Dissecting the microbiome of a polyketide-producing ascidian across the Anvers Island archipelago, Antarctica.

Alison E. Murray[a#], Nicole E. Avalon[b], Lucas Bishop[a], Patrick S.G. Chain[c#], Karen W. Davenport[c], Erwan Delage [d], Armand E.K. Dichosa[c], Damien Eveillard[d], Mary L. Higham[a], Sofia Kokkaliari[b], Chien-Chi Lo[c], Christian S. Riesenfeld[a], Ryan M. Young[b+], Bill J. Baker[b#]

[a] Division of Earth and Ecosystem Science, Desert Research Institute, Reno, Nevada, USA.

[b] Department of Chemistry, University of South Florida, Tampa, Florida, USA.

[c] Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA.

[d] LS2N, Université de Nantes, CNRS, Nantes, France.

Running title: Core microbiome of palmerolide-producing ascidians

Address correspondence to: Alison E. Murray, alison.murray@dri.edu, Bill J. Baker, bjbaker@usf.edu, Patrick S.G. Chain, pchain@lanl.gov

[+]Present address: School of Chemistry, National University of Ireland Galway, Galway, Ireland.

ABSTRACT

The ascidian, *S. adareanum*, from the Antarctic Peninsula near Anvers Island, is known to produce a bioactive compound, palmerolide A (PalA) that has specific activity to melanoma, a particularly invasive and metastatic form of skin cancer. The combined non-ribosomal peptide polyketide structure of PalA has similarities to microbially-produced macrolides which motivated this study utilizing culture-dependent and -independent investigations coupled with PalA detection to improve our understanding of the host-associated microbiome and relationship to PalA. Cultivation efforts yielded seven different bacteria, none of which produced PalA under the conditions tested. The genome sequence was mined for one of the most abundant members of the microbiome, *Pseudovibrio* sp. str. TunPSC04-5.I4, revealing eight biosynthetic gene clusters, none supporting the potential for PalA biosynthesis. PalA was ubiquitous and abundant across a collection of 21 ascidians (3 subsamples each) sampled from seven sites across the Anvers Island archipelago. These 63 samples were used to assess microbiome composition (V3-V4 16S rRNA gene sequence variants) which revealed a core suite of 21 bacteria, 20 of which were distinct from regional bacterioplankton. Co-occurrence analysis yielded several subsystems that may interact functionally and, although the levels of PalA detected were not found to correlate with specific sequence variants, the core members appeared to occur in a preferred optimum and tolerance range of PalA levels. Sequence variant relative abundance and biosynthetic potential of related organisms pointed to a subset of the core membership as potential PalA producers which provides a gateway to identifying the producer of palmerolides in future work.

22 **IMPORTANCE**

23

24 Palmerolide A (PalA), has potential as a chemotherapeutic agent to target melanoma. Uniform,
25 high levels of PalA were present across a regional survey of the PalA-producing Antarctic
26 ascidian, *Synoicum adareanum*, based on multi-lobed colonies sampled in the Anvers Island
27 archipelago. Likewise, we identified a core suite of microorganisms that occur concurrently with
28 the PalA-containing ascidians which spanned four phyla with lineages representing both
29 heterotrophic and chemoautotrophic (nitrifying) metabolisms and are distinct from the
30 bacterioplankton. These represent leads for the PalA producer. Cultivation efforts yielded several
31 isolates; however, none were found to produce PalA under laboratory growth conditions. One of
32 these isolates was a member of the core microbiome of the PalA-producing *S. adareanum*,
33 although the appropriate biosynthetic genes were not found in its genome. Recognition of the
34 ascidian core microbiome informs downstream studies with a target population of potential
35 PalA-producing microorganisms.

36 **INTRODUCTION**

37

38 Palmerolide A (PalA) is in a family of polyketide macrolide natural products that have been
39 isolated from the Antarctic polyclinid ascidian *Synoicum adareanum*. PalA has been found to
40 harbor anticancer properties, with selective activity against melanoma when tested in the
41 National Cancer Institute cell-line panel (1). This result is of particular interest, as there are few
42 natural product therapeutics for this devastating form of cancer. PalA inhibits vacuolar ATPases,
43 which are highly expressed in metastatic melanoma. That this Antarctic ascidian harbors this
44 bioactive compound is perhaps not entirely surprising, as ascidians are known to be rich sources
45 of bioactive natural products (2). However high latitudes are understudied concerning marine
46 natural products overall (3, 4), yet a few Antarctic studies have reported on additional
47 compounds derived from ascidians (5-7). Growing evidence suggests that in many instances,
48 ascidian-associated microbiota are responsible for bioactive natural products (e.g. anticancer,
49 antibacterial or UV-protecting activity (8, 9). We have hypothesized the same to be the case for
50 the biosynthesis of PalA based on its structure.

51 Natural products isolated from ascidians fall into polyketide, terpenoid, peptide, alkaloid
52 and a few other classes of natural products, of which the majority have cytotoxic and
53 antimicrobial activities. The diversity of ascidian-associated microbial producers currently
54 recognized are affiliated with bacterial phyla including Actinobacteria (which dominates the
55 diversity recognized), Cyanobacteria, Firmicutes, Proteobacteria (both Alphaproteobacteria and
56 Gammaproteobacteria) and Verrucomicrobia in addition to many fungi (9, 10). A few of these,
57 e.g. Cyanobacteria-affiliated *Prochloron* spp. (11), Gammaproteobacteria-affiliated *Candidatus*
58 Endoecteinascidia frumentensis (12), Alphaproteobacteria-affiliated *Candidatus*

59 Endolissoclinum faulkneri (13), and Verrucomicrobia-affiliated *Candidatus* Didemnitutus
60 mandela (14), have been identified using cultivation-independent approaches. However, this is
61 most certainly an under-representation of the diversity of ascidian-associated microorganisms
62 with biosynthetic capabilities, given the wealth of ascidian biodiversity (15).

63 One of the primary questions concerning marine benthic invertebrate-microbiome
64 associations is whether they are randomly arranged, or whether there are persistent species-
65 specific host-microbe relationships. Note that here we will use the term 'microbiome,' which
66 encompasses all the microbes and the environment in which they live. The associations may be
67 either obligate or facultative with a functional relationship that spans the spectrum from
68 mutualism to parasitism (16). The vast majority of such studies have been conducted at low- and
69 mid-latitudes from coastal to deep-sea sites. The recent application of cultivation-independent
70 studies enabled by next-generation sequencing has allowed for more extensive comparative
71 biogeographic studies of marine sponge and coral microbiomes. Ascidian-associated
72 microbiomes – studied to a lesser degree – have also been limited to low and mid-latitudes.
73 Current trends suggest that sponges across the world's oceans have a high degree of sponge
74 species-specificity of bacterial microbiomes (17), although representatives of just over 40% of
75 recognized sponge 16S rRNA gene sequence clusters have been detected in the rare biosphere in
76 marine systems (18). Corals also have a small set of conserved, core bacteria across ocean
77 habitats and basins (e.g. 19, 20), in which the structure has been proposed to occur at the level of
78 the individual (21). Ascidians have been less well studied regarding the persistence of host-
79 microbiome association, though there is supporting evidence that ascidian-associated microbiota
80 tend to have species specificity (22) and that their composition is likely linked to secondary

81 metabolite production (8). Overall, habitat and biogeographic forces appear to have strong
82 influences on marine benthic invertebrate microbiome structure and function (23-25).
83 High latitude benthic marine invertebrate-associated microbiome studies are currently
84 limited to the Antarctic, where just the tip of the iceberg has been investigated for different host-
85 microbe associations (26). Antarctic marine invertebrates tend to have a high degree of
86 endemicity at the species level, often display circumpolar distribution, and in many cases have
87 closest relatives associated with deep-sea fauna. Whether endemicity dominates high latitude
88 benthic invertebrate microbiomes is currently not known, nor is the extent of diversity
89 understood within and between different host-associated microbiomes. The few polar host-
90 associated microbiome studies to date have documented a variety of results ranging from low
91 species-specificity in sponge microbiome compositions between sub-Antarctic to South Shetland
92 Island *Mycale* species (27) and higher levels of species specificity were observed in McMurdo
93 Sound sponges (28) and across several Antarctic continental shelf sponge species (29). Webster
94 and Bourne (30) found species-specificity of bacteria across soft corals (*Alcyonium antarcticum*)
95 sampled in McMurdo Sound. Another Cnidarian, the novel ice shelf burrowing sea anemone
96 *Edswardsiella andrillae,* contained novel microbiota, though the composition across a limited set
97 of individuals was only moderately conserved (31).
98 Although several studies have linked ascidians with microorganisms that are natural
99 product producers (32), Antarctic studies have yet to link natural product biosynthesis to a
100 microbial producer. The structural similarities between PalA and microbially synthesized
101 macrolides motivated our initial investigation into host-associated microbiota of *S. adareanum*.
102 We found that bacteria of limited rRNA gene sequence diversity (including Actinobacteria,
103 Bacteroidetes, several Proteobacteria, Verrucomicrobia, and TM7) populate the *S. adareanum*

104 microbiome (33). Also, holobiont DNA screening revealed several ketosynthase sequences
105 homologous to bacterial encoded *trans*-AT polyketide genes. Together, these results motivated
106 further investigation into the *S. adareanum* microbiome given that the initial study was limited to
107 a single individual with genetic potential for PKS production.

108 Here we have designed a study to investigate the potential that there is a core microbiome
109 that persists with PalA producing *S. adareanum* holobionts. We conducted spatial survey of *S.*
110 *adareanum* in which we studied coordinated specimen-level detection and quantitation of PalA
111 along with the host-associated microbiome diversity and community structure across the Anvers
112 Island archipelago on the Antarctic Peninsula (Fig. 1). The results point to a core suite of
113 microbes associated with *S. adareanum*, distinct from the bacterioplankton, which will lead to
114 downstream testing of the hypothesis that the PalA producer is part of the core microbiome.
115 Further, these results advance us a few steps closer to realizing the goal of this study which is
116 compelled by the fact that by identifying the producer, genome sequencing could then provide
117 information on PalA biosynthesis which could lead to the development of a potential therapeutic
118 agent to fight melanoma.

119

120 **RESULTS**

121

122 **Characterization of host-associated cultivated bacteria.** Given our interest in identifying a
123 PalA producing microorganism, we executed a cultivation effort with *S. adareanum* homogenate
124 on three different marine media formulations at 10 °C. 16S rRNA gene sequencing revealed
125 seven unique isolates (of 16 brought into pure culture) at a level of > 99% sequence identity. All
126 but one of the isolates was affiliated with marine Gammaproteobacteria, including five different

127 genera (*Shewanella, Moritella, Photobacterium, Psychromonas and Pseudoalteromonas,* of
128 which nine were highly related). In many cases, we characterized their near neighbors as marine
129 psychrophiles, many from polar habitats (Fig. S1). The exception to this was the isolation of an
130 Alphaproteobacterium associated with the *Pseudovibrio* genus (*Pseudovibrio* sp. str. TunPSC04-
131 5.I4), in which we isolated two nearest neighbors from a sponge and another ascidian. This result
132 identifies the first *Pseudovibrio* sp. cultivated from high latitudes. HPLC screening results of
133 biomass from all sixteen isolates did not reveal the presence of PalA.
134
135 **Variation in holobiont PalA levels across ascidian colonies and collection sites.** Typical
136 procedures for natural products chemistry samples utilize bulk specimen collections for
137 chemistry extraction (~ 30 individual ascidian lobes per extraction in the case of *S. adareanum*).
138 Prior to this study, variation in PalA content at the individual lobe or colony level (inset, Fig. 1)
139 was unknown. Our sampling design addressed within and between colony variation at a given
140 sampling site, as well as between site variation. The sites were constrained to the region that was
141 logistically accessible by zodiac in the Anvers Island Archipelago off-shore of the United States
142 Antarctic Program (USAP)-operated Palmer Station. *S. adareanum* colonies were sampled across
143 seven dive sites (Fig. 1) in which three lobes per multi-lobed colonies were sampled at three
144 colonies per dive site, totaling 63 lobes for PalA comparison. PalA stands out as the dominant
145 peak in all LC/MS analyses of the dichloromethane:methanol soluble metabolite fraction of all
146 samples analyzed (e.g. Fig. 2a). The range in PalA levels varied about an order of magnitude
147 $0.98 \times 10^{11} – 8.12 \times 10^{11}$ ions per g host dry weight across the 63 lobes surveyed. Our study
148 design revealed lobe-to-lobe, intra-site colony level and some site-to-site differences in PalA
149 levels ($p < 0.05$) in the archipelago (Figs 2b, S2). Within a given colony, the lobe-to-lobe

150 variation was often high and significantly different in 17/21 colonies surveyed. Significant
151 differences in PalA levels between colonies were also observed within some sites (Janus Island
152 (Jan), Bonaparte Point (Bon), Laggard Island (Lag), and Litchfield Island (Lit); see Fig 1) in
153 which at least one colony had significantly different levels compared to another colony or both.
154 Despite this, we found differences between some of the sites. Namely Bon was significantly
155 lower than all six other sites. This site is the closest to the largest island, Anvers Island, and
156 Palmer Station. Samples from Killer Whale Rocks (Kil) and Lit sites were also found to have
157 significantly higher PalA levels than Jan, Bon and Norsel Point (Nor), although these did not
158 appear to have a particular spatial pattern or association with sample collection depth.
159
160 ***Synoicum adareanum* microbiome (SaM).** To test the primary hypothesis that PalA is a
161 microbial product, we first needed to understand the nature of conservation in the composition of
162 the host-associated microbiome of *S. adareanum*. We identified the microbiome structure and
163 diversity (based on the V3-V4 region of the 16S rRNA gene) with sections of the 63 samples
164 used for holobiont PalA determinations. It resulted in 461 amplicon sequence variants (ASVs)
165 distributed over 14 bacterial phyla (Table 1). The core suite of microbes, defined as those present
166 in > 80% of samples (referred to as the Core80), included 21 ASVs (six of which were present
167 across all 63 samples). The Core80 ASVs represented the majority of the sequenced Illumina
168 sequence tags (iTags; 95% on average across all 63 samples), in which the first four ASVs
169 dominated the sequence set (Fig. 3). The ASVs present in 50-79% of samples represented the
170 Dynamic50 category and contained 14 ASVs, which represented only 3.3% of the data set. The
171 remaining ASVs fell into the Variable fraction, which included 426 ASVs, representing 1.7 % of
172 the iTag sequences, yet the majority of phylogenetic richness (Table 1). Comparative statistical

173 analyses were conducted with the complete sample set, which was subsampled to the lowest
174 number (9987) of iTags per sample. This procedure was limited by one sample (Bon1b) that
175 underperformed in terms of iTag sequence yield. After the elimination of this sample from the
176 analysis, the 62-sample set had 19,003 iTags per sample with a total of 493 ASVs, the same 21
177 sequences in the Core80 (with seven common to all 62 samples), the same 14 Dynamic50
178 sequences, and a total of 458 Variable ASVs.

179 The Core80 microbiome category included a relatively high degree of phylogenetic
180 novelty, with nearly a third of the membership having low (< 95.0%) sequence identities to
181 nearest neighbors. At the same time, the rest of the Core80 ASVs (14) matched mostly
182 uncultivated marine taxa from polar seas or sediments, in many cases with identities > 97%.
183 Several of the nearest neighboring sequences originated from marine invertebrate microbiomes.
184 The Core80 ASVs were distributed across five phyla (Proteobacteria, Actinobacteria,
185 Nitrospirae, Verrucomicrobia and Bacteroidetes; Table 1) in which seven highly related ASVs
186 were associated with the Gammaproteobacteria genus *Microbulbifer,* and dominated the ASV
187 sequences (Fig. S1). Two other Gammaproteobacteria ASVs in the Core80 were affiliated with
188 the *Endozoicomonas* genus (SaM_ASV7) and *Nitrosospira* (SaM_ASV13). There were several
189 Alphaproteobacteria with nearest neighbors falling in *Pseudovibrio, Hoeflea, Sulfitobacter,* and
190 *Octadecabacter*, genera. A Rhodospiralleles-related sequence was only distantly related to
191 known taxa with an 86.6% sequence identity to its nearest neighbor. A Deltaproteobacteria
192 *Bdellovibrio*-related ASV was also part of the Core80 microbiome (SaM_ASV9) that was also
193 unique, with a sequence identity of 90.3% to its nearest neighbor. The Actinobacteria-affiliated
194 core ASV (SaM_ASV20) was related to an uncultivated Solirubrobacterales sequence. ASV11
195 most closely matched a Nitrospirae-sequence from Arctic marine sediments. There were two

196 Verrucomicrobium-affiliated sequences represented in different families (Puniceicoccaceae,
197 SaM_ASV14, and Opitutacae, SaM_ASV15). Lastly, there were two Bacteroidetes affiliated
198 ASVs: one related to polar strain, *Brumimicrobium glaciale* (SaM_ASV19), and the other to a
199 marine *Lutibacter* strain (SaM_ASV12).

200 Five Dynamic50 ASVs were affiliated with marine Bacteroidetes (Cryomorphaceae and
201 Flavobacteriacae-related), in addition to six Gammaproteobacteria (including four additional
202 *Microbulbifer*-related sequences). There was also a sponge-affiliated relative of a
203 Verrucomicrobia isolate, and a Planktomycetes-related ASV (Table 1, Fig. S1). A number of
204 these ASVs were most closely related to isolates from marine sediments.

205 Interestingly, five sequences identified from earlier cloning and sequencing efforts with
206 this host-associated microbiome (Fig. S1; 33) matched sequences in the Core80 and Dynamic50
207 data sets. Phylogenetic comparisons also revealed that the SaM isolates were distinct from the
208 Core80 and Dynamic50 except for the *Pseudovibrio* sp. TunPSC04-5.I4 isolate, which was
209 present in both the Core80, and the clone and sequencing study.

210 The hierarchical clustering of Core80 ASVs (based on Bray-Curtis similarity) across all
211 63 samples did not reveal strong trends in site or colony specific patterns (Fig. 3). There were
212 eight instances where 2 of 3 lobes paired as closest neighbors, and 4 of 8 primary clusters that
213 included three lobes derived from the same colony. Sample Bon1b clustered apart from them all.
214 In some cases, clusters could be attributed to specific ASVs. For example, cluster 2 (Fig. 3) had
215 the highest relative levels of SaM_ASV15 (an Opitutaceae-affiliated sequence), whereas 3(Fig.
216 3) had the highest relative levels of ASV4 (affiliated with *Microbulbifer* spp.).

