

A Specific Mix of Generalists: Microbial Symbionts in

Mediterranean *Ircinia* spp.

Patrick M. Erwin^{1*}, Susanna López-Legentil², Raúl González-Pech³, Xavier Turon¹

¹Center for Advanced Studies of Blanes (CEAB-CSIC), Accés Cala S. Francesc 14, 17300

Blanes, Girona, Spain

²Department of Animal Biology, University of Barcelona, Diagonal Avenue 643, 08028

Barcelona, Spain

³Universidad Autónoma de Nuevo León, Avenida Universidad s/n, Cd. Universitaria, San

Nicolás de los Garza, Nuevo León, México

*Corresponding author:

Phone: +34 972 336101

Fax: +34 972 337806

Email: patrick@ceab.csic.es

Running Title: *Microbial Symbionts in Mediterranean Ircinia spp.*

Keywords: Sponge, Bacterial Symbionts, Host-Specificity, Phylogenetics, ITS-2, COI

22 ABSTRACT

23 Microbial symbionts form abundant and diverse components of marine sponge
24 holobionts, yet the ecological and evolutionary factors that dictate their community
25 structure are unresolved. Here, we characterized the microbial symbiont communities of
26 3 sympatric host species in the genus *Ircinia* from the NW Mediterranean Sea, using
27 electron microscopy and replicated 16S rRNA gene sequence clone libraries. All *Ircinia*
28 host species harbored abundant and phylogenetically diverse symbiont consortia,
29 comprised primarily of sequences related to other sponge-derived microbes. Community-
30 level analyses of microbial symbionts revealed host species-specific genetic
31 differentiation and structuring of *Ircinia*-associated microbiota. Phylogenetic analyses of
32 host sponges showed a close evolutionary relationship between *I. fasciculata* and *I.*
33 *variabilis*, the 2 host species exhibiting the most similar symbiont communities. In
34 addition, several symbiont OTUs were exclusive to *I. variabilis* and *I. oros*, the 2 host
35 species inhabiting semi-sciophilous communities in more cryptic benthic habitats. The
36 generalist nature of individual symbionts and host-specific structure of entire
37 communities suggest that: 1) a “specific mix of generalists” framework applies to
38 microbial symbionts in *Ircinia* hosts, and 2) factors specific to each host species
39 contribute to the distinct symbiont mix observed in *Ircinia* hosts, including habitat-
40 specific conditions (e.g., irradiance and competition) and host evolutionary relatedness.

41

42 INTRODUCTION

43 Sponges are sessile, filter-feeding invertebrates that inhabit diverse marine
44 ecosystems and host remarkable microbial symbiont populations (Taylor et al., 2007;

1
2
3 45 Webster & Taylor, 2011), in some hosts accounting for up to 35% of sponge biomass
4
5
6 46 (Vacelet et al., 1975) and consisting of hundreds to thousands of symbiont taxa (Webster
7
8 47 et al., 2010; Lee et al., 2011). These diverse symbiont communities may enhance sponge
9
10 48 holobiont metabolism through non-metazoan processes, including photosynthesis (Erwin
11
12 49 & Thacker, 2008), nitrification (López-Legentil et al., 2010b) and sulfate reduction
13
14 50 (Hoffmann et al., 2005), and can produce defensive secondary metabolites (Flatt et al.,
15
16 51 2005) that decrease the susceptibility of host sponges to predation and fouling (Paul &
17
18 52 Ritson-Williams, 2008). In turn, sponge-associated microbes may benefit from the unique
19
20 53 microenvironment within host tissues, potentially nourished by the ammonia-rich end
21
22 54 products of animal metabolism and protected from open-ocean grazing pressures.
23
24
25
26
27 55 Although empirical evidence for symbiont benefit is scarce (Taylor et al., 2007), the high
28
29 56 biodiversity of sponge-associated microbes and their exclusivity to host sponges suggest
30
31 57 these niche habitats are fertile grounds for the evolutionary diversification of marine
32
33 58 microorganisms. To date, few host sponges have been investigated for microbial
34
35 59 symbionts and future research in sponge microbiology will further reveal the impact of
36
37 60 these symbioses on the ecological success of host sponges and the global biodiversity of
38
39 61 marine microorganisms (Taylor et al., 2004).
40
41
42

43 62 The study of sponge microbiology has benefited greatly from the application of
44
45 63 modern DNA sequencing technology, allowing for greater access to elusive symbiont
46
47 64 communities via culture-independent characterization and revealing striking trends in the
48
49 65 distribution and specificity of microbial symbionts. Perhaps most prominent is molecular
50
51 66 evidence for sponge-specific microbial lineages, distinct clusters of sponge-derived
52
53 67 sequences that represent major components of sponge-associated microbial communities
54
55
56
57
58
59
60

1
2
3 68 but are absent, or found in extremely low abundances (i.e., rare biosphere; Webster et al.,
4
5 69 2010; Lee et al., 2011), in ambient bacterioplankton (Taylor et al., 2007). While specific
6
7 70 to host sponge microenvironments, these symbiont lineages often exhibit a generalist
8
9 71 distribution among sponge hosts, occurring in taxonomically distinct hosts from distant
10
11 72 geographic regions (Olson & McCarthy, 2005; Hill et al., 2006; Sipkema et al., 2009) and
12
13 73 prompting early hypotheses on uniform microbial communities in marine sponges
14
15 74 (Hentschel et al., 2002). Typically, sponge-specific clades are common in high-microbial-
16
17 75 abundance (HMA) sponge hosts and absent in low-microbial-abundance (LMA) sponge
18
19 76 hosts (Hentschel et al., 2006; Weisz et al., 2007; Schmitt et al., 2011), where symbiont
20
21 77 communities are less diverse and contain more specialized (host species- or genus-
22
23 78 specific) microbes (Gernert et al., 2005; Holmes & Blanch, 2007; Schmitt et al., 2008).
24
25 79 An additional benefit of sequence-based approaches to characterize the sponge
26
27 80 microbiota is the accessibility of individual datasets and cumulative databases for
28
29 81 comparative studies and meta-analyses. A comprehensive review of sponge-derived 16S
30
31 82 rRNA gene sequences revealed that nearly one-third (32%) of all symbiont sequences
32
33 83 were associated with sponge-specific phylogenetic clades (Taylor et al., 2007).
34
35
36
37
38
39
40

41 From a host perspective, the composition and structure of microbial symbiont
42
43 85 communities have been reported as species-specific, despite the presence of sponge-
44
45 86 specific clusters within these communities (Webster et al., 2010; Lee et al., 2011). The
46
47 87 apparent conflict of distinct, host-specific symbiont communities that consist of shared,
48
49 88 sponge-specific symbiont lineages is explained by: 1) the selective presence (and relative
50
51 89 abundance) of putatively cosmopolitan symbionts among host sponges, and 2) the high
52
53 90 genetic diversity within sponge-specific sequence clusters. Indeed, the widespread
54
55
56
57
58
59
60

1
2
3 91 distribution of many sponge symbionts does not imply these symbionts are a ubiquitous
4
5 92 component of all HMA sponge hosts. For example, symbionts from the sponge-specific
6
7 93 cluster SC8 (*Chloroflexi*) exhibit an inter-ocean distribution, recovered from a Great
8
9 94 Barrier Reef sponge (*Rhopaloeides odorabile*; Webster et al., 2001b) and two Bahamian
10
11 95 sponges (*Agelas digitata* and *Plakortis* sp.; Taylor et al., 2007), and yet are absent in the
12
13 96 HMA sponge *Ircinia ramosa* collected at the same depth and habitat (within 100 m²) as
14
15 97 *R. odorabile* (Webster et al., 2010). Further, the definition of a sponge-specific symbiont
16
17 98 lineage does not require a sequence identity minimum to delineate sequence clusters
18
19 99 (Hentschel et al., 2002). Thus, the high genetic diversity exhibited within these microbial
20
21 100 sequence clusters, in some cases exceeding 20% sequence divergence, can encompass
22
23 101 multiple microbial OTUs from the species to the family level.
24
25
26
27
28

29 102 The factors that drive the observed differences in symbiont communities among
30
31 103 host sponge species are not fully resolved, with previous reports suggesting a role for
32
33 104 both environmental and host-specific factors. Multiple environmental factors have been
34
35 105 hypothesized to influence microbial symbiont composition in marine sponges, including
36
37 106 abiotic (temperature, nutrient levels, heavy metals) and biotic (competition, predation,
38
39 107 disease) factors (Webster et al., 2001a; Mohamed et al., 2008b; Webster et al., 2008a;
40
41 108 Webster et al., 2008b; Anderson et al., 2010; Turque et al., 2010; Angermeier et al.,
42
43 109 2011; Webster et al., 2011); however, several studies have shown that different host
44
45 110 species from the same habitat harbor distinct symbiont communities (Lee et al., 2009b;
46
47 111 Radwan et al., 2010), including the first 2 studies applying pyrosequencing technology to
48
49 112 the characterization of sponge-associated microorganisms (Webster et al., 2010; Lee et
50
51 113 al., 2011). Together, these results indicate that environmental conditions may influence
52
53
54
55
56
57
58
59
60

1
2
3 114 symbiont community structure in some host sponges, yet environmental factors alone are
4
5 115 not sufficient to fully explain the observed patterns of symbiont community structure.
6
7
8 116 Host-specific factors may also affect the composition of microbial symbiont in marine
9
10 117 sponges, although few studies have targeted congeneric hosts from the same habitat (Lee
11
12 118 et al., 2009b) or utilized molecular techniques to resolve host sponge phylogenies
13
14
15 119 (Erpenbeck et al., 2002; Thacker & Starnes, 2003; Thacker et al. 2007; Sipkema et al.,
16
17 120 2009).

18
19
20 121 To investigate the structuring and specificity of microbial symbionts in closely
21
22 122 related HMA host sponges, we studied 3 sympatric species in the genus *Ircinia* from the
23
24 123 Mediterranean Sea. The genus *Ircinia* is a chemically diverse and symbiont rich sponge
25
26 124 taxon that exhibits high species richness, occurring in shallow to deep-water habitats of
27
28 125 tropical and temperate marine environments. The chemical diversity of *Ircinia* spp.
29
30 126 includes diverse terpenoid compounds, most commonly sesterterpenoids (Cafieri et al.,
31
32 127 1972; Liu et al., 2006). These secondary metabolites have been shown to exhibit both
33
34 128 ecological functioning (e.g., anti-predatory activity; Pawlik et al., 2002) and biological
35
36 129 activity (e.g., anti-tumor cytotoxicity; Choi et al., 2004). The microbial diversity of
37
38 130 *Ircinia* spp. is consistent with other HMA host sponges, composed largely of sponge-
39
40 131 specific sequences from *Acidobacteria*, *Actinobacteria*, *Chloroflexi*, *Nitrospira*,
41
42 132 *Poribacteria* and *Proteobacteria* (Schmitt et al., 2007; Schmitt et al., 2008; Mohamed et
43
44 133 al., 2008a, Mohamed et al., 2008b, Mohamed et al., 2010; Webster et al., 2010; Yang et
45
46 134 al., 2011). These studies have focused on the microbiota in Caribbean and Indo-Pacific
47
48 135 host species, whereas the molecular diversity of microbial symbionts in Mediterranean
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 136 *Ircinia* hosts has been addressed by a single study that focused specifically on
4
5
6 137 cyanobacteria (Usher et al., 2004).

7
8 138 In this study, we characterized the microbial communities in the common
9
10 139 Mediterranean species, *Ircinia fasciculata*, *I. variabilis* and *I. oros*, using 16S rRNA gene
11
12 140 sequence clone libraries and compared the richness, diversity, structure and specificity of
13
14 141 symbiont communities among these congeneric and sympatric hosts. Phylogenetic
15
16 142 analyses were conducted to compare symbionts in Mediterranean *Ircinia* spp. with
17
18 143 previously described sponge-associated microbes, including sequences derived from
19
20 144 Caribbean *Ircinia* spp. In addition, we resolved the phylogenetic relationships among the
21
22 145 3 host sponges using ribosomal and mitochondrial DNA markers, thus allowing for the
23
24 146 determination of symbiont specificity within a well-defined host phylogenetic context.
25
26
27
28
29
30
31

32 148 **MATERIAL AND METHODS**

33 34 149 Sample Collection

35
36 150 The marine sponges *Ircinia fasciculata* (PALLAS 1766; Fig. 1a), *I. variabilis*
37
38 151 (SCHMIDT, 1862; Fig. 1b) and *I. oros* (SCHMIDT, 1864; Fig. 1c) were collected from
39
40 152 shallow (< 20 m) littoral zones at 3 neighboring sites (< 12 km apart) along the Catalan
41
42 153 Coast (Spain) in the northwestern Mediterranean Sea. *I. fasciculata* colonies were
43
44 154 sampled at Punta de S'Agulla (Blanes; 41° 40' 54.87" N, 2° 49' 00.01" E), *I. variabilis* at
45
46 155 Mar Menuda (Tossa de Mar; 41° 43' 13.62" N, 2° 56' 26.90" E) and *I. oros* at Punta
47
48 156 Santa Anna (Blanes; 41° 40' 21.48" N, 2° 48' 13.55" E) by SCUBA during 3 consecutive
49
50 157 days in March 2010. At each site, ambient seawater samples (500 ml) were collected
51
52
53 158 simultaneously and in close proximity (< 1 m) to sampled sponges. Sponge and seawater
54
55
56
57
58
59
60

1
2
3 159 samples were transported in an insulated cooler to the laboratory (ca. 2 hrs. transit time),
4
5
6 160 where sponge samples were preserved in 100% ethanol and stored at -20 °C and seawater
7
8 161 samples were concentrated on 0.2 µm filters and stored at -80 °C.
9

10
11 162

12 13 163 Transmission Electron Microscopy (TEM)

14
15 164 To visualize the microbial diversity present in *I. variabilis*, *I. fasciculata* and *I.*
16
17 165 *oros*, small ectosomal and choanosomal tissue pieces (ca. 4 mm³) were dissected and
18
19 166 fixed in a solution of 2.5% glutaraldehyde and 2% paraformaldehyde, buffered with
20
21 167 filtered seawater. Samples were incubated in the fixative mixture overnight at 4°C and
22
23 168 subsequently rinsed with filtered seawater to remove fixative, then dehydrated in a
24
25 169 graded ethanol series and embedded in Spurr resin at room temperature. A Reichert
26
27 170 Ultracut microtome was used to produce ultra-thin sections (ca. 60 nm) that were
28
29 171 contrasted with uranyl acetate and lead citrate for ultrastructural observation (Reynolds,
30
31 172 1963). TEM observations were performed at the Microscopy Unit of the Scientific and
32
33 173 Technical Services of the University of Barcelona on a JEOL JEM-1010 (Tokyo, Japan)
34
35 174 coupled with a Bioscan 972 camera (Gatan, Germany).
36
37
38
39
40
41
42

43
44 175

45 176 DNA Extraction

46 177 Metagenomic DNA extracts were prepared separately for 3 individuals of each
47
48 178 host sponge species (including ectosomal and choanosomal tissue) and 3 samples of
49
50 179 concentrated seawater (1 from each collection site) using the DNeasy® Blood & Tissue
51
52 180 Kit (Qiagen®). Full-strength and 1:10 diluted DNA extracts were used as templates in
53
54 181 PCR amplifications.
55
56
57
58
59
60

1
2
3 1824
5
6 183 16S rRNA Gene Sequence Clone Libraries

7
8 184 The universal bacterial forward primer 8F (5'-AGA GTT TGA TCA TGG CTC
9
10 185 AG-3')(Reysenbach et al., 1994) and reverse primer 1509R (5'-GGT TAC CTT GTT
11
12 186 ACG ACT T-3')(Martínez-Murcia et al., 1995) were used to amplify approximately
13
14 187 1,500 bp fragments of bacterial 16S rRNA gene sequences from all sponge and seawater
15
16 188 extracts. Total PCR reaction volume was 50 µl, including 10 pmol of each primer, 10
17
18 189 nmol of each dNTP, 1X Reaction Buffer (Ecogen) and 5 units of BIOTAQ™ polymerase
19
20 190 (Ecogen). Thermocycler reaction conditions were an initial denaturing time of 2 min at
21
22 191 94°C, followed by 30 cycles of 1 min at 94°C, 0.5 min at 50°C, and 1.5 min at 72°C, and a
23
24 192 final extension time of 2 min at 72°C. To minimize PCR amplification biases, a low
25
26 193 annealing temperature and low cycle number were used and 3 separate reactions were
27
28 194 conducted for each sample. PCR amplification products were gel-purified and cleaned
29
30 195 using the QIAquick Gel Extraction Kit (Qiagen®). Triplicate PCR products were
31
32 196 combined and quantified using a Qubit™ fluorometer and Quant-iT™ dsDNA Assay Kit
33
34 197 (Invitrogen™). Purified PCR products (25 ng) were ligated into plasmids using the
35
36 198 pGEM®-T Vector System (Promega).

