
- Ueda, N., Richards, G.S., Degnan, B.M., Kranz, A., Adamska, M., Croll, R.P., and Degnan,
S.M. (2016) An ancient role for nitric oxide in regulating the animal pelagobenthic life
cycle: evidence from a marine sponge. *Sci Rep-UK* **6**: 37546.
- Vacelet, J., Boury-Esnault, N., Fiala-Medioni, A., and Fisher, C.R. (1995) A methanotrophic
1315 carnivorous sponge. *Nature* **377**: 296–296.
- Valladares, A., Montesinos, M.L., Herrero, A., and Flores, E. (2002) An ABC-type, high-
affinity urea permease identified in cyanobacteria. *Mol Microbiol* **43**: 703–715.
- Vallenet, D., Engelen, S., Mornico, D., Cruveiller, S., Fleury, L., Lajus, A., *et al.* (2009)
MicroScope: a platform for microbial genome annotation and comparative genomics.
1320 *Database* **2009**: bap021.
- Varghese, N.J., Mukherjee, S., Ivanova, N., Konstantinidis, K.T., Mavrommatis, K., Kyrpides,
N.C., and Pati, A. (2015) Microbial species delineation using whole genome sequences.
Nucleic Acids Res **43**: 6761–6771.
- Ventura, M., Turrone, F., Motherway, M.O., MacSharry, J., and van Sinderen, D. (2012) Host–
1325 microbe interactions that facilitate gut colonization by commensal bifidobacteria. *Trends Microbiol* **20**: 467–476.
- Verberkmoes, N.C., Russell, A.L., Shah, M., Godzik, A., Rosenquist, M., Halfvarson, J., *et al.*
(2009) Shotgun metaproteomics of the human distal gut microbiota. *ISME J* **3**: 179–189.
- Wallner, G., Amann, R., and Beisker, W. (1993) Optimizing fluorescent in situ hybridization
1330 with rRNA-targeted oligonucleotide probes for flow cytometric identification of
microorganisms. *Cytometry* **14**: 136–143.

- Wanek, W., Mooshammer, M., Blöchl, A., Hanreich, A., and Richter, A. (2010) Determination of gross rates of amino acid production and immobilization in decomposing leaf litter by a novel ^{15}N isotope pool dilution technique. *Soil Biol Biochem* **42**: 1293–1302.
- 1335 Wang, L., Chen, S., Xu, T., Taghizadeh, K., Wishnok, J.S., Zhou, X., *et al.* (2007) Phosphorothioation of DNA in bacteria by *dnd* genes. *Nat Chem Biol* **3**: 709–710.
- Washburn, M.P., Wolters, D., and Yates, J.R. (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* **19**: 242–247.
- 1340 Weber, E.B., Lehtovirta-Morley, L.E., Prosser, J.I., and Gubry-Rangin, C. (2015) Ammonia oxidation is not required for growth of Group 1.1c soil Thaumarchaeota. *FEMS Microbiol Ecol* **91**: fiv001.
- Webster, N.S. and Thomas, T. (2016) The Sponge Hologenome. *MBio* **7**: e00135–e00116.
- Webster, N.S., Taylor, M.W., Behnam, F., Lückner, S., Rattei, T., Whalan, S., *et al.* (2009) 1345 Deep sequencing reveals exceptional diversity and modes of transmission for bacterial sponge symbionts. *Environ Microbiol* **12**: 2070-2082.
- Wilkesman, J. and Schröder, H.C. (2002) Heat-stable protease from the marine sponge *Geodia cydonium*. *Cell Mol Bio. (Noisy-le-grand)* **48**: 379–383.
- Wilkesman, J.G. and Schröder, H.C. (2007) Analysis of serine proteases from marine sponges 1350 by 2-D zymography. *Electrophoresis* **28**: 429–436.
- Wilkinson, C.R. (1983) Net Primary Productivity in Coral Reef Sponges. *Science* **219**: 410–412.
- Wilkinson, C.R. and Fay, P. (1979) Nitrogen fixation in coral reef sponges with symbiotic cyanobacteria. *Nature* **279**: 527–529.
- 1355 Wilson, M.C., Mori, T., Rückert, C., Uria, A.R., Helf, M.J., Takada, K., *et al.* (2014) An environmental bacterial taxon with a large and distinct metabolic repertoire. *Nature* **506**: 58–62.
- Wuchter, C., Abbas, B., Coolen, M.J.L., Herfort, L., van Bleijswijk, J., Timmers, P., *et al.* (2006) Archaeal nitrification in the ocean. *Proc Natl Acad Sci USA* **103**: 12317–12322.
- 1360 Yan, J., Haaijer, S.C.M., Op den Camp, H.J.M., van Niftrik, L., Stahl, D.A., Könneke, M., *et al.* (2012) Mimicking the oxygen minimum zones: stimulating interaction of aerobic archaeal and anaerobic bacterial ammonia oxidizers in a laboratory-scale model system. *Environ Microbiol* **14**: 3146–3158.

- 1365 You, D., Wang, L., Yao, F., Zhou, X., and Deng, Z. (2007) A novel DNA modification by
sulfur: DndA is a NifS- like cysteine desulfurase capable of assembling DndC as an iron-
sulfur cluster protein in *Streptomyces lividans*. *Biochemistry* **46**: 6126–6133.
- 1370 Zhalnina, K.V., Dias, R., Leonard, M.T., Dorr de Quadros, P., Camargo, F.A.O., Drew, J.C., *et al.* (2014) Genome sequence of *Candidatus Nitrososphaera evergladensis* from group I.1b
enriched from Everglades soil reveals novel genomic features of the ammonia-oxidizing
archaea. *PLoS ONE* **9**: e101648.
- Zaikova, E. (2014) Molecular mechanisms of host-symbiont recognition in a highly specific
sponge-archaeal symbiosis. *University of British Columbia*.
- Zhou, J., Bruns, M.A., and Tiedje, J.M. (1996) DNA recovery from soils of diverse
composition. *Appl Environ Microbiol* **62**: 316–322.

1375