217 Overall the community structures in the *S. adareanum* microbiomes across the 63 lobes
218 surveyed had a high degree of similarity. Bray-Curtis pairwise similarity comparisons between

219 lobes and colonies within each site were higher than 54% in all cases. When comparing the
220 averages of pairwise similarity values within and between colonies, all sites, other than Lag, had
221 higher similarity values within lobes in the same colony (ranging from 69.9 - 82.2%) compared
222 to colonies within a site (66.5-81.0%; Fig. S3), although the differences were small (1.2-6.2%),
223 and variation relatively high in both cases. We performed a two-dimensional tmMDS analysis
224 based on Bray-Curtis similarity to investigate the structure of the microbiome between sites (Fig.
225 4a). The microbiomes sampled at Kil and Lit had the highest overall degree of clustering (>75%
226 similarity) while Kil, Lag, Lit, and Jan samples all clustered at a level of 65%. The microbiomes
227 from DeLaca Island (Del) were the most dissimilar which was supported by SIMPER analysis in
228 which two of the most abundant ASVs in the Core80 were lower than the average across other
229 sites (SaM_ASV1 and 3) while others in the Core80 (SaM_ASV4, 15, and 17) were higher than
230 the average across sites. We also performed a 2D tmMDS on the SaM fractions (Core80,
231 Dynamic50, Variable) with and without permutational iterations which showed similar trends
232 although partitioning of community structures between sites was more evident with the
233 permutations (Fig. S4). Site-based clustering patterns shifted to some degree in the different SaM
234 fractions. For the Core80 alone, Jan samples clustered apart from the others. For Dynamic50,
235 both Jan and Kil were outliers. Finally, for Variable fraction, Kil and Del samples clustered apart
236 from the other sites. Variable fraction was more homogeneous, obscuring any site-to-site
237 variability, while the core displayed tighter data clouds that showed a modest level of dispersion.
238 A PERMANOVA analysis investigated the drivers of variability in community structure
239 in which we tested the role of colony-level, site-level, and stochastic environmental variation.
240 When evaluating the whole community (all sites and ASVs), site-to-site differences explained
241 25% of the variability in the microbiome (Table S1). Colony-to-colony differences explained

242 28%, while the remaining 47% of the variability was unexplained and is likely attributed to
243 stochastic environmental variation. When the SaM fractions were analyzed separately, the most
244 significant difference ($p < 0.05$) was in the Variable fraction of the microbiome in which the site-
245 to-site differences explained only 19.2% of the variation and the residual (stochastic) level
246 increased to 58.4% (Table S1). Conversely, PERMDISP (Fig. 4), a measure of dispersion about
247 the centroid calculated for each site (a measure of β-diversity), revealed differences in the
248 community structures in each SaM fraction across sites, as well as differences between the
249 microbiome fractions. The Core80 had a low level of dispersion (PERMDISP average of 9.1,
250 range: 6.8-12.1), compared to the Dynamic50 (average 28.6, range: 24.5-38.6), which were
251 moderately dispersed about the centroid, and were more variable with differences between the
252 sites more apparent; Bon, Del, and Nor had higher values compared to the other four sites. The
253 Variable fraction had high PERMDISP values (average 54.7, range: 51.6-57.6), representing
254 high differences in β-diversity in which values were relatively close across the seven sites.
255
256 **ASV co-occurrences, and relationship to PalA.** To further investigate the ascidian-microbiome
257 ecology, we performed a network analysis of the ascidian holobiont with a particular emphasis
258 on ASV co-occurrence and ASV ecological niche associated with PalA. The co-occurrence
259 network depicted a sparse graph with 102 nodes and 64 edges (average degree 1.255; diameter
260 15; average path length 5.81) that associated ASVs from similar SaM fractions (assortativity
261 coefficient of 0.41), confirming the robustness of the above discrimination. Upon inspection of
262 the co-occurrence network, a dominant connected component contained 42.1% of ASVs, in
263 which we identified three highly connected modules (via the application of an independent
264 network analysis WGCNA, soft threshold = 9, referred to here as subsystems; Fig. 5). The three

265 subsystems were not significantly associated with PalA (resp. Corr = -0.23; 0.033; -0.17 for
266 subsystems 1, 2 and 3). The network also contained several smaller systems (2-6 nodes), and 30
267 singletons, only one of which was in the Core80 (*Endozoicomonas*-affiliated). The three
268 subsystems included ASVs from the Core80, Dynamic50 and Variable SaM fractions, but were
269 unevenly distributed. Mostly driven by Core80 SaM fractions, Subsystem 1 harbored most of the
270 highly represented ASVs, including the *Microbulbifer*-related ASVs as well as the *Pseudovibrio*
271 ASV. Subsystem 2, interconnecting Subsystems 1 and 3, is mostly driven by Variable fraction
272 ASVs and includes lower relative abundant Core80 *Hoeflea* and the Opitutaceae ASVs, as well
273 as several Bacteroidetes taxa. Finally, Subsystem 3 was smaller. It included several diverse taxa
274 dominated by Dynamic50 fraction, but still including two representatives from the Core80,
275 *Lentimonas*, and Cryomorphaceae-affiliated ASVs. The overall role of Variable ASVs in the
276 network is surprising, as they appear to be critical nodes linking the Subsystems, via 13 edges
277 that lie between Subsystems 2 and 3. Another, perhaps noteworthy smaller connected component
278 is in the lower left of the graph in which three Core80 ASVs, including the chemoautotrophic
279 ammonia and nitrite oxidizers, *Nitrosomonas* and *Nitrospira*, were linked along with an
280 uncharacterized Rhodospiralles-related ASV, and a Variable ASV.

281 To address the first-order question as to whether there was a relationship between the
282 PalA concentration levels detected in the LC/MS analysis and the semi-quantitative ASV
283 occurrences of *S. adareanum* ASVs, we performed three complementary analyses: correlation
284 analysis, weighted co-occurrence networks analysis (WGCNA) and niche robust optima with
285 PalA concentration as a variable. The Pearson correlation between ASV and PalA concentrations
286 ranged from -0.45 to 0.26 at the highest, suggesting little relationship at the gross level of dry
287 weight normalized PalA and ASV relative abundance. This indicated that relative abundance of

288 ASV is not a good predictor of PalA. A complementary WGCNA result showed no significant
289 association between subsystem co-occurrence topology and PalA, pointing to the lack of
290 relationship between microbial community structure and PalA. Finally, another aspect of the
291 relationship between ASV occurrence and PalA levels was explored using the robust optimum
292 method, which estimates the ecological optimum and tolerance range for ASVs about PalA. In
293 this case we calculated the PalA niche optimum for each ASV and ranked them based on the
294 median (**Fig. S5**). Core80 ASVs showed a consistent PalA niche range. Furthermore, the median
295 optima values of the Dynamic50 ASVs lie, for the most part, in lower or higher PalA optima
296 compared to the Core80, with a substantial niche overlap between Core80. The Variable ASVs
297 collectively lie at the lower and higher extremes of the optimum and tolerance range. Altogether,
298 these complementary analyses advocated for not considering an individual nor a community
299 effect on PalA, but rather, acclimation to the high levels of PalA observed that likely rely on
300 unknown metabolic or environmental controls.
301
302 **Culture collection:microbiome and bacterioplankton comparisons.** Representation of isolates
303 in the 16S rRNA gene ASV data set was estimated by comparing the two sequence data sets,
304 albeit different ascidian samples were used for the culture collection and the *S. adareanum*
305 survey. Excepting *Pseudovibrio* sp. TunPSC04-5I4 that was present in the Core80 with a 100%
306 sequence match, two other isolates also had 100% matches to sequences in the Variable SaM
307 (BONS.1.10.24 and BOMB.9.10.19). Three other isolate 16S rRNA sequences (BOMB.3.2.14,
308 BOMB.9.10.16, BOMB.9.10.21) were found to match sequences in the Variable microbiome at a
309 level of 97% or higher. Only the *Pseudoalteromonas*-related isolate and the *Shewanella*-related
310 isolate sequences were not detected with relatives at a level of at least 97% identity to the *S.*

311 *adareanum* microbiome ASVs. The bacterioplankton composition was dominated by
312 Gammaproteobacteria ASVs (47.35% of all ASVs; Table S2) in which six of the isolates
313 (BOMB.9.10.21, BONS.1.10.24, BOMB.9.10.16, BOMB.3.2.20, BOMB.3.2.14,
314 BONSW.4.10.32) matched sequences in the bacterioplankton data set at > 99.2% identity; even
315 at a level of 95% sequence identity the remaining ten isolates did not match sequences in the
316 plankton, including the *Pseudovibrio* sp. str. TunPSC04-5I4 isolate.
317
318 **Microbiome:bacterioplankton comparisons.** Although at a high taxonomic level
319 Proteobacteria and Bacteroidetes phyla dominated the microbiome and bacterioplankton (Table
320 S2), the relative proportions varied and the taxa represented were quite different. Membership
321 between *S. adareanum* microbiome and the bacterioplankton data set indicated a low-level
322 overlap at 100% identity (Fig. S6), with 39 of 604 perfectly matched ASVs. At 100% ASV
323 sequence identity, the results indicate a single, Core80 *S. adareanum* microbiome ASV was a
324 perfect match with the bacterioplankton data set – the *Microbulbifer*-associated sequence that is
325 the most abundant across all 63 *S. adareanum* microbiome data sets. Interestingly, this sequence
326 was only identified in one bacterioplankton sample (IPY-225-9) at a low occurrence (14 of >
327 1.18 million tags distributed across 604 ASVs). There were three (of 14) Dynamic50 ASVs that
328 were perfect matches with the bacterioplankton ASVs. These were affiliated with a poorly
329 classified Bacteroidetes family Flavobacteraceae ASV (77% of SaM samples), a
330 Gammaproteobacteria-associated *Sinobacterium* (73% of SaM samples), and a second
331 Gammaproteobacteria-associated ASV that is associated with *Candidatus* Tenderia (71% of SaM
332 samples). The remaining 35 perfect match ASVs between the two data sets were classified as

333 part of the Variable microbiome. These sequences fell across four phyla and nine classes, 13 of
334 which were well-distributed across the Bacterioplankton data set samples (>50%).
335 At a level of 97% ASV sequence identity, there were three additional matches between
336 the Bacterioplankton data set and the Core80. These included an Alphaproteobacteria-related
337 *Hoeflea* and *Halocynthibacter*-related sequences, and a Nitrospira-related sequence. There were
338 also two other Dynamic50-related ASVs: these were both related to unclassified
339 Flavobacteriaceae. The rest (79 ASVs) of the matches at >97% were affiliated with the Variable
340 fraction of the microbiome.
341
342 **Potential for *Pseudovibrio* sp. str. TunPSC04-5.I4 to produce PalA**. Since it is not known
343 whether PalA is constitutively expressed, or the experimental conditions that would be required
344 to induce its expression if induced, we sequenced the genome of *Pseudovibrio* sp. str.
345 TunPSC04-5I4 even though LC/MS screening of the laboratory culture grown at 10 °C was
346 negative for PalA. The 16S rRNA gene sequence of this strain suggests it is a predominant
347 member of the Core80 (third most abundant ASV across all samples), thus sequencing the
348 genome served as a validation check for the presence of candidate putative PalA biosynthetic
349 gene clusters. The biosynthetic gene cluster (BGC) annotation pipeline (antiSMASH, 34)
350 suggested several BGC pathways including 4 non-ribosomal peptide (NRP), 1 Type I polyketide
351 synthase (PKS), 2 Type 3 PKS, a terpene-producing cluster, and a bacteriocin cluster (Fig. 6**)**.
352 The Type I PKS cluster was not congruent with what would be expected for PalA production.
353
354
355

356 **DISCUSSION**

357

358 The ultimate goal of our research is to identify the palmerolide-producing microorganism and to
359 delineate the biosynthetic gene cluster for PalA so the system can be developed into a pipeline
360 for PalA research and development as a chemotherapeutic agent to treat melanoma. Along this
361 path, this study reports our growing understanding of the microbiome composition of *S.*
362 *adareanum*. To enhance our understanding of the ecology of the PalA-containing ascidian,
363 *Synoicum adareanum*, we investigated the ascidian colonies at an individual lobe level, and
364 compared the ascidian microbiome to the plankton. This comparison allowed us to address
365 questions regarding variability in PalA content and whether a conserved core microbiome occurs
366 across these PalA-containing Antarctic ascidians, thereby supporting the logic that if a microbial
367 producer synthesizes PalA, the producing organism should be present in all PalA-containing *S.*
368 *adareanum* samples. Before this study, however, we did not have quantitative data at the level of
369 the individual ascidian lobe that forms the pedunculated *S. adareanum* colonies (Fig. 1). This
370 discussion focuses on the core microbiome, then takes a broader look at secondary metabolite
371 distributions about the microbiome and in other marine invertebrates as well as the biosynthetic
372 potential of core membership, and concludes with information gained from our initial cultivation
373 effort.

374

375 **Core microbiome.** Ascidian (host)-microbiome specificity is an active area of research.
376 Compared to sponges and corals, for example, ascidian microbiomes are less-well characterized.
377 We found that the Antarctic ascidian *S. adareanum* has a persistent core microbiome across the
378 Anvers Island archipelago that is distinct from the plankton. This dissimilarity between ascidian

379 host-associated microorganisms and bacterioplankton appears to be a consistent observation
380 across the global ocean (e.g. (35-38). The Core80 is comprised of ASVs that numerically
381 dominate the community, as well as those representing only a fraction of a percent of the
382 sequences surveyed. Although ascidian symbioses have not yet been systematically studied in
383 the Antarctic, better-studied lower latitude ascidian microbiomes provide several examples for
384 comparison. The overall trend across ascidian microbiome studies to date suggests that there is a
385 high degree of both geographical as well as host-species level specificity of microbiome
386 composition (e.g. 35, 36, 39). The same appears to be true of *S. adareanum.* Though this study
387 was restricted to a small geographical region, we identified a conserved core of 21 16S rRNA
388 gene sequence types across 63 individual pedunculate lobes studied. We attribute the detection of
389 this high degree of persistent members in part to the uniform homogenization, extraction, and
390 sequencing methodological pipeline applied. Microbiome analysis is sensitive to sequencing
391 depth, quality parameter choices, and algorithmic differences in data-processing pipelines
392 (amplicon sequence variants vs. cluster-derived operational taxonomic units), which can impact
393 direct comparisons between studies. Along these lines, the numerous highly related
394 *Microbulbifer* ASVs would have fallen into a single OTU (97% sequence identity), resulting in a
395 core with 14 members. These limitations aside, our findings are in line with several other
396 ascidian microbiome studies from lower latitudes in terms of the relative size of core
397 membership (where core definitions vary to some degree between studies). For example, *Styela*
398 *plicata*, a solitary ascidian, was reported to have a core membership of 10 (40) to 16 OTUs (41).
399 Other solitary ascidians, including *Herdmania momus* had a core of 17 OTUs (40), while two
400 *Ciona* species ranged from 8-9 OTUs. Temperate colonial ascidians *Botryliodes leachi* and
401 *Botryllus schlosseri* ranged from 10-11 members in their core microbiomes (39). Also, an

402 extensive survey of 10 different ascidian microbiomes (representing both solitary and colonial
403 forms) conducted on the Great Barrier Reef reported core memberships ranging from 2 to 35
404 OTUs (35), while the numbers of individuals surveyed in each case were only 2-3. Note that a
405 few other studies reported much higher numbers of shared OTUs ranging from 93-238 (37, 38);
406 the scale of sequencing was higher in these later studies.

407 Predicted metabolic abilities of the Core80 taxa suggest aerobic heterotrophy,
408 microaerophily and chemoautotrophy are themes amongst the Core80, in which the most
409 abundant ASVs are high molecular weight carbon degraders. The *Microbulbifer* genus has
410 members known to degrade cellulose (42), and perhaps non-coincidently, ascidians are the only
411 known invertebrate capable of cellulose biosynthesis in the marine environment (e.g., 2, 10).
412 From this, we could speculate that the *Microbulbifer* strains associated with *S. adareanum* could
413 occupy a commensal, if not somewhat antagonistic relationship (43). In support of this
414 possibility is the fact that the only overlapping sequence between the Core80 and the
415 bacterioplankton was a *Microbulbifer* sequence, which was a rare sequence in the plankton –
416 suggesting it may be an opportunistic member of the *S. adareanum* microbiome. Besides, free-
417 living and sponge-associated isolates from the *Microbulbifer* genus have been found to produce
418 bioactive compounds including pelagiomicins (44) and parabens (45), respectively. This
419 observation, in the least, suggests that the *Microbulbifer*(s) is/are likely well-adapted to their
420 ascidian host and might be considered a potential PalA producing organism.

421 The ammonia-oxidizing *Nitrosopumulis*-related Thaumarchaeota have been commonly
422 detected in ascidian microbiomes (35, 40, 41, 46) which contrasts phylogenetically, but not in
423 terms of overall function with the ammonia-oxidizing *Nitrosomonas*-ASV that was part of the
424 SaM core. The niche however has been reported to be different for the archaeal and bacterial

425 ammonia oxidizers in which the archaea tend to be found in oligotrophic systems, while the
426 bacteria (e.g. *Nitrosomonas*) can tolerate high levels of dissolved ammonia (reviewed by 47).
427 This result might reflect both the environment, and in situ *S. adareanum* tissue ammonia levels
428 where it may accumulate, as several studies have reported on the high levels of oxidizing
429 ammonia Thaumarchaeota in the coastal waters of the Anvers Island archipelago. This group,
430 however, is numerous only in winter to early spring waters (48-50), while this study was
431 conducted with samples collected in Fall when the ammonia-oxidizing Thaumarchaeota are not
432 abundant in the coastal seawater (48), advocating for the comparisons between the SaM with
433 bacterioplankton collected in both summer and winter periods.

434 In a similar vein, although we did not intentionally conduct a temporal study, the data
435 from samples collected in 2007 and 2011, appear to suggest that a number of the core
436 microorganisms are stable over time. We found several ASVs in 2011 samples that matched (at
437 100% sequence identity) cloned sequences from samples collected in 2007. Stability of the
438 ascidian microbiome over time has been reported in a few studies (32, 26, 46). Studying the
439 persistence of the core membership over the annual cycle would be interesting (and provide
440 compelling evidence for stable relationships) in this high latitude environment where light,
441 carbon production, and sea ice cover are highly variable.

442 The co-occurrence analysis indicated a few subsystems of ASVs that may interact
443 functionally within *S. adareanum*. A small side-network included the two taxa involved with the
444 2-step nitrification process, including the *Nitrosomonas* ASV mentioned above. Even though at
445 present, the functional underpinnings of the host-microbial system have not been studied, the co-
446 occurrence relationships provide fodder for hypothesis testing in the future. One interaction
447 network that warrants mentioning here is the ASVs in Subsystem 1, which harbor several Core80

448 *Microbulbifer* ASVs, and the *Pseudovibrio* ASV are linked to a *Bdellovibrio* ASV that is also a
449 member of the Core80. Members of the *Bdellovibrio* genus are obligate bacterial predators (43)
450 that penetrate the outer membrane and cell wall of their prey. The linkage position in the
451 subsystem is compelling in the sense that the *Bdellovibrio* could potentially control the
452 abundance of the "downstream" less connected members of the network. Lastly, the positions of
453 a couple of Dynamic and several the Variable ASVs in the network, as links between the
454 subsystems, was unexpected. The central positions of these ASVs suggest that they may not be
455 merely stochastic members of the microbiome; that they could play opportunistic, adaptive or
456 ecological roles in the functionality of the microbiome subsystem(s) which potentially participate
457 in different aspects of the holobiont system in particular, by promoting the switch between
458 different ecological modes supported by different subsystems. Such roles were proposed for
459 dynamic members of the *Styela plicata* microbiome (40).
460
461 **Secondary metabolite distributions and bioaccumulation in marine biota.** Although the
462 results of the archipelago spatial ascidian survey did not support a direct relationship between
463 PalA levels and the relative abundance of microbiome ASVs, the results of the PalA niche
464 analysis suggests that the Core80 ASVs occur in a preferred optimum and tolerance range of
465 PalA levels. The lack of specific ASV-PalA patterns may not be entirely surprising, as secondary
466 metabolites result from a complex combination of metabolic reactions that require a fine-tuning
467 to environmental conditions and further metabolic modeling for the sake of understanding.
468 Furthermore, these metabolites have been found to accumulate in the tissues in several different
469 marine invertebrates. The Optimal Defense Theory can be applied to marine invertebrates and
470 reflects the hypothesis that secondary metabolites are distributed in specific tissues based on

471 exposure and anatomic susceptibility for predation (51). For example, nudibranchs sequester
472 toxic compounds, which have been biosynthesized by the gastropod or acquired from their prey.
473 The toxins are concentrated in the anatomical space of their mantles, the most vulnerable portion
474 of their soft, exposed bodies (52-54). Bioaccumulation of secondary metabolites in invertebrates
475 with less anatomical differentiation is also known to occur. In the phylum Porifera, different cell
476 types and layers have been studied to determine spatial and anatomical differences in secondary
477 metabolite concentrations (55-57). Compounds have been found to be concentrated spatially on
478 the surface (e.g., 58) or apical parts of the sponge (59) in some cases. Sponges may be able to
479 differentially bioaccumulate secondary cytotoxic metabolites based on tissues more susceptible
480 to predation (60). Metabolite distribution investigations that are ascidian-specific are less well
481 documented; however, there is also evidence of ascidian secondary cytotoxic metabolite
482 bioaccumulation. The patellazoles, marine macrolides from the ascidian *Lissoclinum patella,*
483 bioaccumulate in the ascidian tissues to concentrations up to seven orders of magnitude higher
484 than their cytotoxic dose in mammalian cell lines (13, 61). Additionally, there are other instances
485 in which bioaccumulation in ascidian host tissues suggests metabolic cooperation of producer
486 and host as well as compound translocation from producer to host (1, 8, 62). Although the PalA
487 levels were normalized to grams of dry lobe weight, tissue-specific spatial localization is a
488 potentially confounding factor in the statistical analyses investigating the ASV:PalA relationship.
489
490 **Biosynthetic potential of the core.** We investigated the natural product biosynthetic potential of
491 the nine genera associated with 15 of 21 Core80 ASVs using antiSMASH (**Table 2, Table S3**).
492 From this, it appears that all genera had at least one relative at the genus level with biosynthetic
493 capacity for either polyketide or non-ribosomal peptide biosynthesis or both. Even though the

494 number of genomes available to survey were highly uneven, there is quite a disparity of
495 biosynthetic capacity between the genera analyzed, thus, it appears that *Pseudovibrio,*
496 *Nitrosomonas, Microbulbifer, Nitrospira* have the greatest capacities (in that order). Likewise,
497 *Microbulbifer*, *Pseudovibrio, Hoeflea and Opitutaceae* might be prioritized as candidate PalA
498 producers based solely on relative abundance ranking (Table 2; 32). Although we did not
499 conduct this analysis for the six ASVs that were classified at best at the family or order level, a
500 few of these might be worth considering as potential producers considering their higher-level
501 relationships with marine natural product producing lineages. For example, marine actinobacteria
502 are classically associated with the production of numerous bioactive natural products (e.g. 63,
503 64), although speculation is difficult with actinobacteria SaM_ASV20 in the core as it is only
504 distantly related to known natural product producers. Likewise, Opitutaceae-related
505 SaM_ASV15 is ranked 7 in terms of average relative abundance and falls in the same family the
506 ascidian-associated *Candidatus* Didemnitutus mandela, which harbors the biosynthetic gene
507 cluster predicted to produce mandelalide, a glycosylated polyketide (14). From this we might
508 prioritize the *Microbulbifer*, *Pseudovibrio,* and Opitutaceae ASVs as being likely PalA
509 producing candidates, with the lower relative abundance *Nitrosomonas* and *Nitrospira* ASVs
510 also holding some potential given the perhaps surprising abundance of biosynthetic gene cluster
511 content in these chemoautotrophic, and generally small genome size taxa.
512
513 **Culture dependent investigations.** The culturing effort succeeded in isolating a *Pseudovibrio*
514 strain that is a crucial member of the Core80. In addition, several other Gammaproteobacteria-
515 affiliated strains which matched sequences in the Variable SaM and the bacterioplankton were
516 cultivated. It is likely that additional media types and isolation strategies could result in

517 additional cultivated diversity. There are a number of taxa with heterotrophic lifestyles in the
518 Core80 that have been brought into pure culture (e.g. *Microbulbifer, Hoeflea*). One challenge we
519 experienced using the nutrient replete media was overgrowth of plates, even at 10 °C.