37
38 199 Individual clones were PCR-screened using vector primers and clones with
39
40 200 approximately 1,500 bp inserts were purified and sequenced at Macrogen Inc. A single
41
42 201 sequencing reaction was performed for all clones to recover the 5'-end of 16S rRNA gene
43
44 202 sequences using the primer 800R (5'-TAC CAG GGT ATC TAA TCC-3'). Ambiguities
45
46 203 on sequencing reaction ends were excluded by trimming sequences at the 5'-end to the
47
48 204 highly conserved *E. coli* site 54 and at the 3'-end to *E. coli* site 754, yielding sequences
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 205 ranging from 613 to 725 bp (average length = 683 bp) that were used for diversity
4
5
6 206 calculations and phylogenetic metrics. In addition, bi-directional sequencing reactions
7
8 207 with vector primers were performed to recover near full-length 16S rRNA gene
9
10 208 sequences (range = 1,423 to 1,523 bp; average = 1,491 bp) of representative clones ($n =$
11
12 209 39) for phylogenetic analyses. Representative clones consisted of bacterial OTUs that
13
14 210 were either: 1) 'common' – occurring at least twice in clone libraries, or 2) 'unique' –
15
16 211 exhibiting greater than 5% sequence divergence from any known sequence. Raw
17
18 212 sequence data were processed in Geneious (Drummond et al., 2010) and low quality
19
20 213 sequencing reads were discarded. Sequences were screened for sequencing anomalies
21
22 214 (e.g., chimeras) using Mallard (Ashelford et al., 2006) and a reference 16S rRNA gene
23
24 215 sequence from *E. coli* (GenBank accession no. U00096). Putative sequence anomalies
25
26 216 were subsequently confirmed or refuted using Pintail (Ashelford et al., 2005) and two
27
28 217 related reference 16S rRNA gene sequences. All confirmed chimeras were removed from
29
30 218 the dataset. High quality sequences are archived in GenBank under accession nos.
31
32 219 JN655200-JN655511.
33
34
35
36
37
38
39
40

41 Operational Taxonomic Unit (OTU) Assignment and Composition

42
43 222 Clone library sequences were ascribed to OTUs calculated at different sequence
44
45 223 identity percentages (99%, 97%, 95%, 90%, 85% and 80%) using the nearest neighbor
46
47 224 algorithm, as implemented in the mothur software package (Schloss et al., 2009). The
48
49 225 observed OTU richness (S_{obs}) for each microbial community was compared across
50
51 226 different OTU thresholds, calculated as total (combined sequence data by source) and
52
53 227 average values (separated sequence data by samples). All subsequent OTU-based metrics
54
55
56
57
58
59
60

1
2
3 228 were conducted using an OTU classification at 99% sequence identity. Representative
4
5 229 sequences from each 99% OTU were analyzed by using a nucleotide-nucleotide BLAST
6
7
8 230 search (Altschul et al., 1990) to find the most closely related sequence, and by using the
9
10 231 Ribosomal Database Project II (Cole et al., 2007) sequence classifier to assess taxonomic
11
12 232 affiliations.
13
14
15 233

17 234 Microbial Community Diversity

19
20 235 The diversity of recovered bacterial communities in *Ircinia* spp. and ambient
21
22 236 seawater were compared using multiple metrics for OTU richness, dominance and
23
24 237 evenness, calculated in the mothur software package. Richness calculations included
25
26 238 observed species richness (99% OTUs), rarefaction analysis and expected species
27
28 239 richness using the Chao1 estimator (S_{Chao1}) to determine both the sampled diversity and
29
30 240 total expected diversity of bacterial communities. In addition, the effect of increase
31
32 241 sequencing effort on OTU richness was estimated using the Boneh calculation (Boneh et
33
34 242 al., 1998). Dominance metrics included Simpson's inverse index ($1/D$) and the Berger-
35
36 243 Parker index (d), the former a reciprocal version of Simpson's heterogeneity index ($D =$
37
38 244 $\sum p_i^2$, where p_i is the proportion of individuals in species i) interpretable as the number of
39
40 245 equally common species that would produce the observed heterogeneity, and the latter a
41
42 246 simple metric that calculates the relative abundance of the most dominant OTU.
43
44 247 Evenness calculations included the Simpson's evenness measure ($E_{1/D}$), based on
45
46 248 Simpson's inverse index ($1/D$), and Smith & Wilson's evenness index (E_{VAR}). Both
47
48 249 indices provide an evenness measure independent of species richness that ranges 0 to 1;
49
50 250 however, $E_{1/D}$ places equal weight on both common and rare taxa, while with E_{VAR}
51
52
53
54
55
56
57
58
59
60

251 common species received higher weighting (i.e., greater influence on the evenness
252 measure) than rare species.

253 Genetic diversity of symbiont communities was compared among host species
254 and seawater sources using nonparametric tests for homogeneity of molecular variance
255 (HOMOVA) and an analysis of molecular variance (AMOVA; Stewart & Excoffier,
256 1996). These tests provide OTU-independent assessments of genetic variation and
257 differentiation within and among bacterial communities inhabiting *Ircinia* spp. and
258 seawater. HOMOVA tests whether the genetic variation observed within each population
259 (i.e., bacterial community) differs among sources (i.e., host sponge species or seawater),
260 thus providing a comparative measure of genetic diversity within each population.

261 AMOVA tests whether the genetic diversity observed within each population differs from
262 the total genetic diversity resulting from pooling the populations, thus providing a
263 comparative measure of genetic differentiation among populations. Distances were
264 calculated for AMOVA using the Tajima and Nei algorithm with $\alpha = 0.05$. Using the
265 Arlequin software package, version 3.5 (Excoffier & Lischer, 2010), a hierarchical
266 partitioning of genetic variation was assessed across different levels (among sources,
267 among replicates within sources and among sequences within each replicate) and
268 pairwise variation among sources was computed as F_{ST} , with statistical significance based
269 on 1000 permutations. Distributions of unique lineages among bacterial communities
270 were examined using a phylogenetic lineage-sorting test (P-test; Martin, 2002), also
271 referred to as the parsimony test (Schloss, 2008).

272

273 Microbial Community Structure & Similarity

1
2
3 274 To determine the distribution of bacterial OTUs within each community, OTU
4
5 275 rank-abundance plots were constructed and compared to a fitted log series and geometric
6
7
8 276 series distributions by calculating the Kolmogorov-Smirnov test statistic (D_{\max}). The
9
10 277 significance of D_{\max} statistics were determined by comparison to critical values calculated
11
12 278 at $\alpha=0.05$ and 0.01. Community similarity among sources was calculated as Bray-Curtis
13
14 279 similarity values and visualized in complete linkage similarity dendrograms using
15
16 280 PRIMER v6 (Plymouth Marine Laboraroty, UK) computer software. Finally, the integral
17
18 281 form of LIBSHUFF (\int -LIBSCHUFF; Schloss et al., 2004) was used to test pairwise
19
20 282 differences in microbial communities from each source. \int -LIBSHUFF analysis was
21
22 283 chosen because it represents a relatively ‘generic’ test of differences between microbial
23
24 284 communities, based on both membership and structure (Schloss, 2008). \int -LIBSHUFF
25
26 285 analyses were implemented in the mothur software package, with significance values
27
28 286 determined by 100,000 randomizations and adjusted using Bonferroni corrections to
29
30 287 maintain an experiment-wise critical value of 0.05 across multiple pairwise comparisons
31
32 288 (Sokal & Rohlf, 1995).

33 34 35 36 37 38 39 40 41 290 Phylogenetic analysis

42
43 291 Phylogenetic analyses of 16S rRNA gene sequences were conducted to determine
44
45 292 the affiliations between sequences recovered from *Ircinia* spp. herein and previously
46
47 293 characterized sponge symbionts. In particular, sequences from recent studies
48
49 294 characterizing sponge-associated bacterial communities in *Ircinia* species from the
50
51 295 Caribbean were targeted. Publicly available datasets for *I. felix* (Schmitt et al., 2007;
52
53 296 Schmitt et al., 2008) and *I. strobilina* (Mohamed et al., 2008b; Yang et al., 2011) were
54
55
56
57
58
59
60

1
2
3 297 retrieved from GenBank and grouped into 99% OTUs, following the methods employed
4
5
6 298 for the clone libraries herein. Representative sequences from each OTU for *I. felix* ($n =$
7
8 299 102) and *I. strobilina* ($n = 156$), top matching sequences from BLAST searches ($n =$
9
10 300 189), and near full-length 16S clones ($n = 39$) and partial 16S clones ($n = 87$) from this
11
12 301 study were aligned to the greengenes reference database (DeSantis et al., 2006) using the
13
14 302 mothur software package, with an outgroup sequence from Archaea (*Haloarcula*
15
16 303 *vallismortis*, GenBank accession no. D50581). Maximum likelihood phylogenetic trees
17
18 304 were constructed in RAxML (Stamatakis et al., 2005) using the General Time Reversible
19
20 305 model with a gamma distribution of variable substitution rates among sites (GTR+G).
21
22 306 Data were resampled using 100 bootstrap replicates and a thorough ML search was
23
24 307 conducted to optimize the topology and recover the best-scoring tree. Due to the variable
25
26 308 length of 16S rRNA gene sequences being compared (422 to 1,526 bp), a binary
27
28 309 backbone constraint tree was constructed from long (>1,000 bp) sequences and used to
29
30 310 restrict topology changes when introducing short (<1,000 bp) sequences into the
31
32 311 phylogeny. This method allowed for: 1) accurate reconstruction of deeper nodes, based
33
34 312 on the most informative sequences, and 2) precise placement of short 16S gene sequence
35
36 313 fragments near terminal nodes, for comparative analysis with previous *Ircinia* sp. datasets
37
38 314 (e.g., excised and sequenced DGGE bands; Schmitt et al., 2007).
39
40
41
42
43
44
45
46
47

316 Molecular Identification of Host Sponges

317 Sponge samples were identified to species using morphological observations,
318 including gross morphology and fiber characteristics. In addition, ribosomal and
319 mitochondrial molecular markers were used to objectively test morphological

1
2
3 320 characterizations, as some controversy surrounds the taxonomic status of *Ircinia*
4
5 321 *fasciculata* and *I. variabilis* (Pronzato et al., 2004) and these species exhibit high
6
7
8 322 phenotypic plasticity that can confound their identification in the field (Maldonado et al.,
9
10 323 2010). A segment of nuclear ribosomal DNA corresponding to the 3'-end of the 5.8S
11
12 324 subunit, the entire second internal transcribed spacer (ITS-2) region and the 5'-end of the
13
14 325 28S subunit were PCR-amplified following the methods of Erwin & Thacker (2007). A
15
16 326 segment of mitochondrial DNA corresponding to the cytochrome oxidase subunit I (COI)
17
18 327 was PCR-amplified using a degenerated version of the universal barcoding forward
19
20 328 primer dgLCO1490 (5'-GGT CAA CAA ATC ATA AAG AYA TYG G-3')(Meyer et al.,
21
22 329 2005) and the reverse primer COX1-R1 (5'-AAT ACT GCT TTT TTT GAT CCT GCC
23
24 330 GG-3'). This primer combination yielded a segment of mtDNA encompassing both the
25
26 331 standard barcoding fragment (i.e., "Folmer" partition)(Folmer et al., 1994, Herbert et al.,
27
28 332 2003) and the I3-M11 partition (Erpenbeck et al., 2006), which spans the internal loop 3
29
30 333 to the transmembrane domain 11. The resulting PCR amplification products were gel-
31
32 334 purified and cleaned using the QIAquick Gel Extraction Kit (Qiagen®) and ligated into
33
34 335 plasmids using the pGEM®-T Vector System (Promega). Individual clones were PCR-
35
36 336 screened using vector primers, purified and sequenced at Macrogen Inc. Bi-directional
37
38 337 sequencing reactions were performed for all clones using vector primers to recover the
39
40 338 entire cloned amplicons. Raw sequence data were processed in Geneious and low quality
41
42 339 sequencing reads were discarded. High quality sequences are archived in GenBank under
43
44 340 accession nos. JN655171-JN655199.