520 We were motivated to sequence the *Pseudovibrio* sp. str. TunPSC04-5I4 genome based
521 on several factors. First, it was present as part of the core microbiome, and ranks highly in terms
522 of relative abundance, and appears to be host-specific. Likewise, this follows the logic presented
523 by Tiarno et al. (32) that bioactive compound producers are often abundant, and host-specific.
524 Second, the *Pseudovibrio* genus has been reported to be an efficient natural product producer
525 (65). Lastly, even though the extracts were surveyed for PalA from str. TunPSC04-5I4 were
526 negative for PalA, the presence of silent (not transcribed) biosynthetic gene clusters has received
527 much attention (e.g. 66), thus required evaluation. The biosynthetic gene clusters identified (8 in
528 total) were interrogated for three elements, which we have predicted would be required for PalA
529 biosynthesis. First – we predict that PalA requires a hybrid non-ribosomal peptide and polyketide
530 pathway. Second – we anticipate that there are two key enzymatic domains leading to placement
531 of two distinct structural features of the polyketide backbone, a carbamoyl transferase that
532 appends a carbamate group at C-11 of the macrolide ring, and an HMGCoA synthetase that
533 inserts a pendant methyl group on an acetate C-1 position of the macrolide structure. None of
534 these potential indicators were identified in the *Pseudovibrio* sp. str. TunPSC04-5.I4 genome.
535 Thus, together with the data showing no correlation between the ASV occurrence and abundance
536 of PalA from tissues, albeit some trends in niche optima of particular taxa, the results suggest
537 this is likely not a PalA producing microorganism.

538 In conclusion, this work has advanced our understanding of the Antarctic ascidian, *S.*
539 *adareanum*, PalA distributions and microbiome in several ways. First, we found PalA to be a

540 dominant product across all samples, with some variation but no coherent trends with the site,
541 sample, or microbiome ASV. The results point to a conserved, core, microbiome represented by
542 21 ASVs, 20 of which appear to be distinct from the bacterioplankton. The phylogenetic
543 distribution of these taxa was diverse, and distinct from other ascidian microbiomes in which
544 organisms with both heterotrophic and chemosynthetic lifestyles are predicted. The co-
545 occurrence analysis suggested the potential for ecologically interacting subsystems that may play
546 out in the future as we improve our understanding of this ascidian-microbiome-natural product
547 biosynthesis system. Likewise, based on the occurrence and likelihood of natural product
548 biosynthesis, we have narrowed down the number of potential producers to less than five taxa.
549 Metagenomics may, therefore, be a means for biosynthetic gene cluster exploration in this
550 interesting Antarctic ascidian.
551
552 **MATERIALS AND METHODS**
553
554 **Cultivation-dependent effort**. *S. adareanum* samples collected by SCUBA in 2004 and 2007
555 were used for cultivation (Table S4). The 2004 specimen were archived in 20% glycerol at -80
556 °C until processing by manual homogenization using sterilized mortar and pestle prior to plating
557 a suspension onto marine agar 2216 plates in 2006. The 2007 samples were homogenized
558 immediately following collection using sterilized mortar and pestle, and suspensions were
559 prepared in 1X marine broth or filter-sterilized seawater then transferred at 4 °C to DRI. Shortly
560 after (within 1 month of collection), the 2007 isolates were cultivated on three types of media in
561 which suspensions initially stored in marine broth were plated onto marine agar (2216), while
562 homogenate preparations stored in seawater were plated onto VNSS agar media (67) and

563 amended seawater plates (3 grams yeast extract (Difco), 5 g peptone (Difco), and 0.2 grams
564 casein hydroslyate (Difco) per liter). Colonies were selected from initial plates, and purified
565 through three rounds of growth on the same media they were isolated on.
566
567 **Field sample collections for cultivation-independent efforts.** Next, a spatial survey of
568 *Synoicum adareanum* was executed in which samples were collected by SCUBA in austral fall
569 between 23 March and 3 April 2011. Seven sampling sites (depths 24.7 - 31 m; Table S5) around
570 the region accessible by Zodiac boat from Palmer Station were selected in which we sampled in
571 a nested design where three multi-lobed colonies were selected from each site, and three lobes
572 per colony were sampled (Fig. 1). Underwater video (68) was taken at each site, then video
573 footage was observed to note general ecosystem characteristics (% cover of major benthic
574 species and algae). In total, 63 *S. adareanum* lobes were sampled (9 from each site). Samples
575 were transported to Palmer Station on ice, and frozen at -80 °C until processing at DRI and USF.
576 Frozen *S. adareanum* lobes were cut longitudinally in half for parallel processing through DNA
577 and palmerolide detection pipelines.
578 Then, to address whether the composition of the *Synoicum adareanum* microbiome
579 (SaM) was distinct from the free-living bacterioplankton (< 2.5-µm fraction), we considered the
580 overlap in membership between the SaM and bacterioplankton in the water column. To
581 accomplish this, we used a reference seawater data set represented by samples that had been
582 collected in February-March 2008 (five samples) and in August-September 2008 (nine samples)
583 from LTER Station B near Anvers Island (east of the Bonaparte Point dive site), and at a few
584 other locations in the region (Table S6). Seawater samples were collected by a submersible pump
585 and acid washed silicone tubing at 10 m at Station B, and using a rosette equipped with 12 L

586 Niskin bottles for the offshore samples at 10 m and 500 m (2 samples each depth). The February-
587 March seawater samples were processed using in-line filtration with a 2.5 μm filter (Polygard,
588 Millipore) to screen larger organisms and bacterioplankton were concentrated using a tangential
589 flow filtration system and the cells were harvested on 25 mm 0.2 μm Supor filters (Millipore).
590 The August-September seawater samples were processed using inline filtration with a 3.0 μm
591 filter (Versapor, Millipore), and then bacterioplankton was collected onto 0.2 μm Sterivex
592 (Millipore) filters using a multichannel peristaltic pump (Masterflex). All filters were immersed
593 in sucrose lysis buffer (69) and stored frozen at -80 °C until extraction.
594
595 **Palmerolide A screening.** Sixty-three frozen *Synoicum adareanum* lobes were cut in half. Half
596 lobes were lyophilized and then exhaustively extracted using dichloromethane for three days,
597 followed by methanol for three days. The extracts were combined and dried on a rotary
598 evaporator. The extracts were filtered, dried down, and reconstituted at 1.0 mg/mL to ensure the
599 injected concentration was consistent. The residue was subjected to Liquid Chromatography-
600 Mass Spectrometry (LC-MS) analysis using a H_2O:ACN gradient with constant 0.05% formic
601 acid. The high-resolution mass spectra were recorded on an Agilent Technologies 6230
602 electrospray ionization Time-of-Flight (ESI-ToF) spectrometer. LC-MS was performed using a
603 C-18 Kinetex analytical column (50 x 2.1 mm; Phenomenex). The presence of PalA was verified
604 using MS/MS on an Agilent Technologies 6540 UHD Accurate-Mass QTOF LC-MS.
605 The *S. adareanum*-associated microbial culture collection was screened for the presence
606 of PalA. Isolates were cultivated in 30 mL volumes, and the resulting biomass was lyophilized
607 and screened by HPLC. Each sample was analyzed in ESI-SIM mode targeting masses of 585
608 amu and 607 amu. The reversed-phase chromatographic analysis consisted of a 0.7 mL/min

609 solvent gradient from 80% H_2O:CH_3CN to 100% CH_3CN with constant 0.05% formic acid using
610 the same C-18 column as above. The analysis was conducted for over 18 min. PalA had a
611 retention time of approximately 14 min. The twenty-four sample sequence was followed by a
612 PalA standard to confirm its analytical characteristics.
613
614 ***S. adareanum*-associated microbial cell preparation.** The outer 1-2 mm of the ascidian tissue,
615 which could contain surface-associated microorganisms, was removed using a sterilized scalpel
616 before sectioning tissue subsamples (~0.2 g) of frozen (-80 °C) *S. adareanum* (½ lobe sections).
617 Tissue samples were diced with a scalpel before homogenization in sterile autoclaved and
618 filtered seawater (1 mL) in 2 mL tubes. Each sample was homogenized (MiniLys, Bertin
619 Instruments) using sterile CK28 beads (Precellys, Bertin Instruments) 3X at 5000 rpm for 20 sec,
620 samples were placed on ice in between each homogenization. Homogenates were centrifuged at
621 500 × g at 4 °C for 5 min to pellet tissue debris. The supernatant was removed to a new tube for
622 a second spin at the same conditions. This supernatant was decanted and the cell suspension
623 centrifuged at 12,000 × g at 4 °C for 5 min to collect the microbial cells. Suspensions were
624 stored on ice, then entered an extraction pipeline in which 12 samples were processed in parallel
625 on the QIAvac 24 Plus Manifold (Qiagen).
626
627 **DNA extractions.** *S. adareanum*-associated microbial cell preparations were extracted with
628 Powerlyzer DNEasy extraction (Qiagen) following manufacturer's instructions starting at with
629 the addition of the lysis solution. Samples were processed in parallel in batches of twelve at a
630 time using the QiaVac 24 Plus Vacuum Manifold (Qiagen). The lysis step (2X at 5000 rpm for
631 60 seconds each with incubation on ice was performed on the MiniLys (Precellys, Bertin

632 Instruments) using the 0.1 mm glass beads that come with the Powerlyzer kit. DNA
633 concentrations of final preparations were estimated using Quant-iT Picogreen dsDNA Assay Kit
634 (Invitrogen) fluorescence detection on a Spectramax Gemini (Molecular Devices).
635 DNA from bacterioplankton samples was extracted following (69), and DNA from
636 bacterial cultures was extracted using the DNeasy Blood and Tissue kit (Qiagen) following
637 manufacturer's instructions. High molecular weight (HMW) genomic DNA was prepared using a
638 standard protocol (70). All DNA concentrations were estimated using Picogreen, and the HMW
639 gDNA was run on an agarose gel (0.8%) to check for quality.
640
641 **16S rRNA gene sequencing.** Illumina tag sequencing for the *S. adareanum* microbiome (SaM)
642 targeted the V3-V4 region of the rRNA gene using primers 341F CCTACGGGNBGCASCAG
643 and 806R GGACTACHVGGGTWTCTAAT. The first round of PCR amplified the V3-V4
644 region using HIFI HotStart Ready Mix (Kapa Biosystems). The first round of PCR used a
645 denaturation temperature of 95 ºC for 3 min, 20 cycles of 95 ºC for 30 sec, 55 ºC for 30 sec and
646 72 ºC for 30 sec and followed by an extension of 72 ºC for 5 minutes before holding at 4 ºC. The
647 second round of PCR added Illumina-specific sequencing adapter sequences and unique indexes,
648 permitting multiplexing, using the Nextera XT Index Kit v2 (Illumina) and HIFI HotStart Ready
649 Mix (Kapa Biosystems). The second round of PCR used a denaturation temperature of 95 ºC for
650 3 minutes, 8 cycles of 95 ºC for 30 seconds, 55 ºC for 30 seconds and 72 ºC for 30 seconds and
651 followed by an extension of 72 ºC for 5 minutes before holding at 4 ºC. Amplicons were cleaned
652 up using AMPure XP beads (Beckman Coulter). A no-template control was processed but did
653 not show a band in the V3-V4 amplicon region and was sequenced for confirmation. A Qubit
654 dsDNA HS Assay (ThermoFisher Scientific) was used for DNA concentration estimates. The

655 average size of the library was determined by the High Sensitivity DNA Kit (Agilent) and the
656 Library Quantification Kit – Illumina/Universal Kit (KAPA Biosystems) quantified the prepared
657 libraries. The amplicon pool sequenced on Illumina MiSeq generated paired end 301 bp reads
658 was demultiplexed using Illumina's bcl2fastq.

659 The bacterioplankton samples sent to the Joint Genome Institute (JGI) for library
660 preparation and paired end (2 x 250 bp) MiSeq Illumina sequencing of the variable region 4 (V4)
661 using primers 515F (5′-GTGCCAGCMGCCGCGGTAA-3′) and 806R (5′-
662 GGACTACHVGGGTWTCTAAT-3′) (71). Sequence processing included removal of PhiX
663 contaminants and Illumina adapters at JGI.

664 The identity of the cultivated isolates was confirmed by16S rRNA gene sequencing using
665 Bact27F and Bact1492R primers either by directly sequencing agarose-gel purified PCR
666 products (Qiagen), or TA cloning (Invitrogen) of PCR fragments into *E. coli* following
667 manufacturer's instructions in which three clones were sequenced for each library, plasmids
668 were purified (Qiagen) at the Nevada Genomics Center, where Sanger sequencing was
669 conducted on an ABI3700 (Applied Biosystems). Sequences were trimmed and quality checked
670 using Sequencer, v. 5.1.

671

672 **Bioinformatic analysis of 16S rRNA gene tag sequences.** We employed a QIIME2 pipeline
673 (72) using the DADA2 plug-in (73) to de-noise the data and generate amplicon sequence variant
674 (ASV) occurrence matrices for the *S. adareanum* microbiome samples. The rigor of ASV
675 determination was used in this instance given the increased ability to uncover variability in the
676 limited geographic study area, interest in uncovering patterns of host-specificity, and ultimately
677 in identifying the conserved, core members of the microbiome, at least one of which may be

678 capable of PalA biosynthesis. Sequence data sets were initially imported into QIIME2 working
679 format and the quality of forward and reverse were checked. Default trimming parameters of
680 quality 2 were used, in addition, the first 10 bases were trimmed, and reads shorter than 250
681 bases were discarded. Next the DADA2 algorithm was used to de-noise the reads (corrects
682 substitution and insertion/deletion errors and infers sequence variants). After de-noising, reads
683 were merged. The ASVs were constructed by grouping the unique full de-noised sequences (the
684 equivalent of 100% OTUs, operational taxonomic units). The ASVs were further curated in the
685 QIIME2-DADA2 pipeline by removing chimeras in each sample individually if they can be
686 exactly reconstructed by combining a left-segment and a right-segment from two more abundant
687 "parent" sequences. A pre-trained SILVA 132 99% 16S rRNA Naive Bayes classifier
688 (https://data.qiime2.org/2019.1/common/silva-132-99-nb-classifier.qza) was used to perform the
689 taxonomic classification. Compositions of the SaM ASVs were summarized by proportion at
690 different taxonomy levels, including genus, family, order, class, and phylum ranks. In order to
691 retain all samples for diversity analysis, we set lowest reads frequency per sample (n = 62
692 samples at 19003 reads; n = 63 samples at 9987 reads) as rarefaction depth to normalize the data
693 for differences in sequence count. ASVs assigned to Eukarya or with unassigned taxa (suspected
694 contaminants) were removed from the final occurrence matrix such that the final matrix read
695 counts were slightly uneven with the lowest number of reads per sample with 9961 reads.
696 The *S. adareanum* microbiome ASVs were binned into Core (highly persistent) if present
697 in ≥80% of samples (Core80), Dynamic if present in 50-79% of samples (Dynamic50) and those
698 that comprise the naturally fluctuating microbiome, or Variable fraction that was defined as those
699 ASVs present in <50% of the samples (19, 20). We used these conservative groupings of the
700 core microbiome due to the low depth of sequencing in our study (20).

701 ASV identities between the SaM, the *S. adareanum* bacterial isolates and the
702 bacterioplankton data sets were compared using CD-HIT (cd-hit-est-2d; http://cd-hit.org). The
703 larger SaM data set which included 19,003 sequences per sample was used for these comparisons
704 to maximize the ability to identify matches; note that this set does exclude one sample, Bon1b
705 which had half as many ASVs, though overall this larger data set includes nearly 200 additional
706 sequences in the Variable fraction for comparison. ASVs with 100 and 97% identity between the
707 pairwise comparisons were summarized in terms of their membership in the Core, Dynamic or
708 Variable fractions of the SaM. Likewise, CD-HIT was used to dereplicate the isolate sequences
709 at a level of 99% sequence identity, and then the dereplicated set was compared against the
710 bacterioplankton iTag data set.
711 Phylogenetic analysis of the *S. adareanum* microbiome ASVs, *S. adareanum* bacterial
712 isolates, and 16S rRNA gene cloned sequences from Riesenfeld et al. (33) was conducted with
713 respect to neighboring sequences identified in the Ribosomal Database Project and SILVA and
714 an archaea outgroup. Two maximum likelihood trees were constructed; the first with the Core80
715 ASVs, and the second with both Core80 and Dynamic50 ASVs. A total of 369 aligned positions
716 were used in both trees. A total of 1000 bootstrap replicates were run in both instances, in which
717 the percentage (≥50 %) of trees in which the associated taxa clustered together are shown next to
718 the branches.
719
720 **Statistical analyses.** T-tests were run in Statistica (v. 13) to determine significance ($p < 0.05$) of
721 site-to-site, within and between colony variation in PalA concentrations. Similarity-matrix and
722 hierarchical clustering analyses were performed using PRIMER v.7 and PERMANOVA+
723 (PRIMER-e). Analyses were performed on the complete microbiome as well as the three

724 microbiome fractions in most cases. The ASV occurrence data was square root transformed for
725 all analyses. A heat map based on the Core80 ASV occurrence was generated with the
726 transformed data, and hierarchical clustering with group average parameter was employed,
727 which was integrated with SIMPROF confidence using 9,999 permutations and a 5%
728 significance level. Bray-Curtis resemblance matrixes were created using ASV occurrences
729 without the use of a dummy variable. To determine basic patterns in community structure within
730 vs. between colony and site differences the significance was determined by t-test using Statistica
731 v. 13. Then threshold metric Multi-Dimensional Scaling (tmMDS) was conducted based on
732 Kruskal fit scheme 1, including 500 iterations, and a minimum stress of 0.001. Similarity profile
733 testing through SIMPROF were performed based on a null hypothesis that no groups would
734 demonstrate differences in ASV occurrences. This clustering algorithm was also used to generate
735 confidence levels on the MDS plot, which were set to 65% and 75%. In addition, 95% bootstrap
736 regions were calculated with 43 bootstraps per group, set to ensure a minimum rho of 0.99. In
737 order to assess the contribution of each factor to the variance of the microbial community in this
738 nested experimental design, Site(Colony), permutational multivariate analysis of variance
739 (PERMANOVA) was used. Site-based centroids were calculated and the PERMDISP algorithm
740 was used to determine the degree of dispersion around the centroid for each site. Overall site-to-
741 site difference in dispersion was determined and pairwise comparisons were also calculated with
742 9,999 permutations used to determine significance ($P(perm) < 0.05$). Exploratory analysis of the
743 major ASV contributors to similarity was performed using the SIMPER procedure based on sites
744 and colonies, with a cut off for low contributions set to 70%.
745 Co-occurrence networks were constructed using filtered ASV occurrence data sets in
746 which the ASV were filtered to only those that were present in at least five samples resulting in a

747 102 ASV data set. The 102 x 63 matrix was provided as input to FlashWeave v1.0 (74) using
748 default parameters, and visualized in Gephi v. 0.9.2 (75). Then to consider whether the ASVs in
749 the Core80, Dynamic50, or Variable fractions of the SaM were affiliated with particular levels of
750 PalA in the ascidian lobes, PalA niche robust optimum and range were computed using the
751 occurrence and dry weight normalized contextual data (76). Weighted gene correlation network
752 analysis (WGCNA package in R; 77) was used to identify modules and their correlation with
753 PalA levels. The matrix was total-sum normalized (78), and WGCNA was used in signed mode.
754 There were few modules detected, although they were not correlated with PalA. Modules were
755 projected on the FlashWeave co-occurrence network and called subsystems.
756
757 ***Pseudovibrio* sp. TunPSC04-5.I4 genome sequencing and biosynthetic gene cluster analysis**.
758 The draft genome of str. TunPSC04-5.I4 was generated using the Pacific Biosciences (PacBio)
759 sequencing technology (79). A PacBio SMRTbell™ library was constructed and sequenced on
760 the PacBio RS platform, which generated 138,864 filtered subreads totaling 669.1 Mbp. All
761 general aspects of library construction and sequencing performed at the DOE Joint Genome
762 Institute (JGI) in Walnut Creek, CA, USA can be found at http://www.jgi.doe.gov. The raw reads
763 were assembled using the Hierarchical Genome Assembly Process (HGAP version: 2.3.0,
764 protocol version 2.3.0, method RS HGAP Assembly.3, smrtpipe.py v1.87.139483; 80). The final
765 draft assembly contained 8 scaffolds, totaling 6.550 Mb in size. The input read coverage was
766 53.1X. The genome was deposited into the Joint Genome Institute's (JGI) database and at
767 GenBank. The biosynthetic gene clusters associated with this genome were analyzed using a
768 web-based annotation and cluster finding tool, antiSMASH (34).