51
52
53 341 For COI fragments, 2 to 4 clones were recovered per sponge individual (5 to 9
54
55 342 clones per species) and used to construct consensus sequences. For rDNA fragments, 2 to
56
57
58
59
60

1
2
3 343 3 clones were recovered per sponge individual (6 to 8 clones per species) and consensus
4
5 344 sequences were processed separately for each clone, due to the potential for intragenomic
6
7
8 345 variation among the multiple ITS-2 copies in tandemly repeated rDNA clusters
9
10 346 (Wörheide et al., 2004). Consensus sequences for rDNA fragments were aligned using
11
12 347 MAFFT (Kato et al., 2002) with outgroup sequences from *Smenospongia aurea*
13
14
15 348 (Dictyoceratida; Thorectidae)(Erwin & Thacker, 2007). Alignment of consensus COI
16
17 349 sequences was unequivocal, due to the protein coding nature of these sequences, and
18
19 350 included the congeneric Caribbean species *I. strobilina* (GenBank accession no.
20
21 351 GQ337013; Erpenbeck et al. 2009) and the outgroup species *Hippospongia lachne*
22
23 352 (Dictyoceratida; Spongiidae)(EU237484; Lavrov et al., 2008). Pairwise genetic distance
24
25 353 matrices (uncorrected p-distance) were constructed using the software package mothur.
26
27 354 For rDNA fragments, maximum likelihood (ML) phylogenies were constructed using
28
29 355 PHYML (Guidon & Gascuel, 2003) and the Hasegawa-Kishino-Yano model with a
30
31 356 gamma distribution of variable substitution rates among sites (HKY+G), as suggested by
32
33 357 FINDMODEL (<http://www.hiv.lanl.gov/content/sequence/findmodel/findmodel.html>)
34
35 358 based on the Akaike information criterion; data were resampled using 100 bootstrap
36
37 359 replicates. Neighbor-Joining trees were constructed using Geneious and the HKY model
38
39 360 of nucleotide substitution; data were resampled using 1000 bootstrap replicates.
40
41
42
43
44
45
46
47

362 RESULTS

363 Transmission Electron Microscopy

364 Electron microscopy observations of host sponge tissue revealed a high density of
365 bacterial cells (Fig. 1). Characteristic of high-microbial-abundance (HMA) sponges,

1
2
3 366 examined sections were comprised primarily of bacterial cells, with only occasional
4
5 367 sponge cells (archeocytes) and structural elements (spongin and collagen fibers). In *I.*
6
7 368 *fasciculata* and *I. variabilis*, ectosomal (peripheral) tissue sections revealed dense
8
9 369 populations of “*Candidatus Synechococcus spongiarum*” identifiable by their
10
11 370 characteristic spiral thylakoid membranes encompassing the perimeter of the cells. *S.*
12
13 371 *spongiarum* cells dominated the ectosomal regions of host tissue and were observed to be
14
15 372 actively dividing, exhibiting several stages of binary fission (Fig. 1d,e). In *I. oros*, no *S.*
16
17 373 *spongiarum* symbionts or other cyanobacterial cells were observed in ectosomal tissue,
18
19 374 rather a high density of heterotrophic bacteria occurred consisting of multiple bacterial
20
21 375 morphotypes, some of which were also showing active cell division (Fig. 1g). Sections
22
23 376 from deeper tissue regions (choanosome) of *I. variabilis* revealed the absence of *S.*
24
25 377 *spongiarum* cells and the proliferation of heterotrophic bacteria cells, many with similar
26
27 378 morphotypes to those observed in *I. oros* ectosomal tissue (Fig. 1f).
28
29
30
31
32
33
34
35

36 Microbial Community OTUs and Coverage

37
38
39 381 Bacterial 16S rRNA gene sequences recovered from *I. fasciculata* ($n = 77$), *I.*
40
41 382 *variabilis* ($n = 80$), *I. oros* ($n = 82$) and ambient seawater ($n = 73$) grouped into 124
42
43 383 OTUs, defined by 99% or greater sequence similarity. Grouping sequences according to
44
45 384 lower identity thresholds reduced the number of OTUs within each source community;
46
47 385 however, the same trend in comparative OTU richness among communities was observed
48
49 386 (Fig. S1). Coverage estimates revealed that sampled sponge-associated bacterial
50
51 387 communities represented the majority of expected diversity (76.3% in *I. variabilis*, 80.5%
52
53 388 in *I. fasciculata*, 81.7% in *I. oros*) and doubling the sampling effort conducted herein
54
55
56
57
58
59
60

1
2
3 389 were predicted to produce few (5 to 6) new OTUs. In contrast, recovered seawater
4
5 390 bacterial communities represented less than half (43.8%) of the total expected diversity
6
7
8 391 and doubling the sampling effort was predicted to produce 12 new OTUs, indicating
9
10 392 more extensive sampling is required to fully characterize bacterioplankton diversity.
11
12 393 Similarly, rarefaction analyses showed the sponge-associated bacterial communities
13
14
15 394 beginning to reach OTU saturation, whereas seawater bacterial communities continued to
16
17 395 rapidly accumulate new OTUs (Fig. S2).
18
19

20 396

21 22 397 Microbial Community Composition

23
24 398 Sponge-associated microbial communities exhibited high diversity and were
25
26
27 399 comprised of 9 phyla, including representatives from 4 classes of *Proteobacteria* (*Alpha-*,
28
29 400 *Beta-*, *Gamma-* and *Deltaproteobacteria*; Table 1). Sequences affiliated with
30
31 401 *Proteobacteria* and *Bacteroidetes* were recovered from all 3 *Ircinia* host species and
32
33
34 402 seawater, whereas *Acidobacteria*, *Nitrospira*, *Chloroflexi* and *Gemmatimonadetes* were
35
36 403 exclusively found in *Ircinia*-associated communities. Within the sponge-associated
37
38 404 microbiota, *Deltaproteobacteria*-affiliated sequences comprised an abundant component
39
40 405 of microbial communities in all 3 host species (>15% total clones), although phylum-
41
42 406 level differentiation of symbiont communities among hosts was apparent, with a high
43
44 407 relative abundance of *Cyanobacteria* in *I. fasciculata*, *Nitrospira* in *I. variabilis* and
45
46 408 *Acidobacteria* in *I. oros* (Table 1). Ambient seawater microbes represented 10 phyla and
47
48 409 included 4 phyla not detected in sponge-associated communities (*Chlorophyta*,
49
50 410 *Firmicutes*, *Planctomycetes* and *Verrucomicrobia*; Table 1).
51
52
53
54
55
56
57
58
59
60

1
2
3 411 The vast majority of 16S rRNA gene sequences recovered from *I. fasciculata*
4
5 412 (87.0%), *I. variabilis* (85.0%) and *I. oros* (81.0%) matched most closely to other sponge-
6
7 413 derived bacterial sequences, generally at high sequence identity levels ($\geq 97\%$; Fig. 2).
8
9 414 Sequences from *Ircinia* spp. not associated with sponge-derived clones were most
10
11 415 commonly matched to sequences from marine sediment ($n = 19$), corals ($n = 4$) and
12
13 416 seawater ($n = 3$). Most sequences obtained from ambient seawater bacteria (98.6%) were
14
15 417 closely related to sequence from other bacterioplankton sources, with nearly all
16
17 418 sequences (90.3%) matching at very high sequence identity levels ($\geq 99\%$; Fig. 2). In
18
19 419 fact, only 3 sequences from seawater clone libraries did not exhibit a top match to another
20
21 420 bacterioplankton clone, with 2 matching to coral-associated bacteria and 1 to a bacterium
22
23 421 from a sulfate-reducing bioreactor.
24
25
26
27
28
29
30
31

32 Microbial Community Diversity

33
34 424 Seawater bacterioplankton communities were clearly differentiated from the
35
36 425 sponge-associated bacterial communities, exhibiting higher OTU richness (observed and
37
38 426 expected), lower dominance indices and higher evenness indices (Table 2). The most
39
40 427 dominant OTU is the seawater community accounted for 8.0% of the total community
41
42 428 and singleton OTUs ($n = 41$) accounted for over 80% of all OTUs. In fact, only 5 OTUs,
43
44 429 (9.4%) were recovered more than twice in seawater clone libraries.
45
46
47

48 430 Among the 3 host sponges, bacterial communities in *I. variabilis* and *I. oros*
49
50 431 exhibited very similar richness and evenness values (Table 2), suggesting a similar
51
52 432 diversity of symbionts among these host species. Dominant OTUs in *I. variabilis* and *I.*
53
54 433 *oros* accounted for 15% of each community, with the top 3 most abundant OTUs
55
56
57
58
59
60

1
2
3 434 accounting for one-third of the total community. Approximately half of the OTUs
4
5 435 recovered from *I. variabilis* ($n = 19$, 55.9%) and *I. oros* ($n = 15$, 45.5%) were singletons.
6
7
8 436 By comparison, the bacterial community in *I. fasciculata* exhibited lower OTU richness
9
10 437 and a less even (more dominant) community structure (Table 2), driven by a single OTU
11
12 438 (corresponding to the cyanobacterium *Synechococcus spongiarum*) that accounted for
13
14 439 over one-fourth of all recovered clones. The top 3 bacterial OTUs accounted for nearly
15
16 440 half (48.1%) of all recovered clones and singleton OTUs were common ($n = 15$, 51.7%).
17
18 441 Notably, although *I. fasciculata* hosted a comparatively less diverse community,
19
20 442 overlapping confidence intervals for index values were observed among all *Ircinia* hosts
21
22 443 (Table 2), indicating similar symbiont diversity across the 3 host species.
23
24
25
26

27 444 Genetic diversity analyses revealed similar trends to OTU-based metrics (Table
28
29 445 3). Microbial communities in seawater exhibited significantly higher genetic variation
30
31 446 compared to those in *Ircinia* sponges (HOMOVA, $P < 0.005$), indicating higher genetic
32
33 447 diversity in the bacterioplankton communities. No significant differences in genetic
34
35 448 variation among communities were observed in pairwise comparisons of host sponge
36
37 449 species (HOMOVA, $P > 0.237$), indicating similar levels of genetic diversity in *Ircinia*-
38
39 450 associated bacterial communities. AMOVA revealed significant genetic differentiation
40
41 451 among all 4 sources ($P < 0.001$), which accounted for the majority (62.1%) of genetic
42
43 452 variation, and among replicates within each source ($P < 0.001$), which accounted for the
44
45 453 remaining genetic variation (37.9%). Further, distinct phylogenetic lineages of symbionts
46
47 454 (P-test, $P < 0.001$) were observed among all pairwise comparisons between source
48
49 455 communities (Table 3).
50
51
52
53
54
55
56
57
58
59
60

1
2
3 457 Microbial Community Structure and Similarity
4

5
6 458 In addition to differences in diversity between bacterioplankton and sponge-
7
8 459 associated communities, minimal overlap in OTU composition was observed between
9
10 460 seawater and sponge-derived sequences. Of the 124 bacterial OTUs recovered, only 2
11
12 461 OTUs were shared between sponges and seawater (Fig. 3), both in *I. variabilis*, and
13
14 462 accounted for a small portion of the total *I. variabilis* (2.5%) and seawater (5.3%) clone
15
16 463 libraries. Comparisons among the sponge-associated bacterial communities revealed
17
18 464 varying levels of OTU overlap and host-specificity (Fig. 3). Four bacterial OTUs were
19
20 465 recovered from all 3 host sponge species and represented dominant symbionts,
21
22 466 accounting for nearly one-fourth of all clones from each host library (23.4% in *I.*
23
24 467 *fasciculata*, 27.5% in *I. variabilis*, and 23.2% in *I. oros*). An additional 7 OTUs were
25
26 468 shared between *I. fasciculata* and *I. variabilis*, accounting for an additional 40.3% and
27
28 469 31.3% of each library, respectively, and 6 OTUs were shared between *I. variabilis* and *I.*
29
30 470 *oros*, accounting for 15.0% and 32.9% of each library, respectively. No additional
31
32 471 microbial OTUs were shared between *I. fasciculata* and *I. oros*. Finally, the majority of
33
34 472 OTUs recovered for each host species' community consisted of specific symbionts (i.e.,
35
36 473 recovered from a single host species)($n = 18$ in *I. fasciculata*, $n = 15$ in *I. variabilis*, n
37
38 474 $=23$ in *I. oros*) and represented rare OTUs, commonly appearing only once (65.5% of
39
40 475 specific OTUs) or twice (25.5%) in clone libraries. Rank-abundance analyses showed
41
42 476 that the presence of few, dominant OTUs and numerous, rare OTUs were consistent with
43
44 477 a log-series distribution of OTUs within each clone library (D_{\max} , $P > 0.05$), and differed
45
46 478 significantly from a geometric distribution (D_{\max} , $P < 0.01$).
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 479 Consistent with patterns of symbiont OTU overlap among host sponges, overall
4
5
6 480 bacterial community similarity values were lowest between *I. fasciculata* and *I. oros*
7
8 481 (16.7%) and the symbionts in these two hosts differed significantly in community
9
10 482 structure (LIBSHUFF, $P < 0.001$; Table 3). The microbiota in *I. variabilis* exhibited
11
12 483 higher similarity to *I. fasciculata* symbiont communities (36.6%), where no significant
13
14 484 difference in symbiont structure were detected ($P > 0.639$; Table 3), than to microbial
15
16 485 symbionts in *I. oros* (31.1%), where significant differences in community structure were
17
18 486 detected ($P < 0.05$; Table 3).
19
20
21

22 487

23 24 488 Phylogenetic analysis

25
26
27 489 Phylogenetic analysis revealed that sequences recovered from the *I. fasciculata*, *I.*
28
29 490 *variabilis* and *I. oros* formed 56 monophyletic sequence clusters (Fig. 4). Nearly half
30
31 491 (48.2%) of these clusters were comprised exclusively of sponge-associated bacterial
32
33 492 sequences, with an additional 11 clusters (19.6%) consisting of sponge and coral-
34
35 493 associated clones. The remaining 18 clades (32.1%) contained non-symbiotic
36
37 494 representatives, most commonly derived from sediment ($n = 8$) and seawater
38
39 495 bacterioplankton ($n = 5$). Previously described bacterial sequences from Caribbean
40
41 496 *Ircinia* spp. were present in 21 of the 56 clades (37.5%), with *I. felix* from Florida ($n =$
42
43 497 16) and *I. strobilina* from the Bahamas ($n = 12$) more commonly presenting related
44
45 498 symbionts than *I. strobilina* from Florida ($n = 3$). Only 1 of the 56 clades (1.8%) was
46
47 499 comprised of sequences exclusively from *Ircinia* spp. (Fig. 4c). Notably, symbiont
48
49 500 sequences from the unrelated host sponge species *Aplysina aerophoba* (Order Verongida)
50
51 501 from the Mediterranean, *Ancorina alata* (Order Astrophorida) from New Zealand,
52
53
54
55
56
57
58
59
60

1
2
3 502 *Rhopaloeides odorabile* (Order Dictyocertida, Family Spongiidae) from Australia and
4
5 503 *Xestospongia muta* (Order Haplosclerida) from the Caribbean were also prevalent in
6
7
8 504 these 56 clusters (32.1%, 19.6%, 14.3% and 10.7% of clusters, respectively; Fig. 4).
9

10 505 Phylogenetic analysis was also used to compare the host-specificity of bacteria
11
12 506 observed among the 3 *Ircinia* spp. with host-specificity on a broader scale. The 4
13
14 507 common bacterial OTUs (i.e., generalist symbionts) present in all 3 Mediterranean *Ircinia*
15
16 508 spp. were related to sequences derived from unrelated host sponge species, non-sponge
17
18 509 (coral) hosts and environmental (sediment) clones (Fig. 4). These generalist symbionts
19
20 510 corresponded to 1 *Deltaproteobacterium* (IRC001) and 3 *Gammaproteobacteria*
21
22 511 (IRC006, IRC012 and IRC019), with only IRC006 forming a sequence cluster comprised
23
24 512 exclusively of sponge-derived clones. The remaining generalist symbiont OTUs formed
25
26 513 sequence clusters with not only sponge-derived clones, but also coral-derived (IRC001)
27
28 514 and sediment-derived (IRC012 and IRC016) bacteria. Further, even the bacterial OTUs
29
30 515 identified as specific to a single species of *Ircinia* in the clone libraries constructed herein
31
32 516 (Fig. 3) were closely related to sequences derived from unrelated sponge hosts and
33
34 517 environmental samples (Fig. 4), suggesting a more generalist distribution of these
35
36 518 symbionts.
37
38
39
40
41
42
43
44
45

46 520 Molecular Identification of Host Sponges

47
48 521 Consensus COI sequences from *I. fasciculata* ($n = 3$) and *I. variabilis* ($n = 3$) and
49
50 522 *I. oros* ($n = 2$) individuals yielded a 1,213 bp fragment encompassing the standard
51
52 523 barcoding region ("Folmer" partition; 676 bp) and the I3-M11 region (537 bp). No intra-
53
54 524 specific variation was observed among host species and *I. fasciculata* and *I. variabilis*
55
56
57
58
59
60

1
2
3 525 exhibited identical sequences across the entire fragment length (Table 4). *I. oros* was
4
5 526 clearly differentiated from *I. fasciculata* and *I. variabilis*, exhibiting 1.8% sequence
6
7
8 527 divergence across the entire fragment with higher variability observed in the Folmer
9
10 528 partition (Table 4). Interestingly, *I. fasciculata* and *I. variabilis* sequences were more
11
12 529 closely related to the Caribbean species *I. strobilina* (0.6% divergence) than the
13
14
15 530 sympatric Mediterranean species *I. oros* (1.8% divergence). The outgroup species
16
17 531 *Hippospongia lachne*, representing a different family in the order Dictyoceratida,
18
19 532 exhibited 4.2 to 4.6% sequence from *Ircinia* spp (Table 4).

20
21
22 533 Consensus rDNA sequences from *I. fasciculata* ($n = 6$), *I. variabilis* ($n = 7$) and *I.*
23
24 534 *oros* ($n = 8$) clones yielded a 650 to 654 bp fragment encompassing the 3'-end of the 5.8S
25
26 535 subunit (50 bp), the entire ITS-2 region (243 to 248 bp) and the 5'-end of the 28S subunit
27
28 536 (356 to 357 bp). Low levels of intra-genomic polymorphisms (IGPs) were observed and
29
30 537 variable sites occurred in all rDNA subunits and regions as inconsistent point mutations
31
32 538 (i.e., occurring at different positions). Similar to COI sequences, partial 28S rDNA
33
34 539 sequences differentiated *I. oros* from *I. fasciculata* (1.16% divergence) and *I. variabilis*
35
36 540 (1.10% divergence), but did not resolve the latter two species, as intra-specific variation
37
38 541 within *I. fasciculata* (average = 0.28%, range = 0.00 – 0.56%) and *I. variabilis* (average =
39
40 542 0.16%, range = 0.00 – 0.56) was equal or greater than inter-specific variability among
41
42 543 these species (average = 0.22%, range = 0.0 – 0.56%). ITS-2 sequences exhibited the
43
44 544 highest variability and clearly differentiated *I. fasciculata* from *I. variabilis* (2.97% \pm 0.38
45
46 545 divergence), with variability among species (range = 2.43 – 3.63%) consistent greater
47
48 546 than variability within *I. fasciculata* (range = 0.40 – 2.02%) and *I. variabilis* (range =
49
50 547 0.00 – 0.00%). Phylogenetic analysis of combined ITS-2 and 28S rDNA sequences
51
52
53
54
55
56
57
58
59
60

1
2
3 548 resolved each species into monophyletic clades with high bootstrap support and showed
4
5 549 *I. fasciculata* and *I. variabilis* were more closely related to each other than to *I. oros* (Fig.
6
7
8 550 S3), similar to the microbial symbiont communities inhabiting these hosts species (Fig.
9
10 551 5).

11 552

12 553 **DISCUSSION**

13
14
15
16
17 554 The Mediterranean sponges *I. fasciculata*, *I. variabilis* and *I. oros* were shown to
18
19
20 555 host dense communities of phylogenetically diverse microbial symbionts, as occurs in
21
22 556 congeneric hosts in the Caribbean and more distantly related HMA host sponges.
23
24 557 Microbial communities associated with *Ircinia* spp. were clearly differentiated from
25
26
27 558 ambient bacterioplankton communities, in terms of richness, diversity and composition,
28
29 559 and were comprised primarily of sponge-specific symbionts related to bacterial sequences
30
31
32 560 derived from related and unrelated host sponge species. Despite the generalist nature of
33
34 561 these symbionts among sponge hosts, each species of *Ircinia* harbored a unique microbial
35
36 562 consortium. These differences in symbiont communities among congeneric and sympatric
37
38 563 sponges suggest that factors specific to each host play a role in structuring microbial
39
40
41 564 symbiont communities in marine sponges.

42
43 565 Determining significant and ecologically relevant differences in complex
44
45 566 microbial communities is aided by the application of modern sequence-based and
46
47 567 phylogenetic statistical methods (Hughes et al., 2001; Martin, 2002; Schloss, 2008). In
48
49 568 the present study, the microbial communities in *I. fasciculata*, *I. variabilis* and *I. oros*
50
51 569 exhibited variable levels of symbiont overlap and specificity. Nearly one-fourth of each
52
53
54 570 symbiont community was shared among all 3 sponge hosts, representing 4 dominant
55
56
57
58
59
60

1
2
3 571 bacterial OTUs; however, the majority of bacterial OTUs were exclusive to a single host
4
5
6 572 species. Coupling symbiont clone libraries with phylogenetic metrics revealed species-
7
8 573 specific differences in symbiont communities among *I. fasciculata*, *I. variabilis* and *I.*
9
10 574 *oros*. While each host species harbored a similar diversity of microbial symbionts, each
11
12 575 symbiont community was comprised of specific phylogenetic lineages and exhibited clear
13
14 576 genetic differentiation based on host species. Further, symbiont community structure
15
16 577 differed significantly between *I. oros* and the other two hosts, but not between *I.*
17
18 578 *fasciculata* and *I. variabilis*. Consistent with these results, *I. fasciculata* and *I. variabilis*
19
20 579 shared the most symbiont sequences (63.7% of clones) and OTUs ($n = 11$) of all pairwise
21
22 580 comparisons among hosts.
23
24
25
26

27 581 Host-specific symbiont structuring in *Ircinia* sponges is particularly notable since
28
29 582 the vast majority these microbial symbiont communities were closely related (>97%
30
31 583 identity) to sequences derived from other host sponges. These hosts included congeneric
32
33 584 hosts from the Caribbean, as well as more distantly related sponges from the
34
35 585 Mediterranean (e.g., *Aplysina aerophoba*) and the Great Barrier Reef (e.g., *Rhopaloeides*
36
37 586 *odorabile*). Indeed, symbionts characterized as ‘specific’ to one *Ircinia* host species
38
39 587 herein were commonly observed in sponge hosts from different oceans when compared to
40
41 588 other sponge-derived microbial sequences. For example, IRC046 represents an
42
43 589 *Actinobacteria*-affiliated symbiont OTU exclusively associated with *I. oros* in the current
44
45 590 study, yet matched closely (99%) to a symbiont sequence from the Indo-Pacific sponge
46
47 591 *Theonella swinhoei*. These observed generalist patterns of specificity among host sponges
48
49 592 for individual symbionts, combined with the host species-specific structure of symbiont
50
51 593 assemblages revealed by community-level analyses, suggest that each host *Ircinia* species
52
53
54
55
56
57
58
59
60

1
2
3 594 harbors a specific mix of generalist sponge symbionts. A similar trend was observed in a
4
5 595 meta-analysis of bacterial symbionts in 10 host sponges that revealed significant pairwise
6
7
8 596 differences in symbiont communities among all host species, despite the predominance of
9
10 597 generalist (i.e., sponge-specific) symbionts comprising each microbiota (Taylor et al.,
11
12
13 598 2007).

14
15 599 The persisting question is which factors account for the observed similarities and
16
17 600 dissimilarities in microbial community structure among different host sponges. To
18
19
20 601 classify the numerous physical, chemical and biological conditions that may structure
21
22 602 symbiont communities, a framework is presented that divides putative factors into 4
23
24 603 categories based on their source (external environment vs. host sponge related) and
25
26
27 604 prevalence among host sponge species (shared vs. exclusive; Fig. 6). The influence and
28
29 605 number of factors in each category depends on the geographic and taxonomic scope of
30
31 606 each study. Congeneric and sympatric host species were examined herein, thus numerous
32
33
34 607 factors were shared among these species, including regional environmental conditions
35
36 608 and a common evolutionary trajectory, and these factors cannot account for the reported
37
38 609 differences in symbiont communities among hosts. However, factors exclusive to
39
40 610 different host species, including habitat-specific conditions and host-specific factors, and
41
42 611 their variable influence on each host may represent a source of differentiation for *Ircinia*-
43
44 612 associated microbial communities.

45
46
47
48 613 The host species investigated exhibit different zonation patterns within littoral
49
50 614 benthic landscapes and environmental factors specific to distinct habitats may play a role
51
52 615 in structuring symbiont communities. *I. variabilis* and *I. oros* occur preferentially in
53
54 616 semi-sciophilous communities, commonly inhabiting vertical relief structures where
55
56
57
58
59
60

1
2
3 617 much of the sponge diversity in these habitats exists. In contrast, *I. fasciculata* is more
4
5 618 prevalent in photophilic communities, characterized by high irradiance conditions and
6
7
8 619 dense algal growth. Irradiance levels are primary factor in structuring sponge
9
10 620 assemblages in the Mediterranean, likely as an indirect consequence of stimulating algal
11
12 621 growth and competitive pressures (Uriz et al. 1992). The distribution of host species
13
14 622 within the local landscape may thus have ecological consequences for sponge-symbiont
15
16 623 interactions and dictate symbiont composition by imposing functional performance
17
18 624 pressures. For example, the dense populations of cyanobacteria in *I. fasciculata* may
19
20 625 contribute to host nutrition and growth via the transfer of photosynthetic by-products
21
22 626 (Arillo et al., 1994; Erwin & Thacker 2008, Freeman & Thacker 2011), thereby
23
24 627 enhancing competitive ability and allowing this species to thrive in algal-dominated
25
26 628 habitats.

27
28
29
30
31 629 In addition to habitat-specific factors, a recent and shared evolutionary past may
32
33 630 influence the composition of symbiotic microbial communities in sponge hosts, as
34
35 631 suggested by the observed correlation between *Ircinia* host phylogenies and symbiont
36
37 632 community similarity. Vertical transmission of microbial symbionts, or direct parent-to-
38
39 633 offspring passage, has been observed in several marine sponges (Usher et al., 2001;
40
41 634 Ereskovsky et al., 2005; Oren et al., 2005; Enticknap et al., 2006; Caralt et al., 2007;
42
43 635 Sharp et al., 2007; Steger et al, 2008; Lee et al., 2009a; Webster et al., 2010), including
44
45 636 *Ircinia* hosts (Schmitt et al., 2007; Schmitt et al., 2008), and provides a mechanism to
46
47 637 maintain a similar microbiota mix in related sponges where successive host generations
48
49 638 are seeded as larvae with diverse microbial symbionts. Thus, while periodic horizontal
50
51 639 transmission of microbial symbionts (i.e., environmental acquisition) may contribute to
52
53
54
55
56
57
58
59
60

1
2
3 640 the homogenization of symbiont communities in unrelated sponge hosts (Schmitt et al.
4
5 641 2008), recurrent vertical transmission may dictate symbiont composition over shorter
6
7
8 642 evolutionary time scales due to legacy effects of the microbial inheritance.
9

10 643 External environmental factors and host sponge related factors are not mutually
11
12 644 exclusive and may act in concert to structure microbial symbiont communities in marine
13
14
15 645 sponges. In the case of *Ircinia* hosts from the Mediterranean, a shared evolutionary past
16
17 646 may contribute more strongly to symbiont structure compared to habitat-specific
18
19
20 647 conditions, considering that the microbiota of *I. variabilis* exhibited higher similarity to
21
22 648 the more closely related species, *I. fasciculata*, than to the neighboring species, *I. oros*.
23
24
25 649 However, a portion of the microbial community in *I. variabilis* was exclusively shared
26
27 650 with *I. fasciculata* (7 OTUs) or *I. oros* (6 OTUs), suggesting an interactive effect of these
28
29 651 factors to produce an intermediate or transitional community between the clearly
30
31
32 652 differentiated symbiont communities in *I. fasciculata* and *I. oros*.
33

34 653 Discerning the contribution of common ancestry among hosts to the microbial
35
36 654 structure of symbiont communities requires accurate resolution of host phylogenies based
37
38
39 655 on molecular data. To date, no universal molecular marker has been discovered that
40
41 656 offers species-level resolution for all sponge taxa, as barcoding genes commonly
42
43
44 657 employed in the molecular identification of other animal taxa (e.g., COI) are conserved
45
46 658 and unable to distinguish different species in some sponge groups (Erpenbeck et al.,
47
48 659 2007; López-Legentil et al., 2010a; Pöppe et al., 2010). Consistent with these findings,
49
50 660 the current study showed that COI mtDNA sequences were unable to differentiate *I.*
51
52 661 *fasciculata* and *I. variabilis*, despite extended coverage of the I3-M11 partition, a region
53
54
55 662 of the COI gene suggested to offer greater resolution for lower metazoan taxa (Erpenbeck
56
57
58
59
60

1
2
3 663 et al., 2006). The utility of the I3-M11 region in differentiating related sponge species
4
5 664 (Reveillaud et al. 2011) and as a population genetics marker for intra-specific
6
7
8 665 differentiation in sponges (López-Legentil & Pawlik, 2009; Xavier et al., 2010), suggest
9
10 666 that this gene region is either highly conserved within the genus *Ircinia* or that *I.*
11
12 667 *fasciculata* and *I. variabilis* are very closely related species. Indeed, some degree of
13
14
15 668 controversy surrounds the taxonomic status of *I. fasciculata* and *I. variabilis* (Pronzanto
16
17 669 et al., 2004) and these species can be difficult to differentiate in the field, resulting in
18
19
20 670 some ecological studies grouping these two putative species into a single taxon
21
22 671 (Maldonado et al., 2010).

23
24
25 672 Analysis of the second internal transcribed spacer region (ITS-2) of the nuclear
26
27 673 ribosomal operon revealed consistent genetic differentiation between *I. fasciculata* and *I.*
28
29 674 *variabilis*, indicating some molecular markers can delineate these species, and resolved
30
31 675 each *Ircinia* host into well-supported monophyletic clades. ITS-2 sequences have also
32
33 676 been utilized in previous phylogenetic analyses (Erwin & Thacker 2007) and population
34
35 677 genetics of marine sponges (Wörheide et al., 2002; Duran et al., 2004). The results herein
36
37 678 suggest this region offers higher fine-scale phylogenetic resolution than COI sequence
38
39 679 data for some sponge taxa. While few species of *Ircinia* have been investigated using
40
41 680 molecular phylogenetics (Pöppe et al., 2010), it is interesting to note that COI sequences
42
43 681 in *I. fasciculata* and *I. variabilis* were more closely related to *I. strobilina* from the
44
45 682 Caribbean (0.6% sequence divergence) than to the sympatric species *I. oros* (1.8%).
46
47 683 Clearly, additional phylogenetic studies and greater sampling of the high species diversity
48
49 684 within this genus are required to resolve the relationships among these species and
50
51 685 determine how these relationships affect symbiont community structure.
52
53
54
55
56
57
58
59
60

1
2
3 686 The sponge-specific nature of individual symbionts, host-specific structure of
4
5 687 symbiont communities and diverse environmental and host-related factors that dictate the
6
7 688 structure of sponge-associated microbes result in a complex picture of host-symbiont
8
9 689 specificity in HMA sponges. Based on the generalist nature of individual symbionts and
10
11 690 host-specific structure of entire communities, our results suggest a ‘specific mix of
12
13 691 generalists’ framework applies to host-symbiont specificity in *Ircinia* spp. from the
14
15 692 Mediterranean, and possibly other HMA sponge hosts. Further, by comparing symbiont
16
17 693 communities in congeneric species from the same environment, the results herein suggest
18
19 694 that factors specific to the each host species play a role in structuring microbial symbiont
20
21 695 communities in marine sponges, including host evolutionary history and habitat-specific
22
23 696 abiotic and biotic environmental factors. Additional research on the spatio-temporal
24
25 697 dynamics of microbial symbionts in sponges and controlled experimental manipulations
26
27 698 of sponge holobionts are required to further unravel the multiple and potentially
28
29 699 interactive factors that structure the complex sponge microbiota.
30
31
32
33
34
35

36 700

37
38
39 701 **ACKNOWLEDGEMENTS**

40
41 702 We thank Dr. M. Uriz (CEAB) for help with sponge taxonomy and sample identification
42
43 703 and F. Crespo (CEAB) for field assistance. This research was supported by the Spanish
44
45 704 Government projects CTM2010-17755 and CTM2010-22218 and by the US National
46
47 705 Science Foundation under grant 0853089.
48
49
50