769 Subsequently, in order to predict the likelihood of Core80 ASVs harboring the potential
770 for PalA biosynthesis we designed a meta-analysis of neighboring genomes found at the
771 Integrated Microbial Genomes (IMG) database (81). The analysis was conducted only for ASVs
772 in which confidence of taxonomic assignment was at the genus level. Therefore, genomes were
773 harvested from IMG that were associated with a total of 9 genera (*Microbulbifer* (16 genomes),
774 *Pseudovibrio* (24 genomes), *Endozoicomonas* (11 genomes), *Nitrosomonas* (19 of 68 total
775 genomes in this genus*), Nitrospira* (14 genomes), *Hoeflea* (7 genomes), *Lutibacter* (12
776 genomes), *Halocynthilibacter* (2 genomes), and in the case of *Lentimonas*, since no genomes
777 were found, we harvested 8 genomes from the *Puniciococcaceae* family). This results in 113
778 genomes that were submitted to antiSMASH for analysis. The genomes and counts of
779 biosynthetic gene clusters assigned to non-ribosomal peptide synthase, polyketide synthase, or
780 hybrid of the two classes were tabulated (Table S2).
781
782 **Data availability.** *Synoicum adareanum* microbiome Illumina sequence information and
783 associated metadata are described under NCBI BioProject PRJNA597083
784 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA597083), and the *S. adareanum* culture
785 collection 16S rRNA gene sequences were deposited in GenBank under MN960541
786 - MN960556. The bacterioplankton sequence information and associated metadata are described
787 under NCBI BioProject PRJNA602715
788 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA602715). The *Pseudovibrio* sp. str.
789 TunPSC04-5.I4 nucleotide sequence data was deposited into the JGI-IMG database (IMG
790 Genome ID # 2636416046), the European Nucleotide Archive (accession # FNLB01000001-
791 FNLB01000008; https://www.ebi.ac.uk/ena/data/view/FNLB01), and NCBI GenBank (accession

792 # FNLB00000000; https://www.ncbi.nlm.nih.gov/genome/?term=FNLB00000000). Records for
793 these Antarctic metadata and associated sequences depositions are also reflected in the mARS
794 database (http://mars.biodiversity.aq).
795
796 **ACKNOWLEDGEMENTS**
797
798 Support for this research was provided in part by the National Institute of Health award
799 (CA205932 to AEM, BJB and PSGC) with additional support from National Science Foundation
800 awards (OPP-0442857, ANT-0838776, PLR-1341339 to BJB and ANT 0632389 to AEM and
801 Postdoctoral Research Fellowship award 0532893 to CSR). Support for the sequencing of the
802 bacterioplankton and *Pseudovibrio* sp. TunPSC04-5.I4 genome was provided as part of the Joint
803 Genome Institute's Community Sequencing Program (JGI-634 to AEM).
804 The assistance of several collaborators and students is also acknowledged including
805 Charles Amsler, Margaret Amsler, Jason Cuce, Bill Dent, Alex Dussaq, Cheryl Gleasner, Alan
806 Maschek, Robert W. Read, Andrew Shilling, Santana Thomas, and the Palmer Station science
807 support staff.
808
809 **REFERENCES**

810 1. Diyabalanage T, Amsler CD, McClintock JB, Baker BJ. 2006. Palmerolide A, a cytotoxic
811 macrolide from the Antarctic tunicate *Synoicum adareanum*. J Am Chem Soc 128:5630-
812 5631.
813 2. Schmidt EW, Donia MS. 2010. Life in cellulose houses: symbiotic bacterial biosynthesis
814 of ascidian drugs and drug leads. Curr Opin in Biotech 21:827-833.
815 3. Nunez-Pons L, Avila C. 2015. Natural products mediating ecological interactions in
816 Antarctic benthic communities: a mini-review of the known molecules. Nat Prod Rep
817 32:1114-30.
818 4. Soldatou S, Baker B. 2017. Cold-water marine natural products, 2006-2016. Nat Prod
819 Rep 34:585-626.

820 5. Franco LH, Joffe EBD, Puricelli L, Tatian M, Seldes AM, Palermo JA. 1998. Indole
821 alkaloids from the tunicate *Aplidium meridianum*. J Nat Prod 61:1130-1132.
822 6. Miyata Y, Diyabalanage T, Amsler CD, McClintock JB, Valeriote FA, Baker BJ. 2007.
823 Ecdysteroids from the antarctic tunicate *Synoicum adareanum*. J Nat Prod 70:1859-1864.
824 7. Seldes A, MF Rodrıguez Brasco, Franco LH, Palermo J. 2007. Identification of two
825 meridianins from the crude extract of the tunicate *Aplidium meridianum* by tandem mass
826 spectrometry. Nat Prod Rep 21:555–563.
827 8. Schmidt EW. 2015. The secret to a successful relationship: lasting chemistry between
828 ascidians and their symbiotic bacteria. Invertebr Biol 134:88-102.
829 9. Chen L, Hu JS, Xu JL, Shao CL, Wang GY. 2018. Biological and chemical diversity of
830 ascidian-associated microorganisms. Mar Drugs 16:362.
831 10. Dou X, Dong B. 2019. Origins and Bioactivities of Natural Compounds Derived from
832 Marine Ascidians and Their Symbionts. Mar Drugs 17:670.
833 11. Schmidt EW, Nelson JT, Rasko DA, Sudek S, Eisen JA, Haygood MG, Ravel J. 2005.
834 Patellamide A and C biosynthesis by a microcin-like pathway in *Prochloron didemni*, the
835 cyanobacterial symbiont of *Lissoclinum patella*. P Natl Acad Sci USA 102:7315-7320.
836 12. Rath CM, Janto B, Earl J, Ahmed A, Hu FZ, Hiller L, Dahlgren M, Kreft R, Yu FA,
837 Wolff JJ, Kweon HK, Christiansen MA, Hakansson K, Williams RM, Ehrlich GD,
838 Sherman DH. 2011. Meta-omic characterization of the marine invertebrate microbial
839 consortium that produces the chemotherapeutic natural product ET-743. ACS Chem Biol
840 6:1244-1256.
841 13. Kwan JC, Donia MS, Han AW, Hirose E, Haygood MG, Schmidt EW. 2012. Genome
842 streamlining and chemical defense in a coral reef symbiosis. P Natl Acad Sci USA
843 109:20655-20660.
844 14. Lopera J, Miller IJ, McPhail KL, Kwan JC. 2017. Increased biosynthetic gene dosage in a
845 genome-reduced defensive bacterial symbiont. mSystems 2:e00096-17.
846 15. Shenkar N, Swalla BJ. 2011. Global diversity of Ascidiacea. PLoS One 6:e20657.
847 16. Tipton L, Darcy JL, Hynson NA. 2019. A developing symbiosis: enabling cross-talk
848 between ecologists and microbiome scientists. Front Microbiol 10:292.
849 17. Schmidt EW, Donia MS, McIntosh JA, Fricke WF, Ravel J. 2012. Origin and variation of
850 tunicate secondary metabolites. J Nat Prod 75:295-304.
851 18. Taylor MW, Tsai P, Simister RL, Deines P, Botte E, Ericson G, Schmitt S, Webster NS.
852 2013. 'Sponge-specific' bacteria are widespread (but rare) in diverse marine
853 environments. ISME Journal 7:438-443.
854 19. Ainsworth TD, Krause L, Bridge T, Torda G, Raina JB, Zakrzewski M, Gates RD,
855 Padilla-Gamino JL, Spalding HL, Smith C, Woolsey ES, Bourne DG, Bongaerts P,
856 Hoegh-Guldberg O, Leggat W. 2015. The coral core microbiome identifies rare bacterial
857 taxa as ubiquitous endosymbionts. ISME Journal 9:2261-2274.
858 20. Hernandez-Agreda A, Leggat W, Bongaerts P, Ainsworth TD. 2016. The microbial
859 signature provides insight into the mechanistic basis of coral success across reef habitats.
860 mBio 7:e00560-16.
861 21. Hernandez-Agreda A, Leggat W, Bongaerts P, Herrera C, Ainsworth TD. 2018.
862 Rethinking the coral microbiome: simplicity exists within a diverse microbial biosphere.
863 mBio 9:e00812-18.
864 22. Chen L, Fu CM, Wang GY. 2017. Microbial diversity associated with ascidians: a review
865 of research methods and application. Symbiosis 71:19-26.

866 23. Burgsdorf I, Erwin PM, Lopez-Legentil S, Cerrano C, Haber M, Frenk S, Steindler L.
867 2014. Biogeography rather than association with cyanobacteria structures symbiotic
868 microbial communities in the marine sponge *Petrosia ficiformis*. Front Microbiol 5.
869 24. Kelly LW, Williams GJ, Barott KL, Carlson CA, Dinsdale EA, Edwards RA, Haas AF,
870 Haynes M, Lim YW, McDole T, Nelson CE, Sala E, Sandin SA, Smith JE, Vermeij
871 MJA, Youle M, Rohwer F. 2014. Local genomic adaptation of coral reef-associated
872 microbiomes to gradients of natural variability and anthropogenic stressors. P Natl Acad
873 Sci USA 111:10227-10232.
874 25. Pantos O, Bongaerts P, Dennis PG, Tyson GW, Hoegh-Guldberg O. 2015. Habitat-
875 specific environmental conditions primarily control the microbiomes of the coral
876 *Seriatopora hystrix*. ISME Journal 9:1916-1927.
877 26. Lo Guidice A, Azzaro M, Schiaparelli S. 2019. Microbial symbionts of Antarctic marine
878 benthic invertebrates, p 277-296. *In* Castro-Sowinski S (ed), The ecological role of
879 microorganisms in the Antarctic enviornment. Springer Polar Sciences, Springer Nature,
880 Sitzerland.
881 27. Cardenas CA, Gonzalez-Aravena M, Font A, Hestetun JT, Hajdu E, Trefault N,
882 Malmbergg M, Bongcarn-Rudloff E. 2018. High similarity in the microbiota of cold-
883 water sponges of the Genus *Mycale* from two different geographical areas. Peerj
884 6:10.7717/peerj.4935.
885 28. Webster NS, Negri AP, Munro M, Battershill CN. 2004. Diverse microbial communities
886 inhabit Antarctic sponges. Environ Microbiol 6:288-300.
887 29. Steinert G, Wemheuer B, Janussen D, Erpenbeck D, Daniel R, Simon M, Brinkhoff T,
888 Schupp PJ. 2019. Prokaryotic diversity and community patterns in Antarctic continental
889 shelf sponges. Front Mar Sci 6:297.
890 30. Webster NS, Bourne D. 2007. Bacterial community structure associated with the
891 Antarctic soft coral, *Alcyonium antarcticum*. FEMS Microbiol Ecol 59:81-94.
892 31. Murray AE, Rack FR, Zook R, Williams MJM, Higham ML, Broe M, Kaufmann RS,
893 Daly M. 2016. Microbiome composition and diversity of the ice-dwelling sea anemone,
894 *Edwardsiella andrillae*. Integr Comp Biol 56:542-555.
895 32. Tianero MDB, Kwan JC, Wyche TP, Presson AP, Koch M, Barrows LR, Bugni TS,
896 Schmidt EW. 2015. Species specificity of symbiosis and secondary metabolism in
897 ascidians. ISME Journal 9:615-628.
898 33. Riesenfeld CS, Murray AE, Baker BJ. 2008. Characterization of the microbial
899 community and polyketide biosynthetic potential in the Palmerolide-producing tunicate,
900 *Synoicum adareanum*. J Nat Prod 71:1812-1818.
901 34. Blin K, Medema MH, Kottmann R, Lee SY, Weber T. 2017. The antiSMASH database, a
902 comprehensive database of microbial secondary metabolite biosynthetic gene clusters.
903 Nucleic Acids Res 45:D555-D559.
904 35. Erwin PM, Pineda MC, Webster N, Turon X, Lopez-Legentil S. 2014. Down under the
905 tunic: bacterial biodiversity hotspots and widespread ammonia-oxidizing archaea in coral
906 reef ascidians. ISME Journal 8:575-588.
907 36. Lopez-Legentil S, Turon X, Espluga R, Erwin PM. 2015. Temporal stability of bacterial
908 symbionts in a temperate ascidian. Front Microbiol 6:1022.
909 37. Evans JS, Erwin PM, Shenkar N, Lopez-Legentil S. 2017. Introduced ascidians harbor
910 highly diverse and host-specific symbiotic microbial assemblages. Sci Rep 7:11033.

38. Evans JS, Erwin PM, Shenkar N, Lopez-Legentil S. 2018. A comparison of prokaryotic symbiont communities in nonnative and native ascidians from reef and harbor habitats. FEMS Microbiol Ecol 94:fiy139.
39. Cahill PL, Fidler AE, Hopkins GA, Wood SA. 2016. Geographically conserved microbiomes of four temperate water tunicates. Env Microbiol Rep 8:470-478.
40. Dror H, Novak L, Evans JS, Lopez-Legentil S, Shenkar N. 2019. Core and dynamic microbial communities of two invasive ascidians: can host-symbiont dynamics plasticity affect invasion capacity? Microb Ecol 78:170-184.
41. Erwin PM, Pineda MC, Webster N, Turon X, Lopez-Legentil S. 2013. Small core communities and high variability in bacteria associated with the introduced ascidian *Styela plicata*. Symbiosis 59:35-46.
42. Gonzalez JM, Mayer F, Moran MA, Hodson RE, Whitman WB. 1997. *Microbulbifer hydrolyticus* gen.nov., sp. nov., and *Marinobacterium georgiense* gen.nov.sp.nov., two marine bacteria from a lignin-rich pulp mill waste enrichment community. Int J Syst Evol Microbiol 47:369-376.
43. Little AE, Robinson CJ, Peterson SB, Raffa KF, Handelsman J. 2008. Rules of engagement: interspecies interactions that regulate microbial communities. Annu Rev Microbiol 62:375-401.
44. Imamura N, Nishuma M, Takader T, Adachi K, Sakai M, Sano H. 1997. New anticancer antibiotics pelagiomicins, produced by a new marine bacterium *Pelagiobacter variabilis*. J Antibiot (Tokyo) 50:8-12.
45. Quevrain E, Domart-Coulon I, Pernice M, Bourguet-Kondracki ML. 2009. Novel natural parabens produced by a Microbulbifer bacterium in its calcareous sponge host *Leuconia nivea*. Environ Microbiol 11:1527-1539.
46. Martínez-García M, Stief P, Díaz-Valdés M, Wanner G, Ramos-Esplá A, Dubilier N, Antón J. 2008. Ammonia-oxidizing Crenarchaeota and nitrification inside the tissue of a colonial ascidian. Environ Microbiol 10:2991-3001.
47. Hatzenpichler R. 2012. Diversity, physiology, and niche differentiation of ammonia-oxidizing archaea. Appl Environ Microbiol 78:7501-10.
48. Murray AE, Preston CM, Massana R, Taylor LT, Blakis A, Wu K, DeLong EF. 1998. Seasonal and spatial variability of bacterial and archaeal assemblages in the coastal waters off Anvers Island, Antarctica. Appl Environ Microbiol 64:2585-2595.
49. Murray AE, Grzymski JJ. 2007. Diversity and genomics of Antarctic marine micro-organisms. Philosophical Transactions Of The Royal Society B-Biological Sciences 362:2259-2271.
50. Grzymski JJ, Riesenfeld CS, Williams TJ, Dussaq AM, Ducklow H, Erickson M, Cavicchioli R, Murray AE. 2012. A metagenomic assessment of winter and summer bacterioplankton from Antarctica Peninsula coastal surface waters. ISME J 6:1901-1915.
51. Rhoades D. 1979. Evolution of plant chemical defense against herbivores. Herbivores: their interaction with secondary plant metabolites. Academic Press, New York.
52. McPhail KL, Davies-Coleman MT, Starmer J. 2001. Sequestered chemistry of the Arminacean nudibranch *Leminda millecra* in Algoa Bay, South Africa. J Nat Prod 64:1183-1190.
53. Carbone M, Gavagnin M, Haber M, Guo YW, Fontana A, Manzo E, Genta-Jouve G, Tsoukatou M, Rudman WB, Cimino G, Ghiselin MT, Mollo E. 2013. Packaging and

956 delivery of chemical weapons: a defensive trojan horse stratagem in chromodorid
957 nudibranchs. PLoS One 8:e62075.
958 54. Winters AE, White AM, Dewi AS, Mudianta IW, Wilson NG, Forster LC, Garson MJ,
959 Cheney KL. 2018. Distribution of defensive metabolites in nudibranch molluscs. J Chem
960 Ecol 44:384-396.
961 55. Schupp P, Eder C, Paul V, Proksch P. 1999. Distribution of secondary metabolites in teh
962 sponge Oceanapia sp. and its ecological implications. Mar Biol 135:573-580.
963 56. Freeman CJ, Gleason DF. 2010. Chemical defenses, nutritional quality, and structural
964 components in three sponge species: *Ircinia felix, I. campana*, and *Aplysina fulva*. Mar
965 Biol 157:1083-1093.
966 57. Roue M, Domart-Coulon I, Ereskovsky A, Dejediat C, Pererz T, Bourguet-Kondracki M-
967 L. 2010. Cellular localization of clathridimine, an antimicrobial 2-aminoimidazole
968 alkaloid produced by the Mediterreanean calcareous sponge *Clathrina clathrus*. J Nat
969 Prod 73:1277-1282.
970 58. Furrow FB, Amsler CD, McClintock JB, Baker BJ. 2003. Surface sequestration of
971 chemical feeding deterrents in the Antarctic sponge *Latrunculia apicalis* as an optimal
972 defense against sea star spongivory. Mar Biol 143:443-449.
973 59. Becerro MA, Paul VJ, Starmer J. 1998. Intracolonial variation in chemical defenses of
974 the sponge *Cacospongia sp.* and its consequences on generalist fish predators and the
975 specialist nudibranch predator *Glossodoris pallida*. Mar Ecol Prog Ser 168:187-196.
976 60. Siriak T, Intaraksa N, Kaewsuwan S, Yuenyongsawad S, Suwanborirux K, Plubrukarn A.
977 2011. Intracolonial allocation of tisoxazole macrolides in the sponge *Pachastrissa nux*.
978 Chem Biodivers 8:238-246.
979 61. Richardson AD, Aalbersberg W, Ireland CM. 2008. The patellazoles inhibit protein
980 synthesis at nanomolar concentrations in human colon tumor cells. Anticancer Drugs
981 16:533-541.
982 62. Gouiffes D, Juge M, Grimaud N, Welin L, Sauviat M, Barbin Y, Laurent D, Roussakis C,
983 Henichart J, Verbist J. 1988. Bistramide A, a new toxin from the urochordata *Lissoclinum*
984 *bistratum* Sluiter: isolation and preliminary characterization. Toxicon 26:1129-1136.
985 63. Subramani R, Aalbersberg W. 2013. Culturable rare Actinomycetes: diversity, isolation
986 and marine natural product discovery. Appl Microbiol Biotechnol 97:9291-321.
987 64. Ziemert N, Lechner A, Wietz M, Millan-Aguinaga N, Chavarria KL, Jensen PR. 2014.
988 Diversity and evolution of secondary metabolism in the marine actinomycete genus
989 Salinispora. Proc Natl Acad Sci USA
990 doi:www.pnas.org/cgi/doi/10.1073/pnas.1324161111:E1130-E1139.
991 65. Romano S. 2018. Ecology and Biotechnological Potential of Bacteria Belonging to the
992 Genus *Pseudovibrio*. Appl Environ Microb 84.
993 66. Zhang X, Elliot H, Elliot MA. 2019. Unlocking the trove of metabolic treasures:
994 activating silent biosynthetic gene clusters in bacteria and fungi. Curr Opin Microbiol
995 51:9-15.
996 67. Holmstrom C, James S, Neilan BA, White DC, Kjelleberg S. 1998. *Pseudoalteromonas*
997 *tunicata* sp. nov., a bacterium that produces antifouling agents. Int J Syst Bacteriol
998 48:1205-1212.
999 68. Baker BJ. 2020. *Synoicum adareanum* sampling u/w video Mar 2011 Palmer Station
1000 Antarctica doi:10.5061/dryad.gxd2547gw, Dryad Data.