51 706

52
53 707
54
55
56
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

708 **REFERENCES**

- 709 Altschul SF, Gish W, Miller W, Myers EW & Lipman DJ (1990) Basic local alignment
 710 search tool. *J Mol Biol* **215**: 403-410.
- 711 Anderson SA, Northcote PT & Page MJ (2010) Spatial and temporal variability of the
 712 bacterial community in different chemotypes of the New Zealand marine sponge
 713 *Mycale hentscheli*. *FEMS Microbiol Ecol* **72**: 328-342.
- 714 Angermeier H, Kamke J, Abdelmohsen UR, Krohne G, Pawlik JR, Lindquist NL &
 715 Hentschel U (2010) The pathology of sponge orange band disease affecting the
 716 Caribbean barrel sponge *Xestospongia muta*. *FEMS Microbiol Ecol* **75**: 218-230.
- 717 Arillo A, Bavestrello G, Burlando B & Sarà M (1992) Metabolic integration between
 718 symbiotic cyanobacteria and sponges: a possible mechanism. *Mar Biol* **117**: 159-162.
- 719 Ashelford KE, Chuzhanova NA, Fry JC, Jones AJ & Weightman AJ (2005) At least 1 in
 720 20 16S rRNA sequence records currently held in public repositories is estimated to
 721 contain substantial anomalies. *Appl Environ Microbiol* **71**: 7724-7736.
- 722 Ashelford KE, Chuzhanova NA, Fry JC, Jones AJ & Weightman AJ (2006) New
 723 screening software shows that most recent large 16S rRNA gene clone libraries
 724 contain chimeras. *Appl Environ Microbiol* **72**: 5734-5741.
- 725 Boneh S, Boneh A & Caron RJ (1998) Estimating the prediction function and the number
 726 of unseen species in sampling with replacement. *J Am Stat Assoc* **93**: 372-379.
- 727 Cafieri F, Fattorusso E, Santacroce C & Minale L (1972) Fasciculatin, a novel
 728 sesterterpene from the sponge *Ircinia fasciculata*. *Tetrahedron* **28**: 1579-1583.
- 729 Caralt S de, Uriz MJ & Wijffels RH (2007) Vertical transmission and successive location
 730 of symbiotic bacteria during embryo development and larva formation in *Corticium*
 731 *candelabrum* (Porifera: Demospongiae). *J Mar Biol Ass UK* **87**: 1693-1699.
- 732 Choi K, Hong J, Lee CO, Kim D, Sim CJ, et al. (2004) Cytotoxic furanosesterterpenes
 733 from a marine sponge *Psammocinia* sp. *J Nat Prod* **67**: 1186-1189.
- 734 Cole JR, Chai B, Farris RJ, Wang Q, Kulam-Syed-Mohideen AS, McGarrell DM,
 735 Bandela AM, Cardenas E, Garrity GM & Tiedje JM (2007) The ribosomal database
 736 project (RDP-II): introducing myRDP space and quality controlled public data.
 737 *Nucleic Acids Res* **35**: D169-D172.
- 738 DeSantis TZ, Hugenholtz P, Larsen N, Rojas M, Brodie EL, Keller K, Huber T, Dalevi
 739 D, Hu P & Andersen (2006) Greengenes, a chimera-checked 16S rRNA gene database
 740 and workbench compatible with ARB. *Appl Environ Microbiol* **72**: 5069-5072.
- 741 Drummond AJ, Ashton B, Buxton S, Cheung M, Cooper A, Duran C, Field M, Heled J,
 742 Kearse M, Markowitz S & et al. (2010) Geneious v5.3. Available from
 743 <http://www.geneious.com>.
- 744 Duran S, Giribet G & Turon X (2004) Phylogeographic history of the sponge *Crambe*
 745 *crambe* (Porifera, Poecilosclerida): range expansion and recent invasion of the
 746 Macaronesian islands from the Mediterranean Sea. *Mol Ecol* **13**: 109-122.
- 747 Enticknap JJ, Kelly M, Peraud O & Hill RT (2006) Characterization of a culturable
 748 alphaproteobacterial symbiont common to many marine sponges and evidence for
 749 vertical transmission via sponge larvae. *Appl Environ Microbiol* **72**: 3724-3732.
- 750 Ereskovsky AV, Gonoboleva E, Vishnyakov A (2005) Morphological evidence for
 751 vertical transmission of symbiotic bacteria in the viviparous sponge *Halisarca*
 752 *dujardini* Johnston (Porifera, Demospongiae, Halisarcida). *Mar Biol* **146**: 869-875.

- 1
2
3 753 Erpenbeck D, Breeuwer JAJ, van der Velde HC & van Soest RWM (2002) Unravelling
4 754 host and symbiont phylogenies of halichondrid sponges (Demospongiae, Porifera)
5 755 using a mitochondrial marker. *Mar Biol* **141**: 377-386.
- 7 756 Erpenbeck D, Hooper JNA & Wörheide G (2006) CO1 phylogenies in diploblasts and the
8 757 'Barcoding of Life' – are we sequencing a suboptimal portion? *Mol Ecol Notes* **6**: 550-
9 758 553.
- 10 759 Erpenbeck D, Duran S, Rutzler K, Paul V, Hooper JNA & Wörheide G (2007) Towards a
11 760 DNA taxonomy of Caribbean demosponges: a gene tree reconstructed from partial
12 761 mitochondrial CO1 gene sequences supports previous rDNA phylogenies and provides
13 762 a new perspective on the systematics of Demospongiae. *J Mar Biol Ass UK* **87**: 1563-
14 763 1570.
- 16 764 Erpenbeck D, Voigt O, Wörheide G & Lavrov DV (2009) The mitochondrial genomes of
17 765 sponges provide evidence for multiple invasions by Repetitive Hairpin-forming
18 766 Elements (RHE). *BMC Genomics* **10**: 59.
- 20 767 Erwin PM & Thacker RW (2007) Phylogenetic analyses of marine sponges within the
21 768 order Verongida: a comparison of morphological and molecular data. *Invert Biol* **126**:
22 769 220-234.
- 23 770 Erwin PM & Thacker RW (2008) Phototrophic nutrition and symbiont diversity of two
24 771 Caribbean sponge-cyanobacteria symbioses. *Mar Ecol Prog Ser* **362**: 139-147.
- 26 772 Excoffier L & Lischer HEL (2010) Arlequin suite ver 3.5: A new series of programs to
27 773 perform population genetics analyses under Linux and Windows. *Mol Ecol Resour* **10**:
28 774 564-567.
- 29 775 Flatt P, Gautschi J, Thacker R, Musafija-Girt M, Crews P & Gerwick WH (2005)
30 776 Identification of the cellular site of polychlorinated peptide biosynthesis in the marine
31 777 sponge *Dysidea (Lamellodysidea) herbacea* and symbiotic cyanobacterium
32 778 *Oscillatoria spongelliae* by CARD-FISH analysis. *Mar Biol* **147**: 761-774.
- 34 779 Folmer O, Black M, Hoeh W, Lutz R & Vrijenhoek R (1994) DNA primers for
35 780 amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan
36 781 invertebrates. *Mol Mar Biol Biotechnol* **3**: 294-299.
- 37 782 Freeman CJ & Thacker RW (2011) Complex interactions between marine sponges and
38 783 their symbiotic microbial communities. *Limnol Oceanogr* **56**: 1577-1586.
- 40 784 Gernert C, Glöckner FO, Krohne G & Hentschel U (2005) Microbial diversity of the
41 785 freshwater sponge *Spongilla lacustris*. *Microb Ecol* **50**: 206-212.
- 42 786 Guidon S & Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large
43 787 phylogenies by maximum likelihood. *Syst Biol* **52**: 696-704.
- 44 788 Hentschel U, Hopke J, Horn M, Friedrich AB, Wagner M, Hacker J & Moore BS (2002)
45 789 Molecular evidence for a uniform microbial community in sponges from different
46 790 oceans. *Appl Environ Microbiol* **68**: 4431-4440.
- 47 791 Hentschel U, Usher KM & Taylor MW (2006) Marine sponges as microbial fermenters.
48 792 *FEMS Microb Ecol* **55**: 167-177.
- 49 793 Herbert PDN, Cywinska A, Ball SL & deWaard JR (2003) Biological identifications
50 794 through DNA barcodes. *Proc R Soc London B* **270**: 313-321.
- 51 795 Hill M, Hill A, Lopez N & Harriott O (2006) Sponge-specific bacterial symbionts in the
52 796 Caribbean sponge, *Chondrilla nucula* (Demospongiae, Chondrosida). *Mar Biol* **148**:
53 797 1221-1230.

- 1
2
3 798 Hoffman F, Larsen O, Theil V, Rapp HT, Pape T, Michaelis W & Reitner J (2005) An
4 799 anaerobic world in sponges. *Geomicrobiol J* **22**: 1-10.
5
6 800 Holmes B & Blanch H (2007) Genus-specific associations of marine sponges with group
7 801 I crenarchaeotes. *Mar Biol* **150**: 759-772.
8 802 Hughes JB, Hellmann JJ, Ricketts TH & Bohannan BJM (2001) Counting the
9 803 uncountable: statistical approaches to estimating microbial diversity. *Appl Environ*
10 804 *Micrbiol* **67**: 4399-4406.
11 805 Katoh K, Misawa K, Kum K & Miyata T (2002) MAFFT: a novel method for rapid
12 806 multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res* **30**:
13 807 3059-3066.
14
15 808 Lavrov DV, Wang X & Kelly M (2008) Reconstructing ordinal relationships in the
16 809 Demospongiae using mitochondrial genomic data. *Mol Phylogenet Evol* **49**: 111-124.
17 810 Lee OO, Chui PY, Wong YH, Pawlik JR & Qian PY (2009a) Evidence for vertical
18 811 transmission of bacterial symbionts from adult to embryo in the Caribbean sponge
19 812 *Svenzea zeai*. *Appl Environ Microbiol* **75**: 6147-6156.
20 813 Lee OO, Wong YH & Qian PY (2009b) Inter- and intraspecific variations of bacterial
21 814 communities associated with marine sponges from San Juan Island, Washington. *Appl*
22 815 *Environ Microbiol* **75**: 3513-3521.
23
24 816 Lee OO, Wang Y, Yang J, Lafi FF, Al-Suwailem A & Qian Pei-Yuan (2011)
25 817 Pyrosequencing reveals highly diverse and species-specific microbial communities in
26 818 sponges from the Red Sea. *ISME J* **5**: 650-664.
27
28 819 Liu Y, Zhang S & Abreu PJM (2006) Heterocyclic terpenes: linear furano- and
29 820 pyrroloterpenoids. *Nat Prod Rep* **23**: 630-651.
30 821 López-Legentil S & Pawlik JR (2009) Genetic structure of the Caribbean giant barrel
31 822 sponge *Xestospongia muta* using the I3-M11 partition of COI. *Coral Reefs* **28**: 157-
32 823 165.
33
34 824 López-Legentil S, Erwin PM, Henkel TP, Loh TL & Pawlik JR (2010a) Phenotypic
35 825 plasticity in the Caribbean sponge *Callyspongia vaginalis* (Porifera: Haplosclerida).
36 826 *Sci Mar* **74**: 445-453.
37 827 López-Legentil S, Erwin PM, Pawlik JR & Song B (2010b) Effects of sponge bleaching
38 828 on ammonia-oxidizing *Archaea*: Distribution and relative expression of ammonia
39 829 monooxygenase genes associated with the barrel sponge *Xestospongia muta*. *Microb*
40 830 *Ecol* **60**: 561-571.
41
42 831 Maldonado M, Sánchez-Tocino L & Navarro C (2010) Recurrent disease outbreaks in the
43 832 corneous demosponges of the genus *Ircinia*: epidemic incidence and defense
44 833 mechanisms. *Mar Biol* **157**: 1577-1590.
45
46 834 Martin AP (2002) Phylogenetic approaches for describing and comparing the diversity of
47 835 microbial communities. *Appl Environ Microbiol* **68**: 3673-3682.
48 836 Martínez-Murcia AJ, Acinas SG & Rodríguez-Valera (1995) Evaluation of prokaryotic
49 837 diversity by restrictase digestion of 16S rDNA directly amplified from hypersaline
50 838 environments. *FEMS Microbiol Ecol* **17**: 247-255.
51
52 839 Meyer CP, Geller JB & Paulay G (2005) Fine scale endemism on coral reefs:
53 840 archipelagic differentiation in turbinid gastropods. *Evolution* **59**: 113-125.
54 841 Mohamed NM, Colman AS, Tal Y & Hill RT (2008a) Diversity and expression of
55 842 nitrogen fixation genes in bacterial symbionts of marine sponges. *Environ Microbiol*
56 843 **10**: 2910-2921.
57
58
59
60

- 1
2
3 844 Mohamed NM, Rao V, Hamann MT, Kelly M & Hill RT (2008b) Monitoring bacterial
4 845 diversity of the marine sponge *Ircinia strobilina* upon transfer into aquaculture. *Appl*
5 846 *Environ Microbiol* **74**: 4133-4143.
- 7 847 Mohamed NM, Saito K, Tal Y & Hill RT (2010) Diversity of aerobic and anaerobic
8 848 ammonia-oxidizing bacteria in marine sponges. *ISME J* **4**: 38-48.
- 9 849 Olson JB & McCarthy PJ (2005) Associated bacterial communities of two deep-water
10 850 sponges. *Aquat Microb Ecol* **39**: 47-55.
- 11 851 Oren ML, Steindler L & Ilan M (2005) Transmission, plasticity and the molecular
12 852 identification of cyanobacterial symbionts in the Red Sea sponge *Diacarnus*
13 853 *erythraenus*. *Mar Biol* **148**: 35-41.
- 14 854 Paul VJ & Ritson-Williams R (2008) Marine chemical ecology. *Nat Prod Rev* **25**: 662-
15 855 695.
- 16 856 Pawlik JP, McFall G & Zea S (2002) Does the odor from sponges of the genus *Ircinia*
17 857 protect them from fish predators? *J Chem Ecol* **28**: 1103-1115.
- 18 858 Pöppe J, Sutcliffe P, Hooper JNA, Wörheide G & Erpenbeck D (2010) COI barcoding
19 859 reveals new clades and radiation patterns of Indo-Pacific sponges of the Family
20 860 Irciniidae (Demospongiae: Dictyoceratida). *PLoS ONE* **5**: e9950.
- 21 861 Pronzato R, Malva R & Manconi R (2004) The taxonomic status of *Ircinia fasciculata*,
22 862 *Ircinia felix*, and *Ircinia variabilis* (Dictyoceratida, Irciniidae). *Boll Mus Ist Biol Univ*
23 863 *Genova* **68**: 553-563.
- 24 864 Radwan M, Hanora A, Zan J, Mohamed NM, Abo-Elmatty DM, Abou-El-Ela SH & Hill
25 865 RT (2010) Bacterial community analyses of two Red Sea sponges. *Mar Biotechnol* **12**:
26 866 350-360.
- 27 867 Reveillaud J, van Soest R, Derycke S, Picton B, Rigaux A & Vanreusel A (2011)
28 868 Phylogenetic relationships among NE Atlantic *Plocamionida* Topsent (1927)
29 869 (Porifera, Poecilosclerida): under-estimated diversity in reef ecosystems. *PLoS ONE* **6**:
30 870 e16533.
- 31 871 Reynolds ES (1963) The use of lead citrate at high pH as an electron-opaque stain in
32 872 electron microscopy. *J Cell Biol* **17**: 208-212.
- 33 873 Reysenbach AL, Wickham GS & Pace NR (1994) Phylogenetic analysis of the
34 874 hyperthermophilic pink filament community in Octopus Spring, Yellowstone National
35 875 Park. *Appl Environ Microbiol* **60**: 2113-2119.
- 36 876 Schloss PD (2008) Evaluating different approaches that test whether microbial
37 877 communities have the same structure. *ISME J* **2**: 265-275.
- 38 878 Schloss PD, Larget BR & Handelsman J (2004) Integration of microbial ecology and
39 879 statistics: a test to compare gene libraries. *Appl Environ Microbiol* **70**: 5485-5492.
- 40 880 Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA,
41 881 Oakley BB, Parks DH, Robinson CJ & et al. (2009) Introducing mothur: open-source,
42 882 platform-independent, community-supported software for describing and comparing
43 883 microbial communities. *Appl Environ Microbiol* **75**: 7537-7541.
- 44 884 Schmitt S, Weisz JB, Lindquist N & Hentschel U (2007) Vertical transmission of a
45 885 phylogenetically complex microbial consortium in the viviparous sponge *Ircinia felix*.
46 886 *Appl Environ Microbiol* **73**: 2067-2078.
- 47 887 Schmitt S, Angermeier H, Schiller R, Lindquist N & Hentschel U (2008) Molecular
48 888 microbial diversity survey of sponge reproductive stages and mechanistic insights into

- 1
2
3 889 the vertical transmission of microbial symbionts. *Appl Environ Microbiol* **74**: 7694-
4 890 7708.
5
6 891 Schmitt S, Deines P, Behnam F, Wagner M & Taylor MW (2011) Chloroflexi bacteria
7 892 are more diverse, abundant and similar in high than in low microbial abundance
8 893 sponges. *FEMS Microbiol Ecol* doi: 10.1111/j.1574-6941.2011.01179.x
9 894 Sharp KH, Eam B, Faulkner DJ & Haygood MG (2007) Vertical transmission of diverse
10 895 microbes in the tropical sponge *Corticium* sp. *Appl Environ Microbiol* **73**: 622-629.
11 896 Sipkema D, Holmes B, Nichols SA & Blanch HW (2009) Biological characterization of
12 897 *Haliclona* (?gellius) sp.: Sponge and associated microorganisms. *Microb Ecol* **58**:
13 898 903-920.
14
15 899 Sokal RR & Rohlf FJ (1995) *Biometry: the principles and practice of statistics in*
16 900 *biological research, 3rd ed.* Freeman, New York, NY.
17 901 Stamatakis A, Ludwig T & Meier H (2005) RAxML-III: a fast program for maximum
18 902 likelihood-based inference of large phylogenetic trees. *Bioinformatics* **21**: 456-463.
19 903 Steger D, Ettinger-Epstein P, Whalan S, Hentschel U, de Nys R, Wagner M & Taylor
20 904 MW (2008) Diversity and mode of transmission of ammonia-oxidizing archaea in
21 905 marine sponges. *Environ Microbiol* **10**: 1087-1094.
22 906 Stewart CN & Excoffier L (1996) Assessing population genetic structure and variability
23 907 with RAPD data: Application to *Vaccinium macrocarpon* (American Cranberry). *J*
24 908 *Evol Biol* **9**: 153-171.
25 909 Taylor MW, Schupp PJ, Dahllöf I, Kjelleberg S & Steinberg PD (2004) Host specificity
26 910 in marine sponge-associated bacteria, and potential implications for marine microbial
27 911 diversity. *Environ Microbiol* **6**: 121-130.
28 912 Taylor MW, Radax R, Steger D, Wagner M (2007) Sponge-associated microorganisms:
29 913 Evolution, ecology, and biotechnological potential. *Microbiol Mol Biol Rev* **71**: 295-
30 914 347.
31 915 Thacker RW & Starnes S (2003) Host specificity of the symbiotic cyanobacterium
32 916 *Oscillatoria spongelliae* in marine sponges, *Dysidea* spp. *Mar Biol* **142**: 643-648.
33 917 Thacker RW, Diaz MC, Rützler K, Erwin PM, Kimble SJA, Pierce MJ & Dillard SL
34 918 (2007) Phylogenetic relationships among the filamentous cyanobacterial symbionts of
35 919 Caribbean sponges and a comparison of photosynthetic production between sponges
36 920 hosting filamentous and unicellular cyanobacteria. *Porifera research: biodiversity,*
37 921 *innovation and sustainability* (Custódio MR, Lôbo-Hajdu G, Hajdu E & Muricy G,
38 922 eds.), pp. 621-626. Museu Nacional, Rio de Janeiro.
39 923 Turque AS, Batista D, Silveira CB, et al. (2010) Environmental shaping of sponge
40 924 associated archaeal communities. *PLoS ONE* **5**: e15774.
41 925 Uriz MJ, Rosell D & Martín D (1992) The sponge population of the Cabrera Archipelago
42 926 (Balearic Islands): characteristics, distribution, and abundance of the most
43 927 representative species. *Mar Ecol* **13**: 101-117.
44 928 Usher KM, Kuo J, Fromont J & Sutton DC (2001) Vertical transmission of
45 929 cyanobacterial symbionts in the marine sponge *Chondrilla australiensis*
46 930 (*Demospongiae*). *Hydrobiologia* **461**: 15-23.
47 931 Usher KM, Fromont J, Sutton DC & Toze S (2004) The biogeography and phylogeny of
48 932 unicellular cyanobacterial symbionts in sponges from Australia and the
49 933 Mediterranean. *Microb Ecol* **48**: 167-177.
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 934 Vacelet J (1975) Etude en microscopie electronique de l'association entre bacteries et
4 935 spongiaires du genre *Verongia* (Dictyoceratida). *J Microsc Biol Cell* **23**: 271-288.
5 936 Webster NS & Taylor MW (2011) Marine sponges and their microbial symbionts: love
6 937 and other relationships. *Environ Microbiol* doi:10.1111/j.1462-2920.2011.02460.x
7 938 Webster NS, Webb RI, Ridd MJ, Hill RT & Negri AP (2001a) The effects of copper on
8 939 the microbial community of a coral reef sponge. *Environ Microbiol* **3**: 19-31.
9 940 Webster NS, Wilson KJ, Blackall LL & Hill RT (2001b) Phylogenetic diversity of
10 941 bacteria associated with the marine sponge *Rhopaloeides odorabile*. *Appl Environ*
11 942 *Microbiol* **67**: 434-444.
12 943 Webster NS, Cobb RE, Negri AP (2008a) Temperature thresholds for bacterial symbiosis
13 944 with a sponge. *ISME J* 1-13.
14 945 Webster NS, Xavier JR, Freckelton M, Motti CA & Cobb R (2008b) Shifts in microbial
15 946 and chemical patterns within the marine sponge *Aplysina aerophoba* during a disease
16 947 outbreak. *Environ Microbiol* **10**: 3366-3376.
17 948 Webster NS, Taylor MW, Behnam F, Lucker S, Rattle T, Whalan S, Horn M & Wagner
18 949 M (2010) Deep sequencing reveals exceptional diversity and modes of transmission
19 950 for bacterial sponge symbionts. *Environ Microbiol* **12**: 2070-2082.
20 951 Webster NS, Cobb RE, Soo R, Anthony SL, Battershill CN, Whalan S & Evans-Illidge E
21 952 (2011) Bacterial community dynamics in the marine sponge *Rhopaloeides odorabile*
22 953 under in situ and ex situ cultivation. *Mar Biotechnol* **13**: 296-304.
23 954 Wiesz JB, Hentschel U, Lindquist N & Martens CS (2007) Linking abundance and
24 955 diversity of sponge-associated microbial communities to metabolic differences in host
25 956 sponges. *Mar Biol* **152**: 475-483.
26 957 Wörheide G, Hooper JNA & Degnan BM (2002) Phylogeography of western Pacific
27 958 *Leucetta 'chagoensis'* (Porifera: Calcarea) from ribosomal DNA sequences:
28 959 implications for population history and conservation of the Great Barrier Reef World
29 960 Heritage Area (Australia). *Mol Ecol* **11**: 1753-1768.
30 961 Wörheide G, Nichols SA & Goldberg J (2004) Intragenomic variation of the rDNA
31 962 internal transcribed spacers in sponges (Phylum Porifera): implications for
32 963 phylogenetic studies. *Mol Phylogenet Evol* **33**: 816-830.
33 964 Xavier JR, van Soest RWN, Breeuwer JAJ, Martins AMF & Menken SBJ (2010)
34 965 Phylogeography, genetic diversity and structure of the poecilosclerid sponge *Phorbast*
35 966 *fictitius* at oceanic islands. *Contrib Zool* **79**: 119-129.
36 967 Yang J, Sun J, Lee OO, Wong YH & Qian PY (2011) Phylogenetic diversity and
37 968 community structure of sponge-associated bacteria from mangroves of the Caribbean
38 969 Sea. *Aquat Microb Ecol* **62**: 231-240.
39 970
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

971 **TABLES**

972

973 Table 1. Composition of microbial communities in *Ircinia* spp. and ambient seawater.

974 Percentage of total clones is shown by bacterial phyla (classes of Proteobacteria shown in

975 parentheses). Values in parenthesis depict the number of 99% OTUs within each lineage.

976

Phylum (Class)	<i>I. fasciculata</i>	<i>I. variabilis</i>	<i>I. oros</i>	Seawater	Total
Proteobacteria	45.4 (16)	61.3 (22)	56.1 (17)	56.2 (28)	54.5 (65)
(α -proteobacteria)	6.5 (4)	11.3 (6)	12.2 (4)	31.5 (16)	15.0 (26)
(β -proteobacteria)	-	1.3 (1)	-	2.7 (2)	1.0 (3)
(γ -proteobacteria)	20.8 (10)	33.8 (11)	28.0 (10)	22.0 (10)	26.1 (31)
(δ -proteobacteria)	18.2 (2)	15.0 (4)	15.9 (3)	-	12.4 (5)
Cyanobacteria	32.5 (2)	7.5 (1)	-	1.4 (1)	10.2 (3)
Acidobacteria	5.2 (3)	5.0 (3)	18.3 (3)	-	7.3 (8)
Bacteroidetes	9.1 (3)	3.8 (2)	3.7 (2)	9.6 (7)	6.7 (14)
Bacillariophyta	-	2.5 (2)	9.8 (5)	8.2 (4)	5.4 (11)
Nitrospira	2.6 (1)	15.0 (1)	2.4 (1)	-	5.1 (2)
Actinobacteria	-	3.8 (2)	3.7 (2)	12.3 (3)	4.8 (7)
Chloroflexi	3.9 (3)	-	4.9 (2)	-	2.2 (5)
Chlorophyta	-	-	-	4.1 (3)	1.0 (3)
Gemmatimonadetes	1.3 (1)	1.3 (1)	1.2 (1)	-	1.0 (3)
Firmicutes	-	-	-	1.4 (1)	0.3 (1)
Planctomycetes	-	-	-	1.4 (1)	0.3 (1)
Verrucomicrobia	-	-	-	1.4 (1)	0.3 (1)
Uncertain Classification	-	-	-	4.0 (2)	1.0 (2)

977

978

979 Table 2. Diversity metrics comparing the richness, dominance and evenness of microbial
980 communities in *Ircinia* spp. and seawater. Lower and upper 95% confidence intervals are
981 shown in parentheses where available.

982

Diversity Metric	Measure	<i>I. fasciculata</i>	<i>I. oros</i>	<i>I. variabilis</i>	Seawater
Richness	Observed OTUs	29	33	34	51
	S_{obs}				
	Expected OTUs	40	43	53	188
	S_{Chao1}	(32 – 64)	(36 – 65)	(40 – 91)	(105 – 394)
Dominance	Berger-Parker	26.0%	14.6%	15.0%	8.2%
	d				
	Simpson Index	10.3	19.4	19.9	58.4
	$1/D$	(7.1 – 19.1)	(13.6 – 34.0)	(14.0 – 33.9)	(36.0 – 155.2)
Evenness	Smith & Wilson	0.14	0.24	0.20	0.70
	E_{var}				
	Simpson Index	0.32	0.48	0.47	0.64
	$E_{1/D}$				

983

984

985

986 Table 3. Pairwise statistical comparisons of genetic diversity and community structure of
987 bacterial communities in *Ircinia fasciculata* (IF), *I. variabilis* (IV), *I. oros* (IO) and
988 seawater (H2O).

989

Test	Statistic	IF – IV	IV – IO	IO – IF	IF – H2O	IV – H2O	IO – H2O
AMOVA	F _{ST}	0.0312	0.0203	0.0444	0.0827	0.0794	0.0692
	P-value	***	***	***	***	***	***
HOMOVA	B	0.0018	0.0500	0.0697	1.3499	1.2813	0.8456
	P-value	0.862	0.321	0.237	***	***	**
P-Test	Score	17.0	40.0	27.0	15.0	16.0	33.0
	P-value	***	***	***	***	***	***
LIBSHUFF	ΔC _{XY}	0.0018	0.0043	0.0181	0.0686	0.0787	0.0911
	ΔC _{YX}	0.0014	0.0031	0.0367	0.0850	0.0661	0.0691
	P-value _{XY}	0.639	*	***	***	***	***
	P-value _{YX}	0.802	0.172	***	***	***	***

990

991 * = $P < 0.05$, ** = $P < 0.005$, *** = $P < 0.001$

992

993

994 Table 4. Pairwise genetic distance (p-distance) among *Ircinia* spp. and *Hippospongia*
 995 *lachne* (Order Dictyoceratida, Family Spongiidae) for the 5'-end 'Folmer' partition
 996 (upper right; 676 bp fragment) and the I3-M11 partition (lower left; 537 bp fragment) of
 997 the mitochondrial gene cytochrome oxidase subunit I. Values shown as percentages.
 998

	<i>I. fasciculata</i>	<i>I. variabilis</i>	<i>I. strobilina</i>	<i>I. oros</i>	<i>H. lachne</i>
<i>I. fasciculata</i>	-	0	0.74	1.18	5.47
<i>I. variabilis</i>	0	-	0.74	1.18	5.47
<i>I. strobilina</i>	0.37	0.37	-	1.33	5.33
<i>I. oros</i>	2.61	2.61	2.24	-	5.47
<i>H. lachne</i>	3.17	3.17	2.79	3.54	-

999

1000

1
2
3 1001 **FIGURE LEGENDS**

4
5 1002

6
7 1003 **Figure 1.** *In situ* photographs and electron micrographs of the external morphology of

8
9 1004 host sponge species (A–C) and density and diversity of bacterial symbionts (D–G).

10
11 1005 Underwater photographs highlight the distinct coloration, growth forms and surface

12
13 1006 texture of *I. fasciculata* (A), *I. variabilis* (B) and *I. oros* (C) individuals sampled in this

14
15 1007 study. Representative micrographs show the density the symbiotic cyanobacterium,

16
17 1008 *Synechococcus spongiarum*, in ectosomal (peripheral) tissues of *I. fasciculata* (D) and *I.*

18
19 1009 *variabilis* (E) that is absent in *I. oros* tissue (G). Heterotrophic bacterial symbionts

20
21 1010 dominate the choanosomal (interior) tissue of *I. variabilis* (F) and exhibit similar

22
23 1011 morphotypes to symbiotic bacteria present in *I. oros* (G). Scale bars represent 2 cm (A–

24
25 1012 C), 2 μm (D, E) and 1 μm (F, G).

26
27 1013

28
29 1014 **Figure 2.** Similarity of bacterial 16S rRNA gene sequences recovered from *Ircinia*

30
31 1015 *fasciculata* (A), *I. variabilis* (B), *I. oros* (C) and seawater (D) to sequences in the

32
33 1016 GenBank database. Top matches from BLAST searches are grouped by sequence identity

34
35 1017 bins and sequence source: sponge-associated, coral-associated, seawater, sediment and

36
37 1018 other.

38
39 1019

40
41 1020 **Figure 3.** Specificity of microbial communities in *I. fasciculata*, *I. variabilis* and *I. oros*.

42
43 1021 Pie charts depict the percentage of total clones corresponding to each specificity

44
45 1022 category; numbers denote the total OTUs accounting for each partition.