1001 69. Massana R, Murray AE, Preston CM, DeLong EF. 1997. Vertical distribution and
1002 phylogenetic characterization of marine planktonic Archaea in the Santa Barbara
1003 Channel. Appl Environ Microbiol 63:50-56.
1004 70. Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Struhl K. 1999.
1005 Short protocols in molecular biology, 4th ed. John Wiley and Sons, Inc., New York.
1006 71. Caporaso JG, Lauber C, Walters W, Berg-Lyons D, Lozupone C, Turnbaugh P, Fierer N,
1007 Knight R. 2011. Global patterns of 16S rRNA diversity at a depth of millions of
1008 sequences per sample. Proc Natl Acad Sci USA 108:4516-4522.
1009 72. Bolyen E, Rideout JR, Dillon MR, Bokulich N, Abnet CC, Al-Ghalith GA, Alexander H,
1010 Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn
1011 CJ, Brown CT, Callahan BJ, Caraballo-Rodriguez AM, Chase J, Cope EK, Da Silva R,
1012 Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M,
1013 Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K,
1014 Guo JR, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S,
1015 Jarmusch AK, Jiang LJ, Kaehler BD, Bin Kang K, Keefe CR, Keim P, Kelley ST,
1016 Knights D, et al. 2019. Reproducible, interactive, scalable and extensible microbiome
1017 data science using QIIME 2. Nat Biotechnol 37:852-857.
1018 73. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP. 2016.
1019 DADA2: High-resolution sample inference from Illumina amplicon data. Nat Methods
1020 13:581-+.
1021 74. Tackmann J, Matias Rodrigues JF, von Mering C. 2019. Rapid inference of direct
1022 interactions in large-scale ecological networks from heterogeneous microbial sequencing
1023 data. Cell Syst 9:286-296 e8.
1024 75. Bastian M, Heymann S, Jacomy M. 2009. Gephi: an open source software for exploring
1025 and manipulating networks. International AAAI Conference on Weblogs and Social
1026 Media.
1027 76. Cristóbal E, Ayuso SV, Justel A, Toro M. 2013. Robust optima and tolerance ranges of
1028 biological indicators: a new method to identify sentinels of global warming. Ecol Res
1029 29:55-68.
1030 77. Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network
1031 analysis. BMC Bioinformatics 9:559.
1032 78. Paulson JN, Stine OC, Bravo HC, Pop M. 2013. Differential abundance analysis for
1033 microbial marker-gene surveys. Nat Methods 10:1200-2.
1034 79. Eid J, Fehr A, Gray J, Luong K, Lyle J, Otto G, Peluso P, Rank D, Baybayan P, Bettman
1035 B, Bibillo A, Bjornson K, Chaudhuri B, Christians F, Cicero R, Clark S, Dalal R,
1036 Dewinter A, Dixon J, Foquet M, Gaertner A, Hardenbol P, Heiner C, Hester K, Holden
1037 D, Kearns G, Kong XX, Kuse R, Lacroix Y, Lin S, Lundquist P, Ma CC, Marks P,
1038 Maxham M, Murphy D, Park I, Pham T, Phillips M, Roy J, Sebra R, Shen G, Sorenson J,
1039 Tomaney A, Travers K, Trulson M, Vieceli J, Wegener J, Wu D, Yang A, Zaccarin D, et
1040 al. 2009. Real-Time DNA sequencing from single polymerase molecules. Science
1041 323:133-138.
1042 80. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland
1043 A, Huddleston J, Eichler EE, Turner SW, Korlach J. 2013. Nonhybrid, finished microbial
1044 genome assemblies from long-read SMRT sequencing data. Nature Methods 10:563-569.
1045 81. Markowitz VM, Chen I-MA, Palaniappan K, Chu K, Szeto E, Grechkin Y, Ratner A,
1046 Jacob B, Huang J, Williams P, Huntemann M, Anderson I, Mavromatis K, Ivanova NN,

1047 Kyrpides NC. 2012. IMG: the integrated microbial genomes database and comparative
1048 analysis system. Nucleic Acids Res 40:D115-22.
1049

1050 **TABLES AND FIGURES**

1051

1052 Table 1. Relative proportions (average +/- standard deviations, n=63) of Phyla (and class for the
1053 Proteobacteria) across the different microbiome fractions.

1054

Phlya or Class	Whole Community	Core80	Dynamic50	Variable
Proteobacteria				
Gammaproteobacteria	71.99 ± 6.64	73.28 ± 6.33	51.3 ± 23.16	43.71 ± 23.63
Alphaproteobacteria	22.9 ± 5.39	23.83 ± 5.93		23.83 ± 19.7
Deltaproteobacteria	0.17 ± 0.1	0.16 ± 0.1		1.11 ± 2.00
Bacteroidetes	2.83 ± 2.14	0.79 ± 0.69	46.55 ± 22.91	17.4 ± 14.69
Verrucomicrobia	1.56 ± 2.8	1.59 ± 2.93		2.83 ± 3.97
Nitrospirae	0.27 ± 0.23	0.29 ± 0.24		0.02 ± 0.17
Planctomycetes	0.12 ± 0.13		2.15 ± 3.06	4.6 ± 12.84
Actinobacteria	0.1 ± 0.08	0.05 ± 0.05		3.15 ± 3.53
Patescibacteria	0.02 ± 0.09			0.84 ± 2.05
Dadabacteria	0.02 ± 0.03			1.22 ± 2.03
Uncl. Bacteria	0.009 ± 0.017			0.72 ± 1.53
Dependentiae	0.004 ± 0.018			0.34 ± 1.99
Chlamydiae	0.002 ± 0.006			0.19 ± 0.83
Acidobacteria	0.000 ± 0.003			0.02 ± 0.13
Chloroflexi	0.000 ± 0.003			0.02 ± 0.13
Epsilonbacteraeota	0.000 ± 0.001			0.01 ± 0.07

1055

1056 Table 2. Taxonomic affiliations of core microbiome, relative abundance rank, and potential of
1057 affiliated genus in natural product gene cluster biosynthesis. Taxonomy is shown according to
1058 genome taxonomy database (GTDB) classification and NCBI taxonomy is included
1059 (GTDB/NCBI) where they differ. Biosynthetic potential only calculated for ASVs with Genus-
1060 level taxonomic assignments was based on representative genome biosynthetic gene cluster
1061 content in the same genus (See Fig. S3 for list of genomes). ASVs in bold ranked in the top 10.
1062 Where more than one ASV was found per genus, the average relative abundance and standard
1063 deviations were summed. n=63 individuals.
1064

ASV_ID	Phylum, highest taxonomic assignment	Average Relative abundance (%)	Rank	Nearest neighbor % identity	NRP BGC	PKS BGC	Combined NRP-PKS
SaM_ASV1, 2, 4, 5, 10, 17, 18	**Proteobacteria, Microbulbifer**	**77.54 ± 21.86**	**1, 2, 4, 5, 8, 9, 10**	**97.42**	**+**	**+**	
SaM_ASV7	Proteobacteria, Endozoicomonas	0.47 ± 0.51	13	96.71	+	+	
SaM_ASV13	Proteobacteria, Nitrosomonas	0.46 ± 0.35	14	99.77	+	+	+
SaM_ASV3	**Proteobacteria, Pseudovibrio**	**19.92 ± 4.74**	**3**	**98.75**	**+**	**+**	**+**
SaM_ASV6	**Proteobacteria, Hoeflea**	**1.59 ± 1.36**	**6**	**99.25**	**+**	**+**	**+**
SaM_ASV16	Proteobacteria, Halocynthiibacter	0.63 ± 0.64	11	99.75	+		
SaM_ASV11	Nitrospirota/ Nitrospirae, Nitrospira	0.27 ± 0.23	15	98.32	+	+	+
SaM_ASV12	Bacteroidota/ Bacteroidetes, Lutibacter	0.50 ± 0.55	12	94.54	+		+
SaM_ASV14	Verrucomicrobiota/ Verrucomicrobia, Lentimonas	0.16 ± 0.22	19	99.77	+		+
Sam_ASV21	Proteobacteria, Rhodobacteraceae	0.18 ± 0.26	18	98.75			
SaM_ASV8	Proteobacteria, Rhodospirillales	0.22 ± 0.22	17	86.60			
SaM_ASV9	Bdellovibrionota/ Proteobacteria, Bdellovibrionaceae	0.15 ± 0.09	20	90.35			
SaM_ASV19	Bacteroidetes, Cryomorphaceae	0.24 ± 0.23	16	89.10			
SaM_ASV15	**Verrucomicrobiota/ Verrucomicrobia, Opitutaceae**	**1.34 ± 2.77**	**7**	**90.14**			
SaM_ASV20	Actinobacteria, Solirubrobacterales	0.05 ± 0.05	21	91.80			

1065

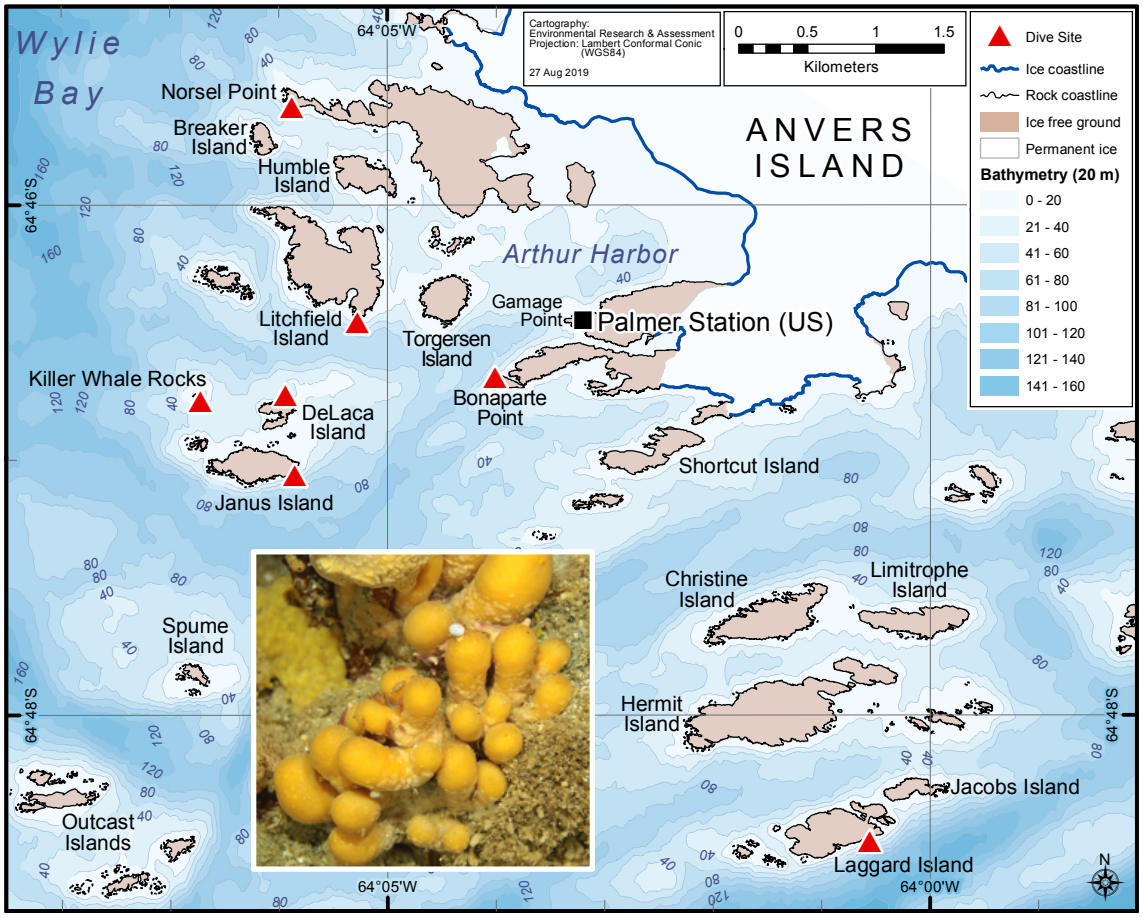


Fig. 1. Bathymetric map of the study area off Anvers Island. *Synoicum adareanum* collection sites are shown with red triangles. The map was generated by Environmental Research & Assessment, Cambridge, UK, using Arthur Harbor bathymetry data from the PRIMO survey project 2004-06 (Dr. Scott Gallagher and Dr. Vernon Asper). Inset: Colonial ascidian, *S. adareanum*, which occurs in clusters of multiple lobes connected by a peduncle which together comprise a colony on the seafloor, collected at depths ranging from 24-31 m. Sampling site abbreviations in text: Bonaparte Point (Bon); DeLaca Island (Del); Janus Island (Jan); Killer Whale Rocks (Kil); Laggard Island (Lag); Litchfield (Lit); Norsel Point (Nor).

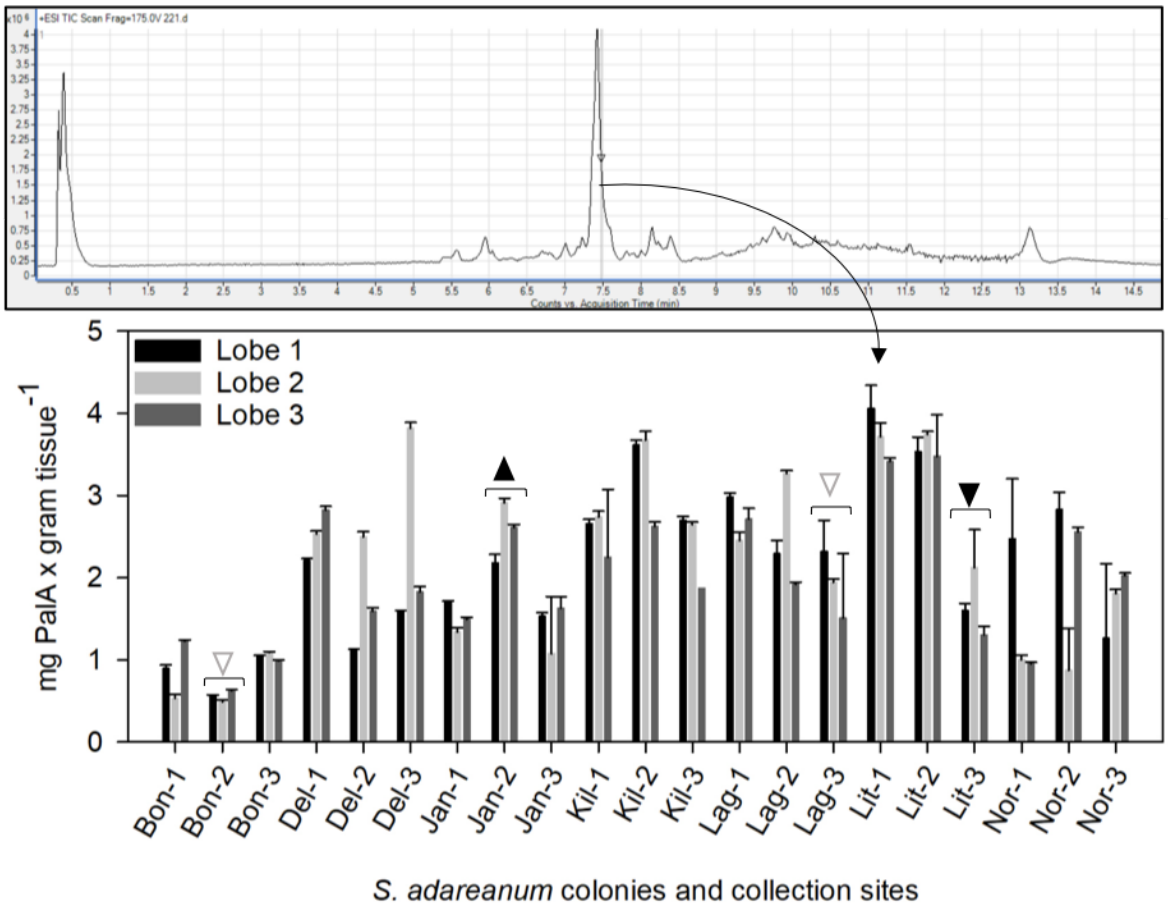


Fig. 2. Palmerolide A (PalA) detection in *S. adareanum* holobionts. (a) Mass spectra derived from sample Lit-1a. The PalA peak dominates the dichloromethane-methanol fraction of the *S. adareanum* extract. (b) Levels of PalA normalized to tissue dry weight detected by mass spectrometry in *S. adareanum* holobiont tissues (three lobes per colony) surveyed in three colonies per site across the Anvers Island archipelago. Error represents individual lobe technical replication (standard deviation). Colonies with significant differences in PalA levels within a site are indicted with triangles, in which the direction of point indicates a significantly higher or lower colony. Filled triangles = significance ($p < 0.05$) in comparison to the other two colonies, and open triangles are those that were different from only one of the two colonies. Most colonies had significant lobe-lobe differences in PalA concentration, and some site-level differences were observed (Fig. S2).

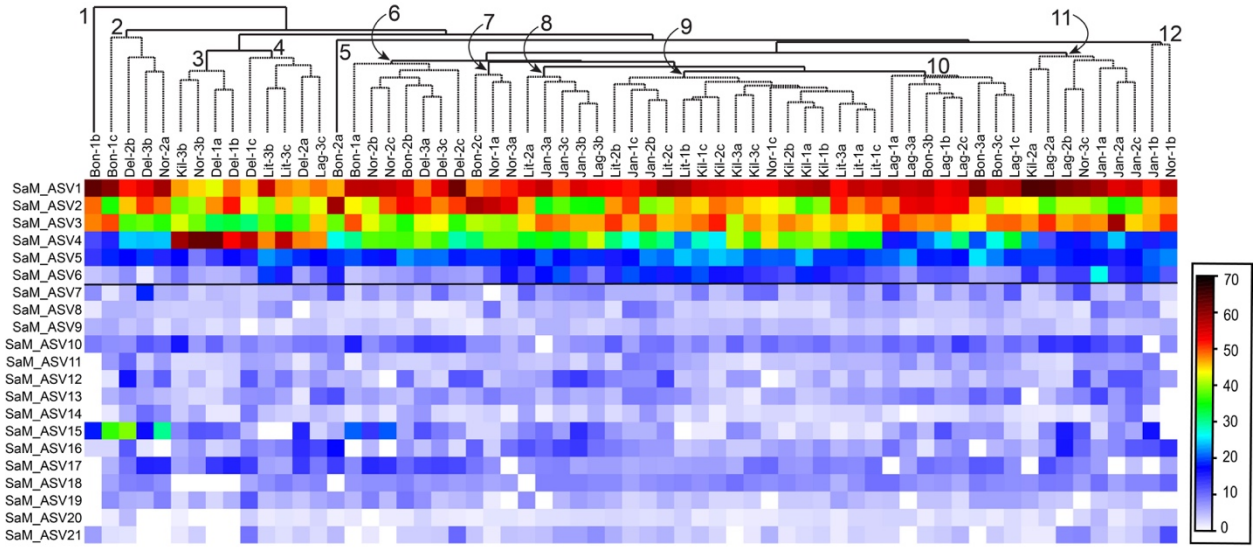


Fig. 3. Heatmap of square root transformed ASV occurrence data for the core microbiome. ASVs (ranked and numbered) are shown on the y-axis, and 63-samples were hierarchically clustered, shown on the y-axis (site-based; square root transformed abundance data). Nodes with significant clusters are indicated from left to right ($p < 0.05$); order of clustering inside the node was not significant. The horizontal line drawn below SaM_ASV6 demarcates those ASVs that were present in all 63 samples.