46
47 1023

1
2
3 1024 **Figure 4.** Phylogeny of bacterial 16S rRNA gene sequences from *Ircinia fasciculata*, *I.*
4
5 1025 *oros*, *I. variabilis* and ambient seawater. Tree topology was constructed using maximum
6
7 1026 likelihood criteria and numbers on nodes depict bootstrap support (< 50% not shown).
8
9 1027 Terminal nodes denote the sequence source, bold values correspond to sequences from
10
11 1028 this study and indicate the OTU and host species followed by the total number of clones
12
13 1029 (in parentheses). For condensed clades (white triangles), the total number of sequences
14
15 1030 (in parentheses) and bootstrap support (%) are shown. Gray boxes delineate clusters
16
17 1031 containing sequences from *Ircinia* spp. herein, with black triangles indicating *Ircinia*-
18
19 1032 derived sequences from other studies. Asterisks (*) indicate clones for which near full-
20
21 1033 length sequences were recovered. BAH = Bahamas, FLO = Florida.
22
23
24
25
26

27 1034
28
29 1035 **Figure 5.** Phylogenetic tree of host sponges (left) and similarity dendrogram of bacterial
30
31 1036 symbiont communities (right) in *I. fasciculata*, *I. variabilis* and *I. oros*. Tree topology
32
33 1037 was constructed using maximum likelihood (ML) analysis of ITS-2/28S rDNA gene
34
35 1038 sequences. ML (upper) and neighbor-joining (lower) bootstrap support values are shown
36
37 1039 on internal node labels; bold values indicate support for species-level clades (gray boxes).
38
39 1040 Scale bar represents 0.03 substitutions per site. Dendrogram (right) was constructed from
40
41 1041 Bray-Curtis community similarity values among microbial symbionts in each host
42
43 1042 species. Scale represents percentage similarity values. Arrows point to outgroup taxa
44
45 1043 (left) and seawater bacterial communities (right). Full sponge phylogeny available as
46
47 1044 supplemental material (Fig. S3).
48
49
50
51
52

53 1045
54
55
56
57
58
59
60

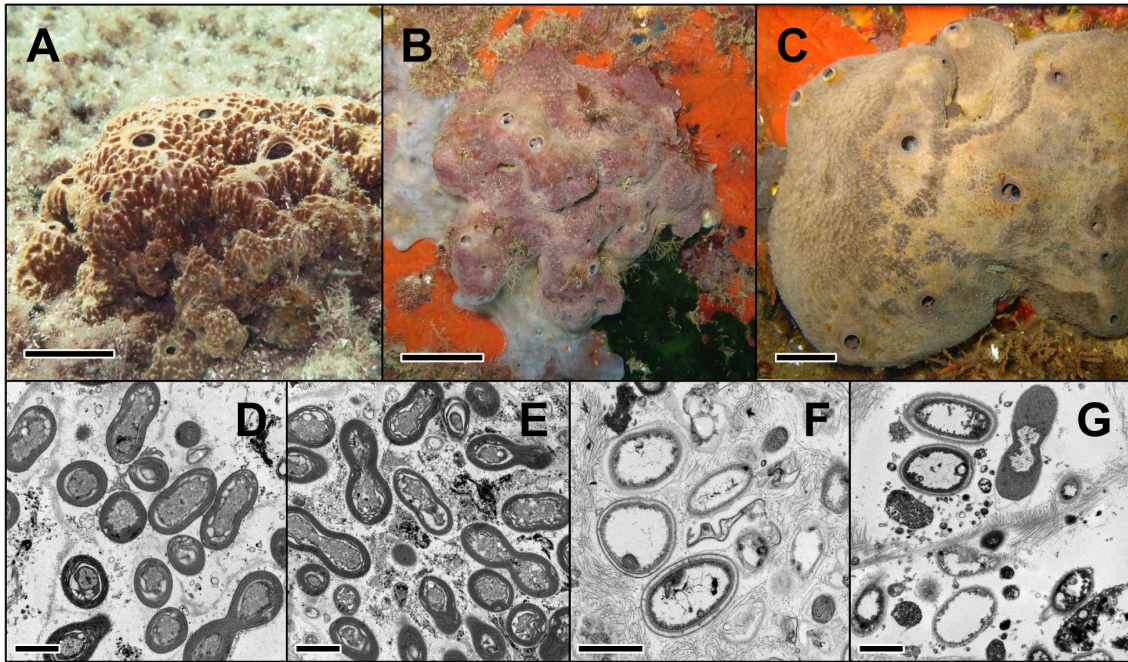
1
2
3 1046 **Figure 6.** Theoretical framework of the putative factors that structure microbial symbiont
4
5
6 1047 communities in marine sponges. Factors are first divided based on their source, occurring
7
8 1048 in the ambient external environment or within the host sponge microenvironment. Factors
9
10 1049 are further categorized by their impact on the host sponges investigated (shared among all
11
12 1050 hosts vs. exclusive to one or a subset of hosts) to distinguish factors that may homogenize
13
14 1051 (shared) or differentiate (exclusive) microbial symbiont communities among hosts.
15
16

17 1052

18
19 1053
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

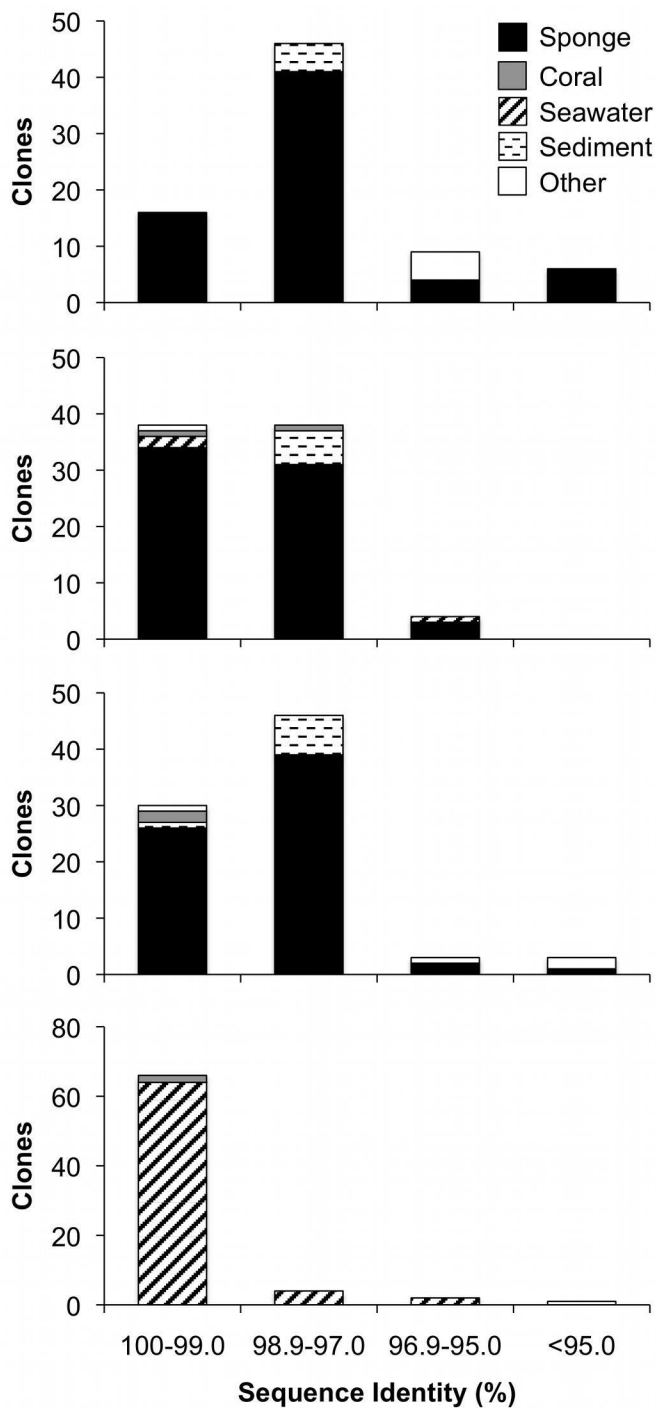
1054 FIGURES



1055

1056 Fig. 1

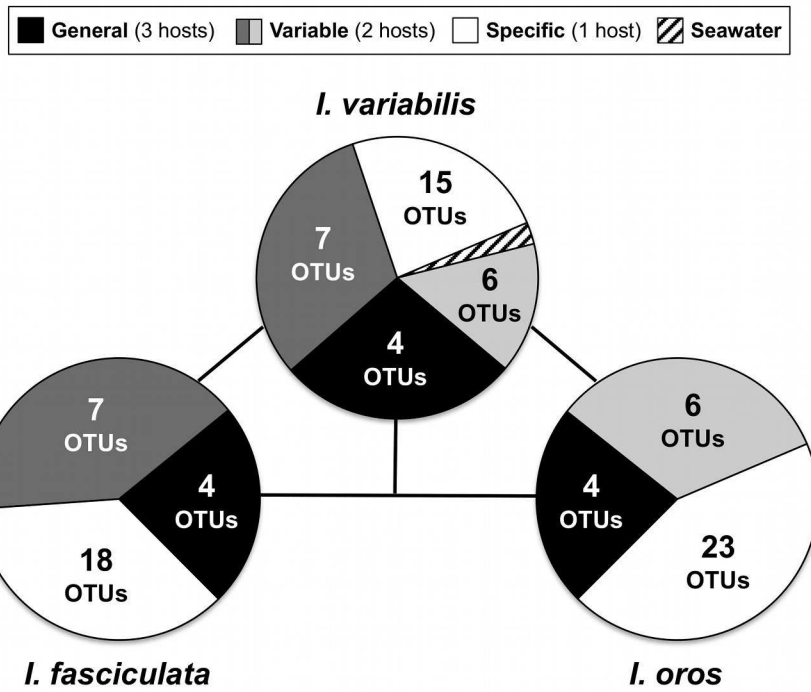
1057



review

1058

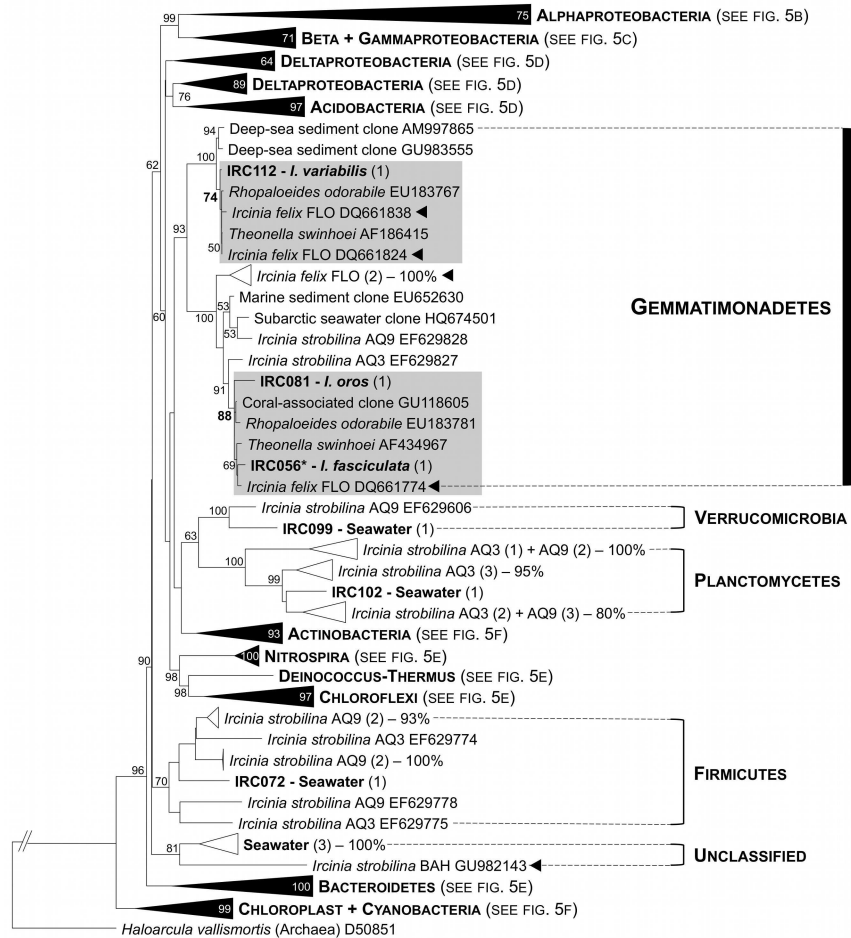
1059 Fig. 2



1060

1061 **Fig. 3**

1062

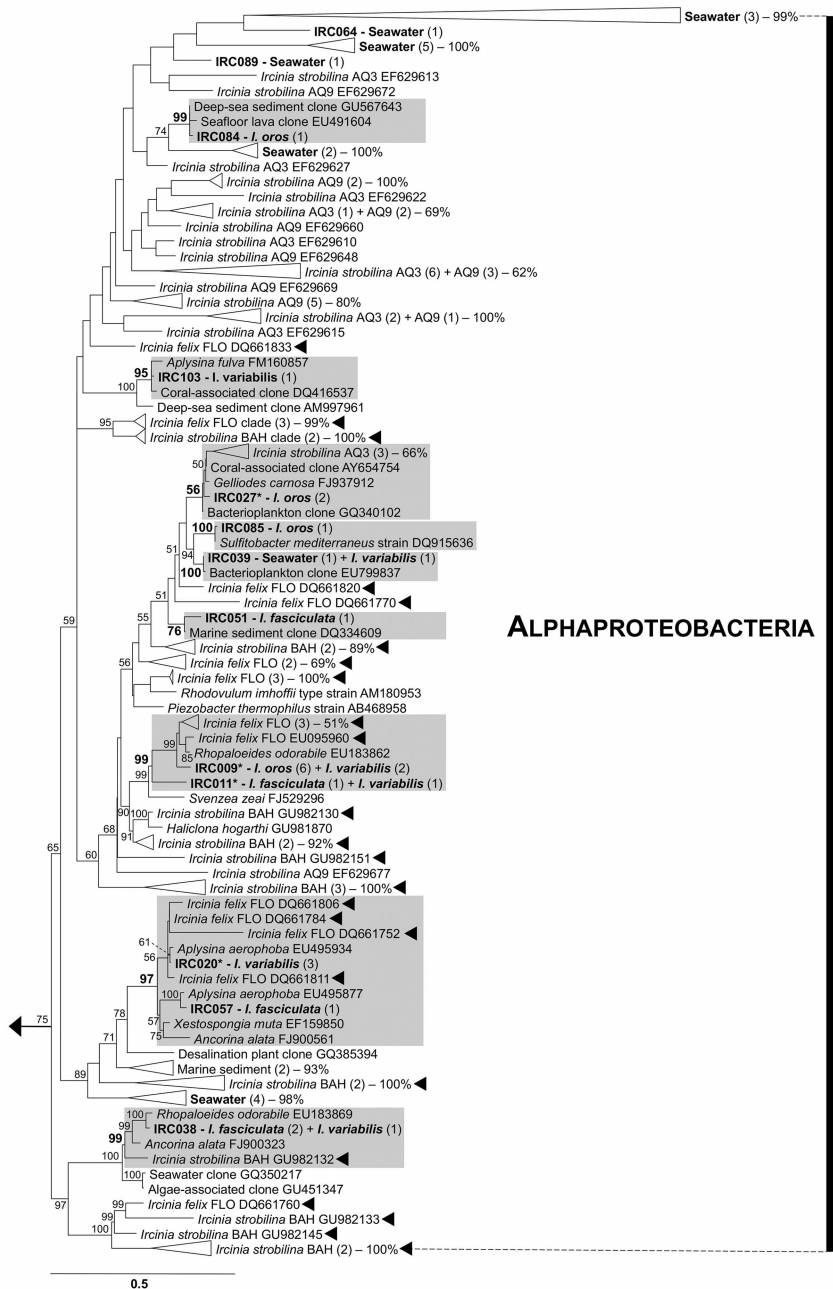


1063

1064 **Fig. 4a**

1065

review



1066

1067 **Fig. 4b**

1068

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

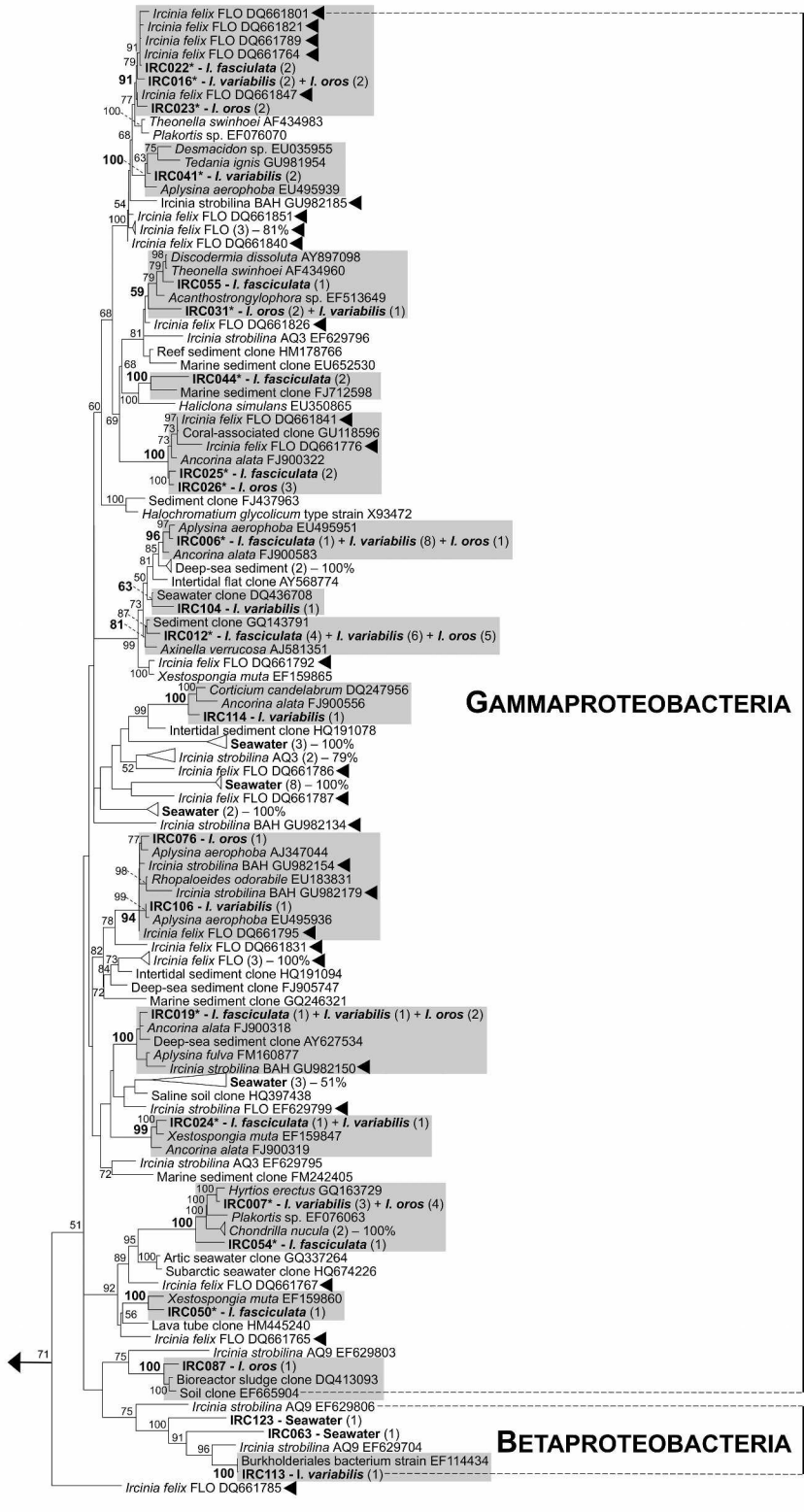
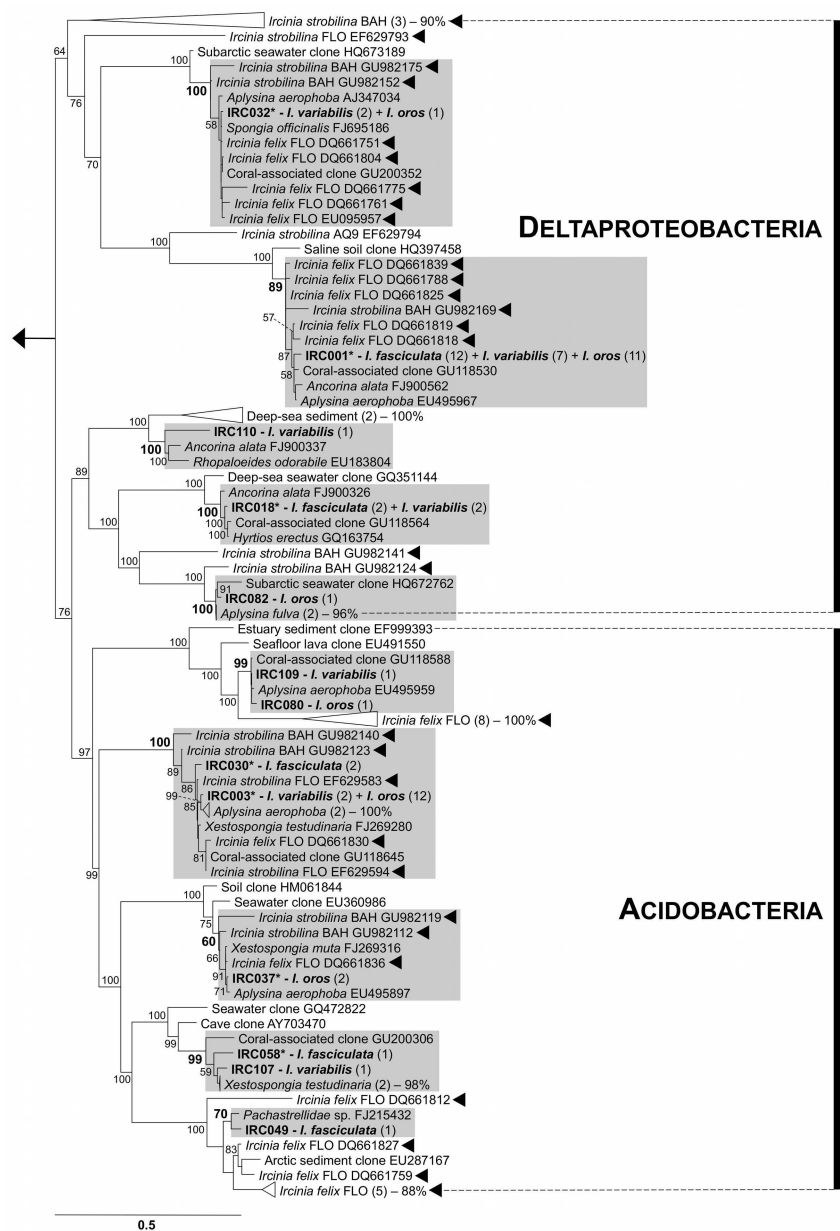
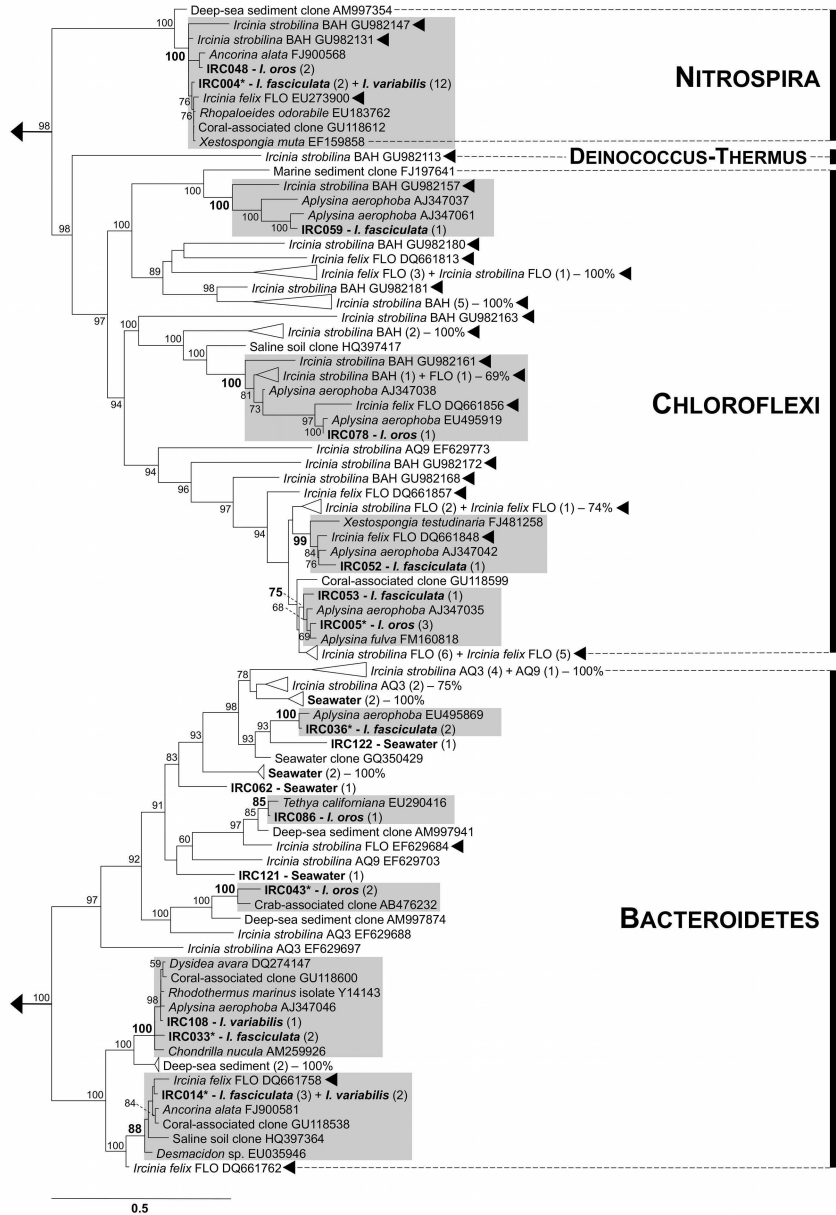


Fig. 4c





1073

1074 **Fig. 4e**

1075

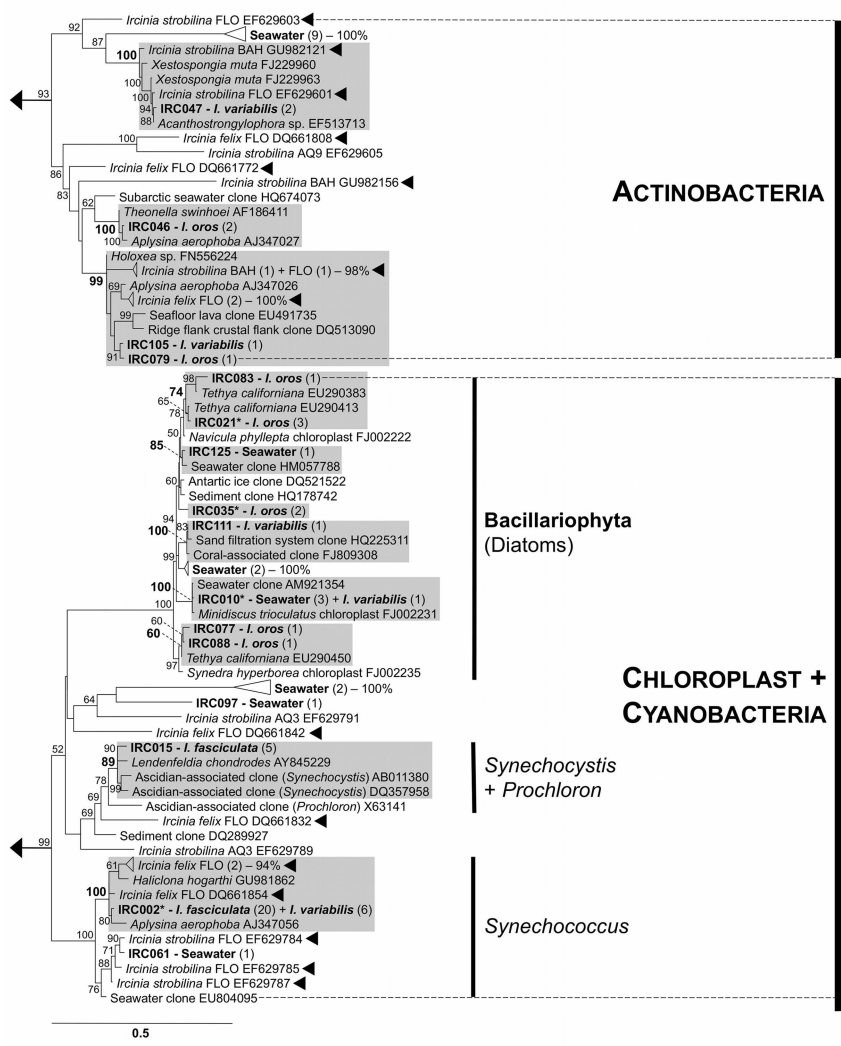
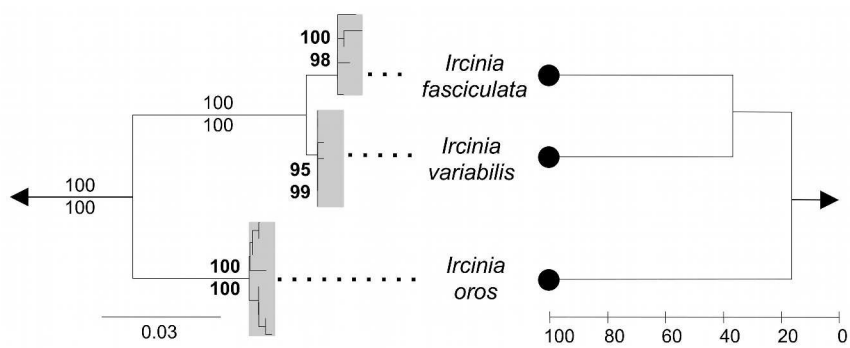


Fig. 4f

view

1076
 1077
 1078



1079

1080 **Fig. 5**

1081

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

	Shared	Exclusive
Environment	<p><u>Regional conditions</u></p> <ul style="list-style-type: none"> • Temperature <ul style="list-style-type: none"> • Salinity • POM <p>[Coastal NW Mediterranean Sea]</p>	<p><u>Habitat specific</u></p> <ul style="list-style-type: none"> • Competition • Irradiance • Predation <p>[Photophilic vs. sciophilous communities]</p>
Host Sponge	<p><u>Internal microenvironment</u></p> <ul style="list-style-type: none"> • Nutrients • Oxygen • Irradiance <p>[choanosome vs. ectosome, internal gradients]</p>	<p><u>Host species-specific</u></p> <ul style="list-style-type: none"> • Evolutionary History • Microbial Inheritance <p>[Species relatedness, vertical transmission]</p>

1082

1083 **Fig. 6**

er Review