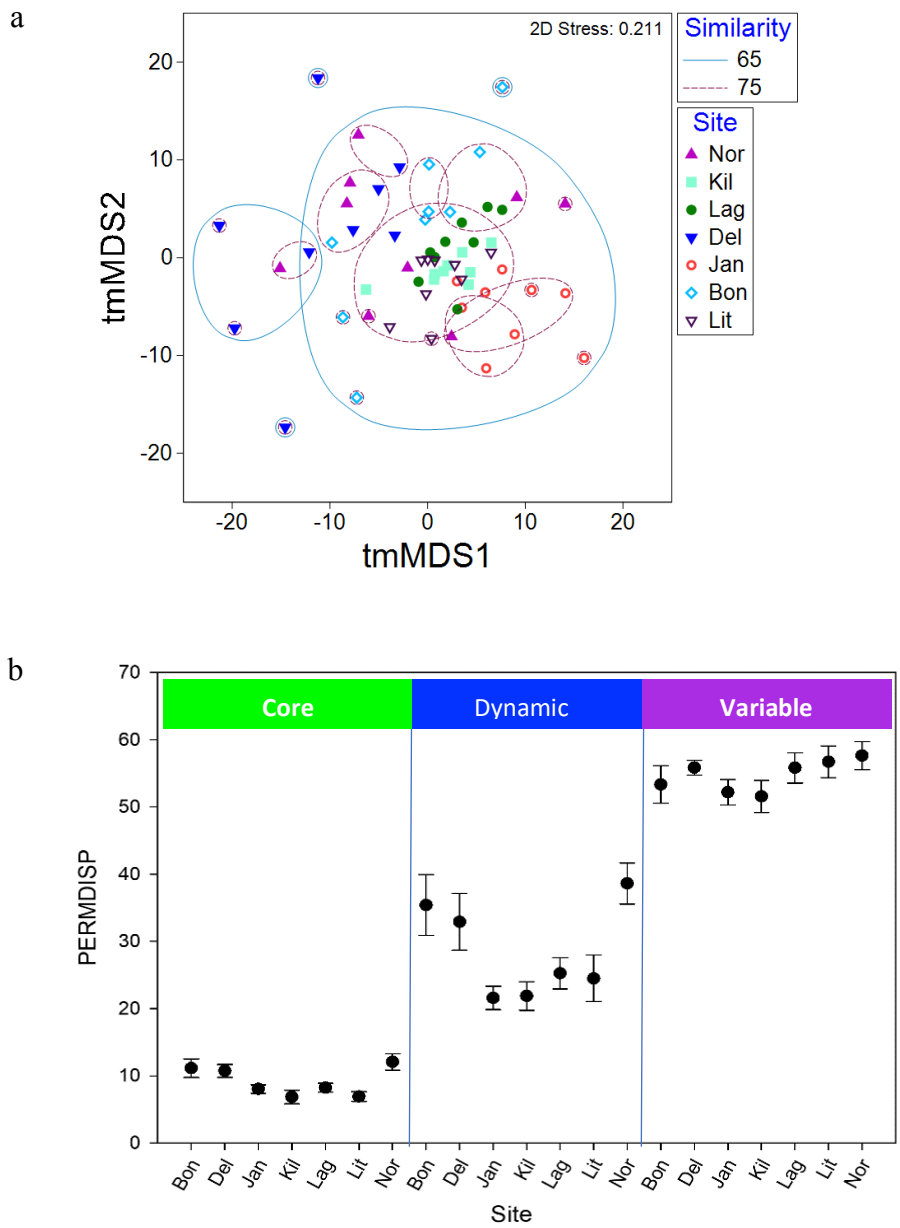


Fig. 4. Similarity relationships amongst the *S. adareanum* microbiome samples in the Anvers Island archipelago. (a) tmMDS of Bray-Curtis similarities of square root transformed ASV occurrence data representing the microbiome of the 63 *S. adareanum* lobe samples using the complete ASV occurrence profiles. Microbiome samples with significant levels of similarity are shown (see legend). (b) Beta-diversity across Anvers Island archipelago sites represented by PERMDISP (9999 permutations) reveals differences between the highly persistent core, dynamic and variable portions of the *S. adareanum* microbiome (standard error shown). The degree of dispersion (variance) around the centroid changes significantly ($p < 0.0001$) for the different microbiome classifications, which the lowest levels of dispersion are found in the core microbiome.

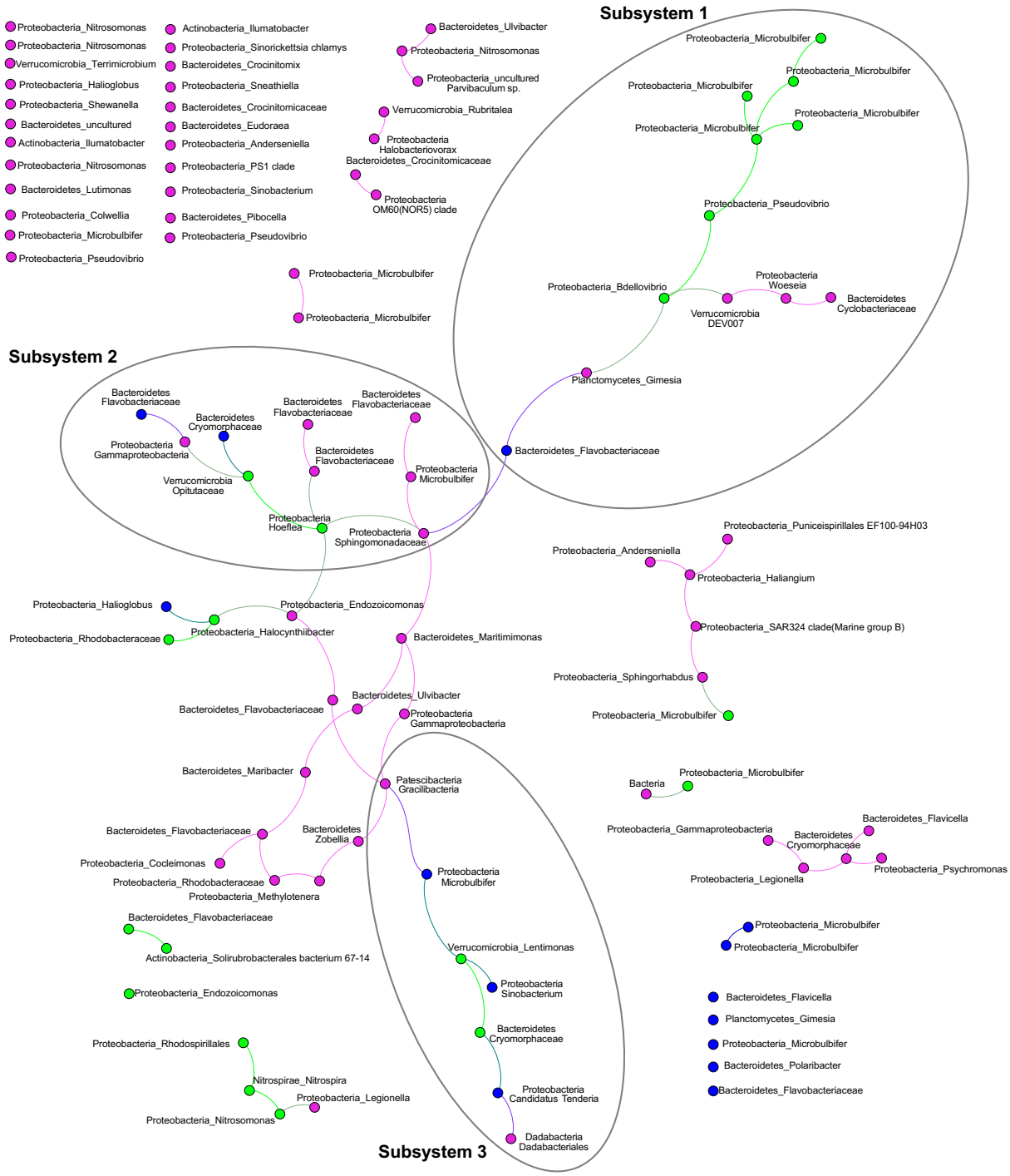


Fig. 5. ASV co-occurrence network. The largest connected component of the co-occurrence network (seeded with ASVs found in at least 5 samples, 102 in total) identified three subsystems. Node colors represent the microbiome fractions (Core80, green; Dynamic50, blue; Variable, pink). Taxonomic identities of the ASVs are shown next to the nodes, with the phylum_highest level taxon identified shown.

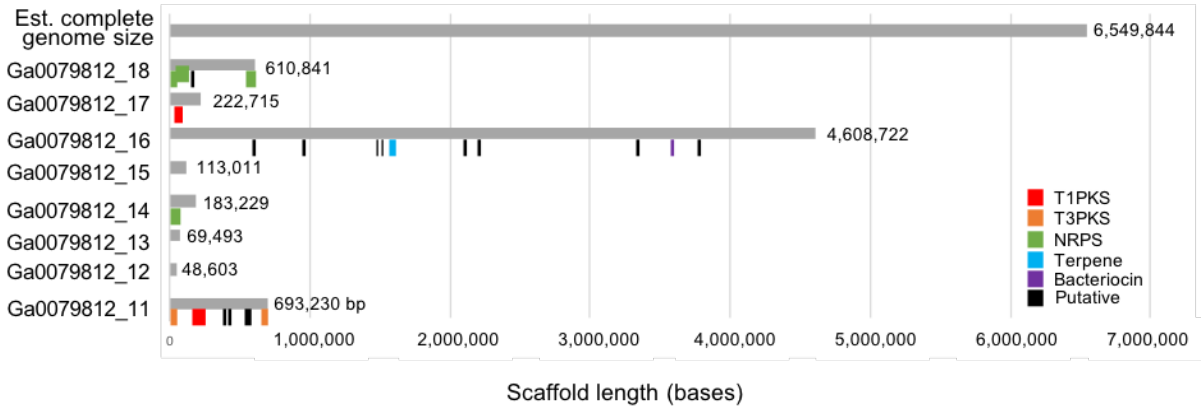


Fig. 6. *Pseudovibrio* sp. TunPSC04-5.I4 genome map showing relative positions of biosynthetic gene clusters (BGC) on genome scaffolds. The IMG contig ID is listed to the left of the contigs, this corresponds to scaffolds in GenBank FNLB01000001.1 - FNLB01000008.1. Scaffold lengths are listed at the right of the grey bars. The estimated genome size is shown while the contig order is not currently known. BGC structural categories are identified by color shown in the key.

1159 **SUPPLEMENTAL MATERIALS**

1160

1161 Table S1. PERMANOVA allows the components of this nested experimental design, site and
1162 colony, to be evaluated as drivers of variability. Note, all values presented in the table are
1163 statistically significant (P(perm)< 0.05).

1164

	All ASVs (%)	Core (%)	Dynamic (%)	Variable (%)
Site-to-Site	25.0	26.2	28.4	19.2
Colony-to-Colony	27.8	31.5	27.2	22.3
Residual	47.2	42.3	44.4	58.4

1165

1166 Table S2. Taxonomic distribution of bacterioplankton ASVs. Phyla and assigned classes are
1167 indicated, along with average representation across the 15-sample data set. Average values below
1168 0.005 are represented as 0.00%.
1169

Phylum_Class	Average Representation
Acidobacteria_Acidobacteriia	0.00%
Acidobacteria_Subgroup 26	0.00%
Acidobacteria_Subgroup 6	0.01%
Actinobacteria_Acidimicrobiia	0.96%
Actinobacteria_Actinobacteria	0.05%
AncK6_Unclassified	0.01%
Atribacteria_JS1	0.00%
Bacteroidetes_Bacteroidia	13.10%
Chlamydiae	0.03%
Chloroflexi_Dehalococcoidia	0.49%
Chloroflexi_JG30-KF-CM66	0.03%
Oxyphotobacteria/Plastid	7.39%
Dadabacteria_Dadabacteriia	0.06%
Dependentiae_Babeliae	0.00%
Epsilonbacteraeota_Campylobacteria	0.32%
Fibrobacteres_Chitinivibrionia	0.00%
Fibrobacteres_Fibrobacteria	0.00%
Firmicutes_Bacilli	0.03%
Firmicutes_Clostridia	0.01%
Fusobacteria_Fusobacteriia	0.03%
Gemmatimonadetes_BD2-11 terrestrial group	0.12%
Hydrogenedentes_Hydrogenedentia	0.00%
Kiritimatiellaeota_Kiritimatiellae	0.00%
Latescibacteria_Latescibacteria	0.00%
Lentisphaerae_Lentisphaeria	0.01%
Lentisphaerae_Oligosphaeria	0.00%
Lentisphaerae_Unclassified	0.00%
Margulisbacteria_Unclassified	0.04%
Marinimicrobia (SAR406 clade)_Unclassified	2.72%
Nitrospinae_Nitrospinia	0.31%
Nitrospinae_P9X2b3D02	0.00%
Nitrospinae_Nitrospira	0.00%
Unclassified_Omnitrophicaeota	0.00%
Patescibacteria_Gracilibacteria	0.03%

Patescibacteria_Parcubacteria	0.00%
PAUC34f_Unclassified	0.13%
Planctomycetes_BD7-11	0.01%
Planctomycetes_Brocadiae	0.00%
Planctomycetes_OM190	0.03%
Planctomycetes_Phycisphaerae	0.10%
Planctomycetes_Pla3 lineage	0.01%
Planctomycetes_Pla4 lineage	0.00%
Planctomycetes_Planctomycetacia	0.15%
Planctomycetes_SGST604	0.00%
Planktomycetes_unclassified	0.00%
Poribacteria_Unclassified	0.00%
Proteobacteria_Alphaproteobacteria	15.57%
Proteobacteria_Deltaproteobacteria	2.81%
Proteobacteria_Gammaproteobacteria	47.35%
Proteobacteria_Unclassified	0.00%
Stramenopiles	0.00%
Spirochaetes_Spirochaetia	0.00%
Tenericutes_Mollicutes	0.00%
Verrucomicrobia_Verrucomicrobiae	0.28%
Unclassified_Bacteria	0.07%
Thaumarchaeota_Nitrososphaeria	6.29%
Euryarchaeota_Halobacteria	0.00%
Euryarchaeota_Thermoplasmata	1.39%
Nanoarchaeota_Woesearchaeia	0.00%

1171 Table S3. Bacterial genomes related to Core80 SaM genera with selected biosynthetic gene
1172 clusters identified using antiSMASH. Genome sequence identifiers are indicated (initially
1173 identified in the Integrated Microbial Genomes database, IMG) in addition to NCBI accession
1174 numbers, where possible, along with the number of gene clusters identified that were related to
1175 polyketide biosynthesis (PKS) or non-ribosomal peptide synthesis (NRPS).

	IMG taxon_oid	NCBI BioSample Accession	NCBI Bioproject Accession	NCBI assembly Accession	Genome Name	T1PKS	T3PKS	transAT-PKS NRPS	T1PKS T3PKS and T3PKS T1PKS	T1PKS bacteriocin	NRPS	NRPS-Like	NRPS-T1PKS	T1PKS NRPS-like	NRPS T1PKS terpene	NRPS T1PKS bacteriocin
Endozoicomonas	2839589109	SAMN08174373	PRJNA422318	GCA_002864045.1	*Endozoicomonas acroporae* Acr-14											
	2651870311	SAMN03407264	PRJNA279233	GCA_001562015.1	*Endozoicomonas arenosclerae* Ab112							1				
	2651870310	SAMN03444428	PRJNA279233	GCA_001562005.1	*Endozoicomonas arenosclerae* E-MC227											
	2627853506	SAMN03956932	PRJNA291958	GCA_001646945.1	*Endozoicomonas ascidiicola* AVMART05						1					
	2775507270	SAMN03978955	PRJNA292506	GCA_001647025.2	*Endozoicomonas atrinae* WP70											
	2574179787	SAMN02851610	PRJNA252578	GCA_000710775.1	*Endozoicomonas elysicola* MKT110											
	2515154160	SAMN02440436	PRJNA169813	GCA_000373945.1	*Endozoicomonas elysicola* MKT110, DSM 22380											
	2687453173	SAMN04155633	PRJNA66389	GCA_001583435.1	*Endozoicomonas montiporae* CL-33											
	2574179789	SAMN02851611	PRJNA252578	GCA_000722565.1	*Endozoicomonas montiporae* LMG 24815											
	2574179790	SAMN02851612	PRJNA252578	GCA_000722635.1	*Endozoicomonas numazuensis* DSM 25634											
	2681813502				putative Gammaproteobacteria family (bin U_52270)		1									
Endozoicomonas-sum					**Genome count: 11**	**0**	**1**	**0**	**0**	**0**	**1**	**1**	**0**	**0**	**0**	**0**
Nitrosomonas	2737471630	SAMN08777538	PRJNA444064	GCA_003046585.1	*Nitrosomonas aestuarii* Nm36											
	2675903044	SAMN05421882	PRJNA330477	GCA_900106545.1	*Nitrosomonas communis* Nm110							1	1			
	2588253501	SAMN02743940	PRJNA185244	GCA_900143275.1	*Nitrosomonas cryotolerans* ATCC 49181							1				
	637000195	SAMEA3138193	PRJNA52	GCA_000009145.1	*Nitrosomonas europaea* ATCC 19718	1										
	637000196	SAMN02598323	PRJNA13913	GCA_000014765.1	*Nitrosomonas eutropha* C91	1										
	2675903041	SAMN05421881	PRJNA330480	GCA_900107165.1	*Nitrosomonas halophila* Nm1						1	2	1			
	2671180082	SAMN05216325	PRJNA323340	GCA_900110145.1	*Nitrosomonas marina* Nm22											
	2834530931	SAMEA4505502	PRJEB15545	GCA_900103035.1	*Nitrosomonas mobilis* ERS1404681											
	2671180218	SAMN05421880	PRJNA330481	GCA_900114795.1	*Nitrosomonas nitrosa* Nm146						1	1	1			
	2737471629	SAMN08777537	PRJNA444079	GCA_003050805.1	*Nitrosomonas oligotropha* Nm49	1							1			
	2675903042	SAMN05421755	PRJNA329788	GCA_900107265.1	*Nitrosomonas* sp. 33 Nm33						1				1	
	2675903038	SAMN05428952	PRJNA329786	GCA_900100485.1	*Nitrosomonas* sp. 41 Nm132							1			1	
	2675903037	SAMN05421690	PRJNA329790	GCA_900111165.1	*Nitrosomonas* sp. 51 Nm51	1						1				
	650716066	SAMN00002570	PRJNA32989	GCA_000175095.2	*Nitrosomonas* sp. AL212											
	2834527210	SAMN08707727	PRJNA438189	GCA_003011925.2	*Nitrosomonas* sp. APG5								1			
	2747843217				*Nitrosomonas* sp. HKU-PRO10						2					
	650716067	SAMN02232044	PRJNA52837	GCA_000219585.1	*Nitrosomonas* sp. IS79A3	1										
	2724679654	SAMN06298227	PRJNA370067	GCA_003664085.1	*Nitrosomonas* sp. Nm120	1							1			
	2671180029	SAMN05216406	PRJNA323344	GCA_900105875.1	*Nitrosomonas ureae* Nm10							1	1			
Nitrosomonas-sum					**Genome count: 19**	**6**	**0**	**0**	**0**	**0**	**5**	**8**	**7**	**0**	**2**	**0**
Halocynthilibacter	2775507242	SAMN03252591	PRJNA269208	GCA_000812665.2	*Halocynthiibacter arcticus* PAMC 20958							1				
	2636415494	SAMN03254441	PRJNA269555	GCA_000812685.1	*Halocynthiibacter namhaensis* RA2-3											
Halocynthilibacter-sum					**Genome count: 2**	**0**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**	**0**
Nitrospira	2684623072	SAMEA3560316	PRJEB10818	GCA_001458695.1	*Candidatus* Nitrospira inopinata ENR4							1				
	2690315930	SAMEA3638996	PRJEB11446	GCA_001458775.1	*Candidatus* Nitrospira nitrificans COMA2								1			1
	2690315890	SAMEA3638995	PRJEB11445	GCA_001458735.1	*Candidatus* Nitrospira nitrosa COMA1	1							1			
	2654587964	SAMN06268482	PRJNA366768	GCA_001303205.1	*Nitrospira* bacterium SG8_3							1				
	2654587963	SAMN06268482	PRJNA366768	GCA_001303305.1	*Nitrospira* bacterium SG8_35_1							1				
	2654587962	SAMN06268482	PRJNA366768	GCA_001303405.1	*Nitrospira* bacterium SG8_35_4											
	2654587939	SAMN06268417	PRJNA366184	GCA_001303745.1	*Nitrospira* bacterium SM23_35							1				
	649633030	SAMEA2272290	PRJEA46433	GCA_000196815.1	*Nitrospira defluvii*						1		1			
	2757321055	SAMEA17758918	PRJEB18128	GCA_900169565.1	*Nitrospira japonica* NJ11			1			1		1			
	2833481757	SAMEA4623285	PRJEB26290	GCA_900403705.1	*Nitrospira lenta* N. lenta BS10						1					
	2596583682	SAMN05660474	PRJNA262287		*Nitrospira marina* Nb-295							1				
	2639762945	SAMN03702441	PRJNA283178	GCA_001273775.1	*Nitrospira moscoviensis* NSP M-1						1	1				
	2833524137	SAMEA17758168	PRJEB18127	GCA_900170025.1	*Nitrospira* sp. ND1						2		1			
	2724679713				*Nitrospira* sp. QS2205	1					15	1				
Nitrospira-sum					**Genome count: 14**	**2**	**0**	**1**	**0**	**0**	**21**	**7**	**5**	**0**	**0**	**1**
Lutibacter	2724679779	SAMN06265371	PRJNA363559	GCA_900188235.1	*Lutibacter agarilyticus* DSM 29150											
	2622736608	SAMN04488111	PRJNA303615	GCA_900188355.1	*Lutibacter flavus* DSM 27993											
	2622736611	SAMN04488006	PRJNA303385	GCA_900116115.1	*Lutibacter maritimus* DSM 24450											
	2710264748	SAMN06265796	PRJNA364304	GCA_003384935.1	*Lutibacter oceani* 325-5											
	2837442332	SAMN09780835	PRJNA485024	GCA_003426875.1	*Lutibacter oceani* JCM 30924											
	2693429896	SAMN05444411	PRJNA331369	GCA_900106565.1	*Lutibacter oricola* DSM 24956											
	2663763042	SAMN04299808	PRJNA304382	GCA_001543325.1	*Lutibacter profundi* LP1											
	2839499322	SAMN08813454	PRJNA447935	GCA_003260195.1	*Lutibacter* sp. 1KV19											
	2597490376	SAMN03159350	PRJNA265337	GCA_007827455.1	*Lutibacter* sp. Hel_I_33_5											
	2718218033	SAMN05726020	PRJNA341557	GCA_001761325.1	*Lutibacter* sp. LPB0138											
	2823967136	SAMN12024057	PRJNA546748		*Lutibacter* sp. SG786								1			
	2839507965	SAMN09759694	PRJNA484105	GCA_003449015.1	*Lutibacter* sp. SM1352		1									
Lutibacter-sum					**Genome count: 12**	**0**	**1**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**0**	**0**
Puniceicoccaceae	2786546861	SAMEA2622923		GCA_002689065.1	*Puniceicoccaceae* bacterium ARS1008											

	IMG taxon_oid	NCBI BioSample Accession	NCBI Bioproject Accession	NCBI assembly Accession	Genome Name	T1PKS	T3PKS	transAT-PKS NRPS	T1PKS T3PKS and T3PKS T1PKS	T1PKS bacteriocin	NRPS	NRPS-Like	NRPS-T1PKS	T1PKS NRPS-like	NRPS T1PKS terpene	NRPS T1PKS bacteriocin
	2786546862	SAMEA2622923		GCA_002690565.1	*Puniceicoccaceae* bacterium MED739											
	2786546866	SAMEA2620855		GCA_002704545.1	*Puniceicoccaceae* bacterium MED785											
	2786546863	SAMEA2620230		GCA_002704415.1	*Puniceicoccaceae* bacterium MED802											
	2786546868	SAMEA2620929		GCA_002703145.1	*Puniceicoccaceae* bacterium MED840											
	2786546867	SAMEA2620929		GCA_002729725.1	*Puniceicoccaceae* bacterium NP23											
	2786546865	SAMEA2621812		GCA_002722335.1	*Puniceicoccaceae* bacterium SP237	**1**										
	2786546864	SAMEA2621812		GCA_002717975.1	*Puniceicoccaceae* bacterium SP4032											
Puniceicoccaceae -sum					**Genome count: 8**	**1**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**	**0**
Hoeflea	2740891889	SAMN05877838	PRJNA346220	GCA_900220985.1	*Hoeflea halophila* KCTC 23107											
	2770939490	SAMN09074944	PRJNA463379	GCA_003182275.1	*Hoeflea marina* DSM 16791											
	2663763045	SAMN02436119	PRJNA19311	GCA_000154705.2	*Hoeflea phototrophica* DFL-43							1				
	2517434009	SAMN02441719	PRJNA169687	GCA_000372965.1	*Hoeflea* sp. 108									3		
	2617271323	SAMN03069733	PRJNA261244	GCA_000759435.1	*Hoeflea* sp. BAL378											
	2654588011	SAMN03375744	PRJNA276494	GCA_001011155.1	*Hoeflea* sp. IMCC20628											
	2788500370	SAMEA2620929		GCA_002707605.1	*Hoeflea* sp. SAT27	1										
Hoeflea-sum					**Genome count: 7**	**1**	**0**	**0**	**0**	**0**	**0**	**1**	**0**	**3**	**0**	**0**
Pseudovibrio	2615840717	SAMN04488518	PRJNA303594	GCA_900114245.1	*Pseudovibrio ascidiaceicola* DSM 16392				3							
	2744054490	SAMN04147116	PRJNA298121	GCA_001623255.1	*Pseudovibrio axinellae* Ad2		1	2	1		1					
	2684622848	SAMN05421798	PRJNA330493	GCA_900110875.1	*Pseudovibrio axinellae* DSM 24994			2			1					
	2693429873	SAMN05444141	PRJNA331415	GCA_900116575.1	*Pseudovibrio denitrificans* DSM 17465					2	1		1			
	2681812989	SAMD00000429	PRJDB661	GCA_001310815.1	*Pseudovibrio denitrificans* JCM 12308				6	2			1			
	2744054961	SAMN04192297	PRJNA299034	GCA_001561995.1	*Pseudovibrio hongkongensis* UST20140214-015B											
	2841905807	SAMN05729687	PRJNA341851	GCA_001941685.1	*Pseudovibrio* sp. Ab134				3		1					
	2744054491	SAMN04147117	PRJNA298121	GCA_001623225.1	*Pseudovibrio* sp. Ad13				3		1					
	2744054492	SAMN04147118	PRJNA298121	GCA_001623245.1	*Pseudovibrio* sp. Ad14				3		1					
	2744054493	SAMN04147119	PRJNA298121	GCA_001623285.1	*Pseudovibrio* sp. Ad26				3		1					
	2744054494	SAMN04147120	PRJNA298121	GCA_001623075.1	*Pseudovibrio* sp. Ad37				3		1					
	2744054495	SAMN04147121	PRJNA298121	GCA_001623065.1	*Pseudovibrio* sp. Ad46				3		2		1			
	2744054496	SAMN04147122	PRJNA298121	GCA_001623095.1	*Pseudovibrio* sp. Ad5				3		1					
	2841887973	SAMEA19232668	PRJEB18465	GCA_900143565.1	*Pseudovibrio* sp. Alg231-02				3		1					
	2606217185			GCA_900312785.1	*Pseudovibrio* sp. AU243				6							
	2511231065		PRJNA73563	GCA_000236645.1	*Pseudovibrio* sp. FO-BEG1				3				1			
	2728369655	SAMD00000393	PRJDB1714	GCA_001310715.1	*Pseudovibrio* sp. JCM 19062	1	1				2	1				
	647533200	SAMN02436269	PRJNA19345	GCA_000156235.1	*Pseudovibrio* sp. JE062				3							
	2651869879	SAMN03217285	PRJNA268231	GCA_001431305.1	*Pseudovibrio* sp. POLY-S9				3		3		1			
	2636416046	SAMN04515695	PRJNA311457	GCA_900104145.1	*Pseudovibrio* sp. Tun.PSC04-5.I4	2	2				4					
	2744054497	SAMN04147123	PRJNA298121	GCA_001623145.1	*Pseudovibrio* sp. W64				3		1					
	2744054498	SAMN04147124	PRJNA298121	GCA_001623085.1	*Pseudovibrio* sp. W74				3		1					
	2744054499	SAMN04147125	PRJNA298121	GCA_001623165.1	*Pseudovibrio* sp. WM33				3		1					
	2791355020	SAMN04192298	PRJNA299035	GCA_001562055.1	*Pseudovibrio stylochi* UST20140214-052											
Pseudovibrio-sum					**Genome count: 24**	**3**	**4**	**4**	**58**	**4**	**24**	**1**	**5**	**0**	**0**	**0**
Microbulbifer	2772190780	SAMN06311449	PRJNA371746	GCA_001999945.1	*Microbulbifer agarilyticus* GP101						1	1				
	2547132385	SAMN02470156	PRJNA67843	GCA_000220505.2	*Microbulbifer agarilyticus* S89						1					
	2751185758	SAMN04334609	PRJNA305828	GCA_001750105.1	*Microbulbifer aggregans* CCB-MM1						1					
	2619618825	SAMN04487965	PRJNA303547	GCA_900129095.1	*Microbulbifer donghaiensis* CGMCC 1.7063											
	2791355005	SAMN05214359	PRJNA324768	GCA_002009015.1	*Microbulbifer mangrovi* DD-13											
	2667527380	SAMN05216562	PRJNA323126	GCA_900107725.1	*Microbulbifer marinus* CGMCC 1.10657											
	2834404499	SAMN08463771	PRJNA432855	GCA_002959965.1	*Microbulbifer pacificus* LD25											
	2824280218	SAMN12025179	PRJNA546710		*Microbulbifer rhizosphaerae* CECT 8799	1					1					
	2609459990	SAMN02641169	PRJNA236345	GCA_000708675.1	*Microbulbifer* sp. HZ11											
	2690315924	SAMN04349613	PRJNA294922	GCA_001639145.1	*Microbulbifer* sp. Q7							2				
	2744054596	SAMN04419353	PRJNA309016	GCA_001641755.2	*Microbulbifer* sp. WRN-8						1					
	2728369117	SAMN04376035	PRJNA307149	GCA_001507605.1	*Microbulbifer* sp. ZGT114						1					
	2687453240	SAMN04556710	PRJNA315367	GCA_001617625.1	*Microbulbifer thermotolerans* DAU221						2					
	2599185213	SAMN05660479	PRJNA262321	GCA_900112305.1	*Microbulbifer thermotolerans* DSM 19189						2					
	2519899636	SAMN02441043	PRJNA185297	GCA_000380565.1	*Microbulbifer variabilis* ATCC 700307						1	3				
	2667528184	SAMN05216212	PRJNA323127	GCA_900100355.1	*Microbulbifer yueqingensis* CGMCC 1.10658											
Microbulbifer-sum					**Genome count: 16**	**1**	**0**	**0**	**0**	**0**	**11**	**6**	**0**	**0**	**0**	**0**

1178 Table S4. *Synoicum adareanum* collections and preparations for microbiome cultivation.
1179 Temperature was estimated using dive computer.
1180

S. adareanum sample ID	Collection Site	GPS Location	Date	Depth (m)	Temperature (°C)
PSC04	Norsel Point	S 64° 45.638' W 64° 05.874'	2004	28.3	n.d.
P7-AM17	Bonaparte Point	S 64° 46.662' W 64° 03.986'	2-May-07	23.4	-1.1

1181

Table S5. *Synoicum adareanum* collections for palmerolide A and microbiome characterization by V3-4 rRNA gene tag sequencing. Temperature was estimated using dive computer. Qualitative analysis of dive videos was used to determine benthic cover at dive sites.

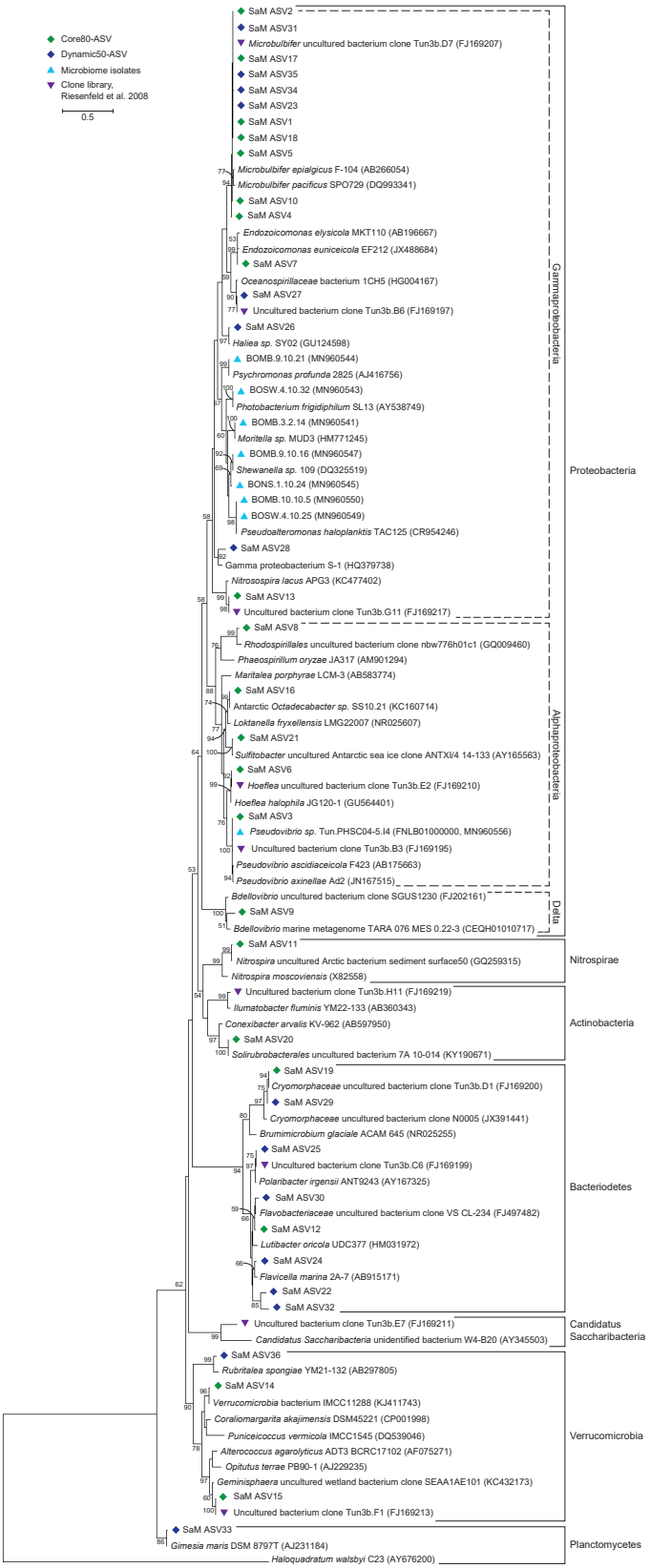
Sample	Dive Site	Collection Date	GPS Location	Depth (m)	Water Temp (°C)	Presence of Red Algae	Presence of Brown Algae	Presence of Sponge	Presence of Bryozoan
Bon-1	Bonaparte Point	24-Mar-11	S 64° 46.679', W 64° 04.013'	26.2	2	+	+	+	+
Bon-2	Bonaparte Point	24-Mar-11	S 64° 46.679', W 64° 04.013'	25.3	1	+	+	+	+
Bon-3	Bonaparte Point	24-Mar-11	S 64° 46.679', W 64° 04.013'	26.5	1	+	--	--	+
Del-1	DeLaca Island	28-Mar-11	S 64° 46.749', W 64° 05.945'	25.0	1	--	+	--	--
Del-2	DeLaca Island	28-Mar-11	S 64° 46.749', W 64° 05.945'	22.0	1	--	+	--	+
Del-3	DeLaca Island	28-Mar-11	S 64° 46.749', W 64° 05.945'	22.9	1	+	+	+	--
Jan-1	Janus Island	23-Mar-11	S 64° 47.062', W 64° 05.857	24.1	1	+	+	--	--
Jan-2	Janus Island	23-Mar-11	S 64° 47.062', W 64° 05.857	24.1	1	+	+	+	--
Jan-3	Janus Island	23-Mar-11	S 64° 47.062', W 64° 05.857	24.1	1	+	+	+	--

Kil-1	Killer Whale Rocks	2-Apr-11	S 64° 46.771', W 64° 06.725'	26.5	1	+	+	+	+
Kil-2	Killer Whale Rocks	2-Apr-11	S 64° 46.771', W 64° 06.725'	25.6	1	+	+	+	--
Kil-3	Killer Whale Rocks	2-Apr-11	S 64° 46.771', W 64° 06.725'	25.3	1	+	+	+	--
Lag-1	Laggard Island	31-Mar-11	S 64° 48.497', W 64° 00.564'	26.2	1	+	+	--	+
Lag-2	Laggard Island	31-Mar-11	S 64° 48.497', W 64° 00.564'	24.7	1	+	--	--	+
Lag-3	Laggard Island	31-Mar-11	S 64° 48.497', W 64° 00.564'	26.5	1	+	+	--	--
Lit-1	Litchfield Island	3-Apr-11	S 64° 46.464', W 64° 05.280	26.2	1	+	+	+	+
Lit-2	Litchfield Island	3-Apr-11	S 64° 46.464', W 64° 05.280	22.6	1	+	+	--	+
Lit-3	Litchfield Island	3-Apr-11	S 64° 46.464', W 64° 05.280	23.8	1	--	+	+	--
Nor-1	Norsel Point	27-Mar-11	S 64° 45.620', W 64° 05.883'	28.4	2	+	+	+	+
Nor-2	Norsel Point	27-Mar-11	S 64° 45.620', W 64° 05.883'	25.6	1	+	+	--	+

1186 Table S6. Picoplankton collections used in v4 rRNA gene tag sequencing.
1187

Picoplankton sample ID	Collection Site	GPS Location	Date	Depth (m)	Temperature (°C)
Palmer PEL 1 2-22-08	B-C	S 64° 47.009' W 64° 46.656'	22-Feb-08	10	0.5
Palmer PEL 2 2-25-08	B-C	S 64° 47.009' W 64° 46.656'	25-Feb-08	10	1.5
Palmer PEL 4 2-2-08	B-C	S 64° 47.009' W 64° 46.656'	29-Feb-08	10	1.6
Palmer PEL 5 3-3-08	B-C	S 64° 47.009' W 64° 46.656'	3-Mar-08	10	1.6
Palmer PEL 6 3-6-08	B-C	S 64° 47.009' W 64° 46.656'	6-Mar-08	10	1.6
IPY-206-3	B	S 64° 46.769' W 64° 04.353'	24-Jul-08	10	-0.9
IPY-210-4	B	S 64° 46.769' W 64° 04.353'	28-Jul-08	10	-0.7
IPY-225-9	B	S 64° 46.769' W 64° 04.353'	12-Aug-08	10	-1.5
IPY-243-14	Palmer Deep	S 64° 55.556' W 64° 82.962'	30-Aug-08	500	-1.7
IPY-243-15	Palmer Deep	S 64° 55.556' W 64° 82.962'	30-Aug-08	10	-1.7
IPY-245-16	Gerlache Strait	S 64° 38.628' W 62° 53.343'	1-Sep-08	500	-0.7
IPY-245-17	Gerlache Strait	S 64° 38.628' W 62° 53.343'	1-Sep-08	10	-1.8
IPY-246-19	LaPeyrére Bay	S 64° 25.275' W 63° 16.974'	2-Sep-08	10	-1.1
IPY-256-21	B	S 64° 46.769' W 64° 04.353'	12-Sep-08	10	n.d.

1188


Core80-ASV
Dynamic50-ASV
Microbiome isolates
Clone library, Riesenfeld et al. 2008
0.5
SaM ASV2
SaM ASV31
Microbulbifer uncultured bacterium clone Tun3b.D7 (FJ169207)
SaM ASV17
SaM ASV35
SaM ASV34
SaM ASV23
SaM ASV1
SaM ASV18
SaM ASV5
Microbulbifer epialgicus F-104 (AB266054)
Microbulbifer pacificus SPO729 (DQ993341)
SaM ASV10
SaM ASV4
Endozoicomonas elysicola MKT110 (AB196667)
Endozoicomonas euniceicola EF212 (JX488684)
SaM ASV7
Oceanospirillaceae bacterium 1CH5 (HG004167)
SaM ASV27
Uncultured bacterium clone Tun3b.B6 (FJ169197)
SaM ASV26
Haliea sp. SY02 (GU124598)
BOMB.9.10.21 (MN960544)
Psychromonas profunda 2825 (AJ416756)
BOSW.4.10.32 (MN960543)
Photobacterium frigidiphilum SL13 (AY538749)
BOMB.3.2.14 (MN960541)
Moritella sp. MUD3 (HM771245)
BOMB.9.10.16 (MN960547)
Shewanella sp. 109 (DQ325519)
BONS.1.10.24 (MN960545)
BOMB.10.10.5 (MN960550)
BOSW.4.10.25 (MN960549)
Pseudoalteromonas haloplanktis TAC125 (CR954246)
SaM ASV28
Gamma proteobacterium S-1 (HQ379738)
Nitrosospira lacus APG3 (KC477402)
SaM ASV13
Uncultured bacterium clone Tun3b.G11 (FJ169217)
Gammaproteobacteria
SaM ASV8
Rhodospirillales uncultured bacterium clone nbw776h01c1 (GQ009460)
Phaeospirillum oryzae JA317 (AM901294)
Maritalea porphyrae LCM-3 (AB583774)
SaM ASV16
Antarctic Octadecabacter sp. SS10.21 (KC160714)
Loktanella fryxellensis LMG22007 (NR025607)
SaM ASV21
Sulfitobacter uncultured Antarctic sea ice clone ANTXI/4 14-133 (AY165563)
SaM ASV6
Hoeflea uncultured bacterium clone Tun3b.E2 (FJ169210)
Hoeflea halophila JG120-1 (GU564401)
SaM ASV3
Pseudovibrio sp. Tun.PHSC04-5.I4 (FNLB01000000, MN960556)
Uncultured bacterium clone Tun3b.B3 (FJ169195)
Pseudovibrio ascidiaceicola F423 (AB175663)
Pseudovibrio axinellae Ad2 (JN167515)
Alphaproteobacteria
Bdellovibrio uncultured bacterium clone SGUS1230 (FJ202161)
SaM ASV9
Bdellovibrio marine metagenome TARA 076 MES 0.22-3 (CEQH01010717)
Delta
Proteobacteria
SaM ASV11
Nitrospira uncultured Arctic bacterium sediment surface50 (GQ259315)
Nitrospira moscoviensis (X82558)
Nitrospirae
Uncultured bacterium clone Tun3b.H11 (FJ169219)
Ilumatobacter fluminis YM22-133 (AB360343)
Conexibacter arvalis KV-962 (AB597950)
SaM ASV20
Solirubrobacterales uncultured bacterium 7A 10-014 (KY190671)
Actinobacteria
SaM ASV19
Cryomorphaceae uncultured bacterium clone Tun3b.D1 (FJ169200)
SaM ASV29
Cryomorphaceae uncultured bacterium clone N0005 (JX391441)
Brumimicrobium glaciale ACAM 645 (NR025255)
SaM ASV25
Uncultured bacterium clone Tun3b.C6 (FJ169199)
Polaribacter irgensii ANT9243 (AY167325)
SaM ASV30
Flavobacteriaceae uncultured bacterium clone VS CL-234 (FJ497482)
SaM ASV12
Lutibacter oricola UDC377 (HM031972)
SaM ASV24
Flavicella marina 2A-7 (AB915171)
SaM ASV22
SaM ASV32
Bacteriodetes
Uncultured bacterium clone Tun3b.E7 (FJ169211)
Candidatus Saccharibacteria unidentified bacterium W4-B20 (AY345503)
Candidatus Saccharibacteria
SaM ASV36
Rubritalea spongiae YM21-132 (AB297805)
SaM ASV14
Verrucomicrobia bacterium IMCC11288 (KJ411743)
Coraliomargarita akajimensis DSM45221 (CP001998)
Puniceicoccus vermicola IMCC1545 (DQ539046)
Alterococcus agarolyticus ADT3 BCRC17102 (AF075271)
Opitutus terrae PB90-1 (AJ229235)
Geminisphaera uncultured wetland bacterium clone SEAA1AE101 (KC432173)
SaM ASV15
Uncultured bacterium clone Tun3b.F1 (FJ169213)
Verrucomicrobia
SaM ASV33
Gimesia maris DSM 8797T (AJ231184)
Haloquadratum walsbyi C23 (AY676200)
Planctomycetes

1189

1190 Fig. S1. Maximum likelihood phylogenetic tree based on X aligned bases. Archaeon,
1191 *Haloquadratum walsbyi* was used as an outgroup. Core (green diamonds) and dynamic (blue
1192 diamonds) ASVs are shown, along with cultivated isolates from the *Synoicum adareanum*
1193 microbiome (light blue triangles) and cloned rRNA gene sequences from an earlier study with *S.*
1194 *adareanum* collected in 2006 (Riesenfeld et al. 2008). *[note upon final submission, this figure*
1195 *will appear as a standalone PDF in which it will be in full size]*

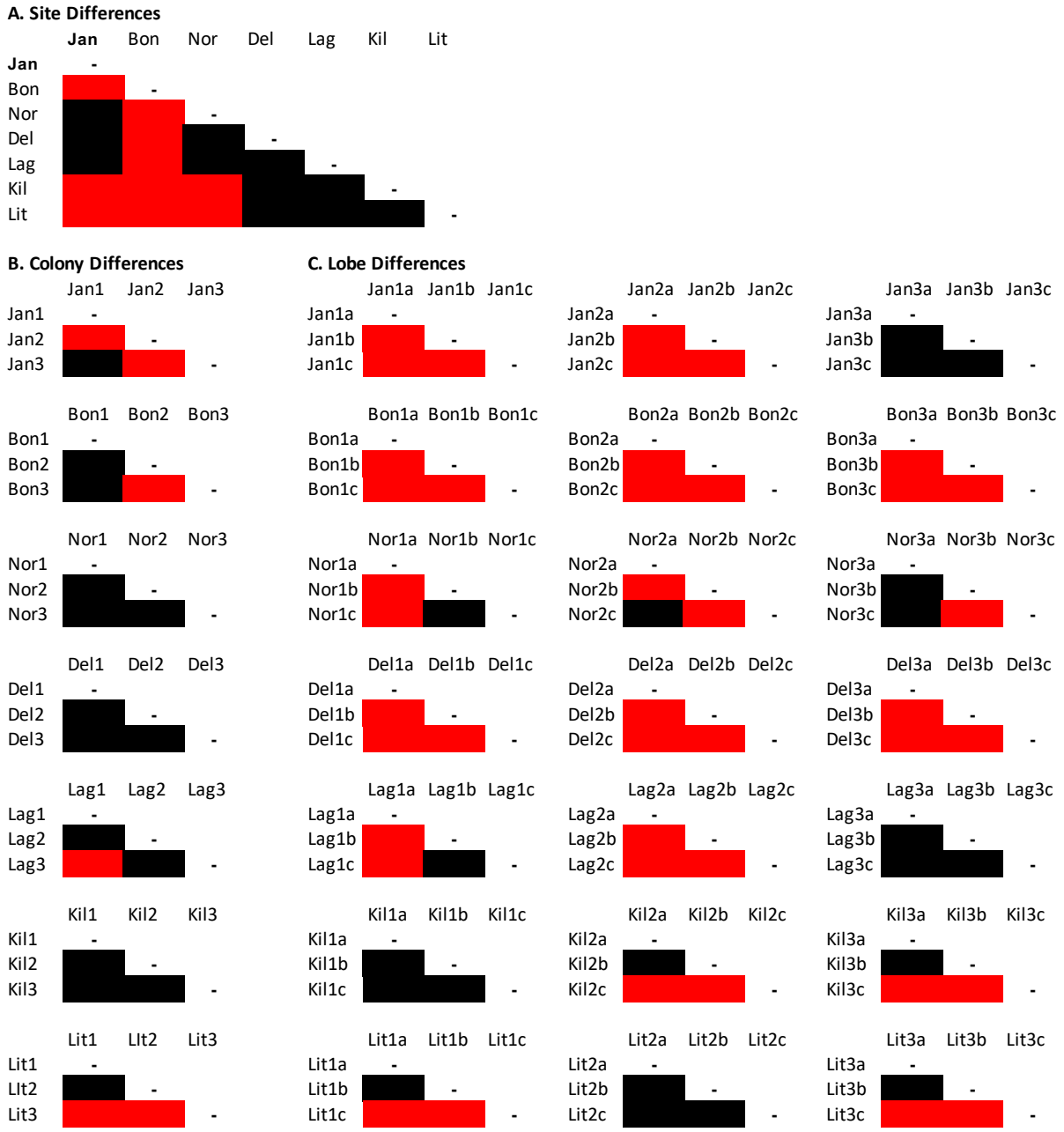


Fig. S2. Results of pairwise t-tests of PalA levels determined by mass spectrometry in which significant ($p \leq 0.05$) paired comparisons are indicated in red, and nonsignificant in black. n=9 for the site to site comparisons in (a); n=3 for colony to colony comparisons in (b); and n=3 for lobe:lobe comparisons in (c).

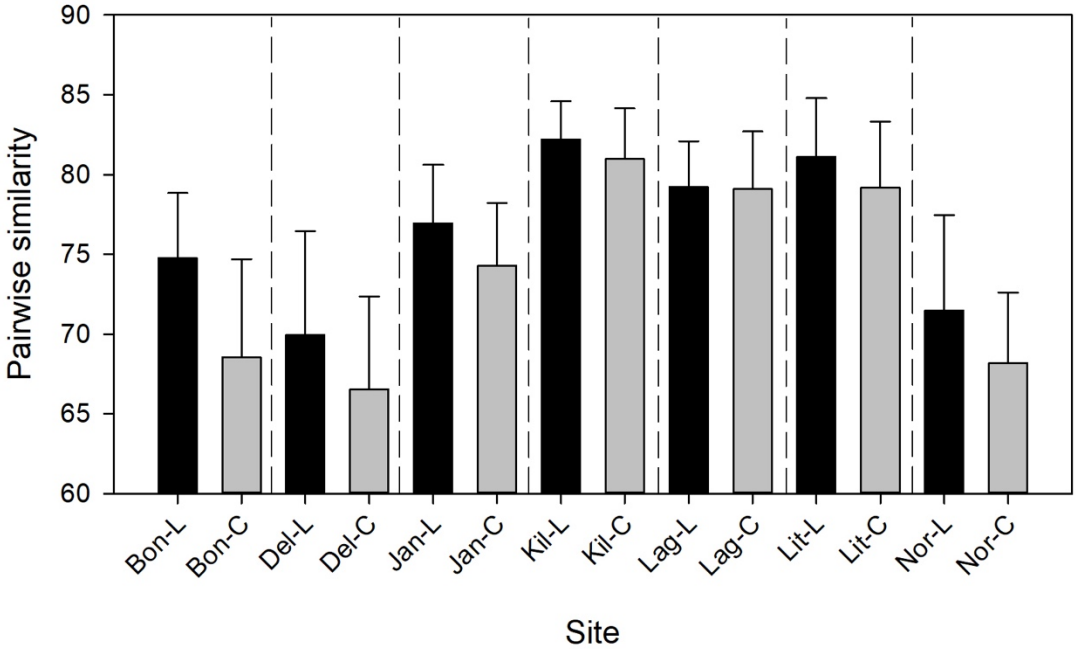


Fig. S3. Average pairwise similarity within (labels with -L, black bars) and between *S. adareanum* colony microbiome community structures (labels with -C, gray bars) at each dive site (abbreviations for sites as in Fig. 3). Standard deviation shown, n=9 for lobes and n=27 for colonies. Bray Curtis similarity was calculated with square-root transformed occurrence data.

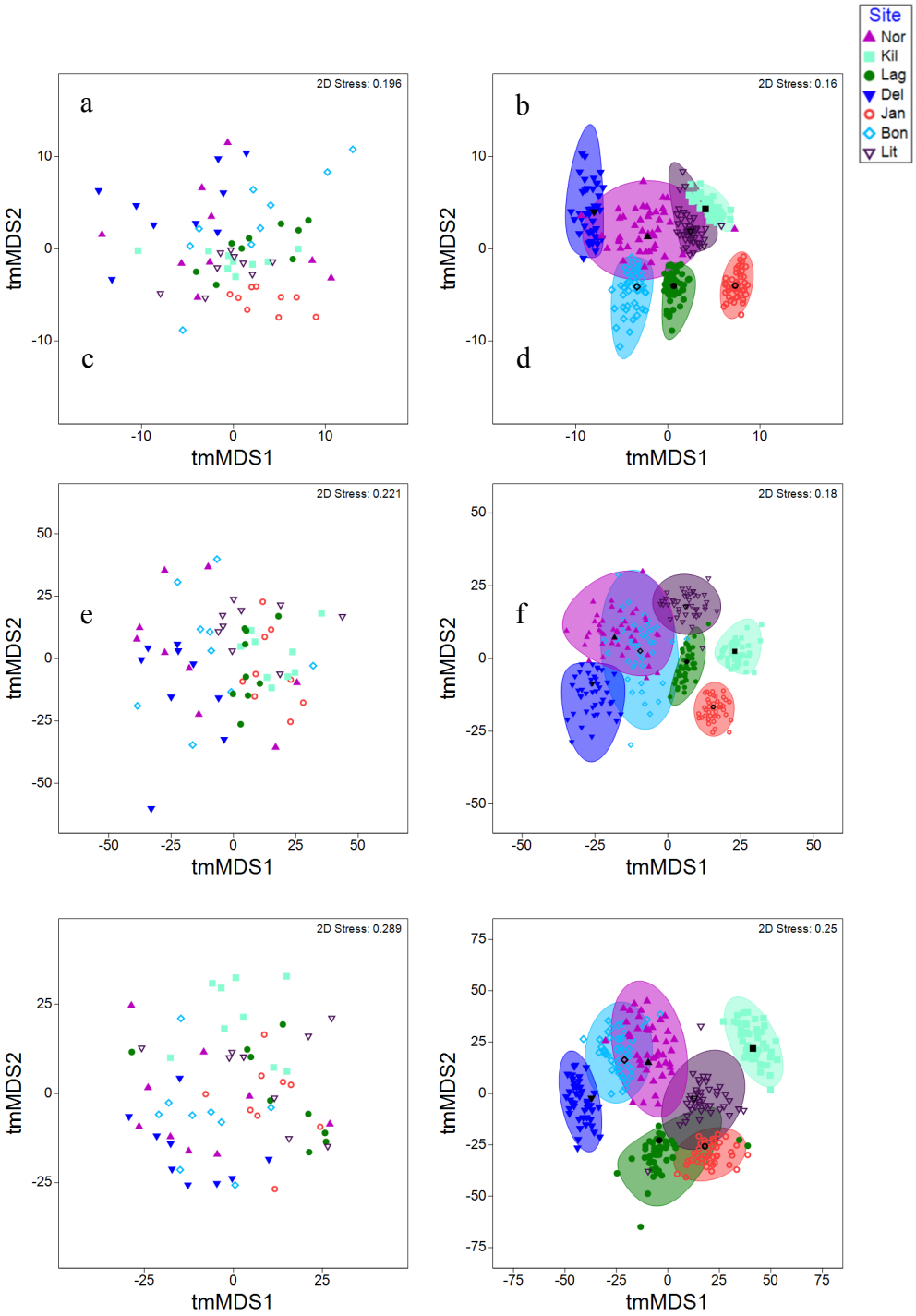


Fig. S4. tmMDS plots (Bray-Curtis; square root transformed ASV occurrence data) representing the microbiome of the 63 *S. adareanum* samples using the ASV occurrence profiles for the core (a), dynamic (c), and variable (e) microbiome fractions. Bootstrapped data MDS plots (43 bootstraps) for subgroups are also provided (b, d, and f); the resulting data cloud for each site is shaded to aid visualization of the cluster. The site-based centroid is demarcated by its respective symbol in black.

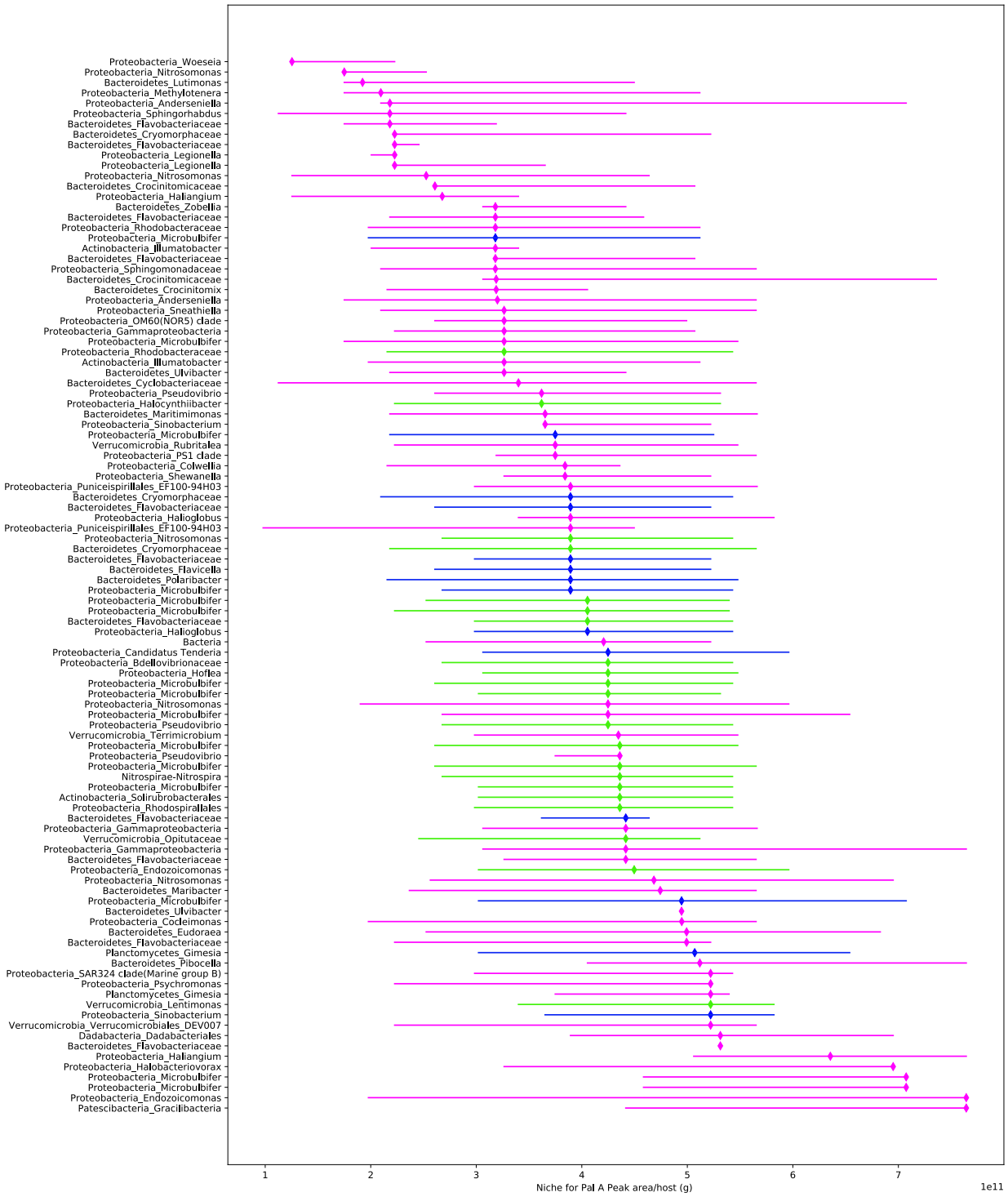


Fig. S5. PalA niche optimum for *S. adareanum* microbiome ASVs. PalA range and median values (diamond symbols) are shown according to microbiome membership (Core80, green; Dynamic50, blue; Variable, pink). Data were ordered from low to high PalA niche values (i.e. ASVs found in low PalA-containing tissues at the top, and those found associated with high PalA tissues at the bottom). Taxonomic phyla and highest level of assignment are indicated.