marine drugs

MDPI

Research Article

Uncovering the core microbiome and distributions of palmerolide in *Synoicum adareanum* across the Anvers Island archipelago, Antarctica

Alison Murray [1*], Nicole Avalon [2], Lucas Bishop[1], Karen W. Davenport[3], Erwan Delage [4], Armand E.K. Dichosa[3], Damien Eveillard[4], Mary L. Higham[1], Sofia Kokkaliari[2], Chien-Chi Lo[3], Christian S. Riesenfeld[1], Ryan M. Young[4^], Patrick S.G. Chain[3*], Bill J. Baker[2*]

[1] Division of Earth and Ecosystem Science, Desert Research Institute, Reno, Nevada, USA; Alison.murray@dri.edu (A.E.M); bishoplucas95@gmail.com (L.B.); mary.higham@dri.edu (M.L.H); csriesenfeld@gmail.com (C.S.R.)

[2] Department of Chemistry, University of South Florida, Tampa, Florida, USA; neavalon@usf.edu (N.E.A.); skokkaliari@usf.edu (S.K.); ryan.young@nuigalway.ie (R.M.Y.); bjbaker@usf.edu (B.J.B.)

[3] Bioscience Division, Los Alamos National Laboratory, Los Alamos, New Mexico, USA; kwdavenport@lanl.gov (K.W.D); armand@lanl.gov (A.E.K.D.); chienchi@lanl.gov (C-C.L.); pchain@lanl.gov (P.S.G.C.)

[4] LS2N, Université de Nantes, CNRS, Nantes, France; erwan.delage@univ-nantes.fr (E.D.); damien.eveillard@univ-nantes.fr (D.E.)

* Correspondence: alisonemurray@gmail.com; bjbaker@usf.edu, pchain@lanl.gov

Received: date; Accepted: date; Published: date

Abstract: Polar marine ecosystems hold the potential for bioactive compound biodiscovery, based on their untapped macro- and microorganismal diversity. Characterization of polar benthic marine invertebrate-associated microbiomes is limited to few studies. This study was motivated by our interest in better understanding the microbiome structure and composition of the ascidian, *Synoicum adareanum*, in which the bioactive macrolide that has specific activity to melanoma, palmerolide A (PalA), was found. PalA bears structural resemblance to a combined nonribosomal peptide polyketide, that has similarities to microbially-produced macrolides. We conducted a spatial survey to assess both PalA levels and microbiome composition in *S. adareanum* in a region of the Antarctic Peninsula near Anvers Island (64° 46'S, 64° 03'W). PalA was ubiquitous and abundant across a collection of 21 ascidians (3 subsamples each) sampled from seven sites across the Anvers Island archipelago. The microbiome composition (V3-V4 16S rRNA gene sequence variants) of these 63 samples revealed a core suite of 21 bacteria, 20 of which were distinct from regional bacterioplankton. Co-occurrence analysis yielded several potentially interacting subsystems 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. Taking these results together with an analysis of biosynthetic potential of related microbiome taxa indicates a core microbiome with substantial promise for natural product biosynthesis that likely interact with the host and with each other.

Keywords: Antarctica; ascidian; microbiome; microbial diversity; palmerolide A; co-occurrence

1. Introduction

Microbial partners of marine invertebrates play intrinsic roles in the marine environment at both the individual (host survival) and community (species distribution) levels. Host-microbe relationships are mediated through complex interactions that can include nutrient exchange, environmental adaptation, and production of defensive metabolites. These functional interactions are tied to the structural nature (diversity, biogeography, stability) of host and microbiome, and the ecological interactions between them. Studies of sponges, corals, and to a lesser degree, ascidians have revealed strong trends in invertebrate host species specificity to particular groups of bacteria and archaea. These studies have documented an underlying layer of diversity (e.g., [1-3]) in which

habitat and biogeography appear to have strong influences on the microbiome structure and function [4-6].

The vast majority of host-microbiome studies have been conducted at low- and mid-latitudes from coastal to deep-sea sites. High latitude benthic marine invertebrate-associated microbiome studies are currently limited to the Antarctic, where just the tip of the iceberg has been investigated for different host-microbe associations [7] and ecological understanding is sparse. Antarctic marine invertebrates tend to have a high degree of endemicity at the species level, often display circumpolar distribution, and in many cases have closest relatives associated with deep-sea fauna. Whether endemicity dominates the microbiomes of high latitude benthic invertebrate is currently not known, nor is the extent of diversity understood within and between different host-associated microbiomes. Likewise, reports of core (conserved within a host species) microbiomes within Antarctic invertebrate species are sparse.

The few polar host-associated microbiome studies to date have documented varying trends in host-species specificity, with generally low numbers of individuals surveyed. For example, low species-specificity was reported in sponge microbiome compositions between different sub-Antarctic and South Shetland Island Mycale species [8] which shared 74% of the OTUs, possibly representing a cross-*Mycale* core microbiome. On the contrary, high levels of microbiome-host species specificity and shared core sequences within a species was observed in five McMurdo Sound sponge species [9]. The same was found across several Antarctic continental shelf sponge species [10]. Webster and Bourne [11] also found conserved bacterial taxa dominated by microorganisms in the class Gammaproteobacteria across the soft coral, *Alcyonium antarcticum*, sampled at three sites in McMurdo Sound. Another cnidarian, the novel ice shelf burrowing sea anemone *Edswardsiella andrillae*, contained novel microbiota, though the composition across a limited set of individuals was only moderately conserved in which some specimen were dominated by an OTU associated with the phylum Tenericutes and others, a novel OTU in the class Alphaproteobacteria [12]. Lastly, a single representative of the Antarctic ascidian *Synoicum adareanum* revealed limited rRNA gene sequence diversity, including the phyla Actinobacteria, Bacteroidetes, several Proteobacteria, Verrucomicrobia, and TM7 [13], though persistence of these taxa across individuals was not studied.

Alcyonium antarcticum (= *A. paessleri*) and *Synoicum adareanum* are both reported to be rich in secondary metabolites. The soft coral *A. antarcticum* produces sesquiterpenes that are unusual in bearing nitrate ester functional groups [14], while the ascidian *S. adareanum* is known to produce a family of macrolide polyketides, the palmerolides, which have potent activity against melanoma [15]. The role of the microbial community in contributing to host defensive chemistry, microbe-chemistry interactions and niche optimization, as well as microbe-microbe interactions, are unknown in these high latitude environments.

Here we have designed a study to investigate whether a core microbiome persists among *S. adareanum* holobionts that may inform our understanding of palmerolide origins. We conducted a spatial survey of *S. adareanum* in which we studied coordinated specimen-level quantitation of the major secondary metabolite, palmerolide A (PalA) along with the host-associated microbiome diversity and community structure across the Anvers Island archipelago (64° 46'S, 64° 03'W) on the Antarctic Peninsula (Figure 1). The results point to a core suite of microbes associated with PalA-containing *S. adareanum*, distinct from the bacterioplankton, which will lead to downstream testing of the hypothesis that the PalA producer is part of the core microbiome.

2. Results

2.1 *Variation in holobiont PalA levels across ascidian colonies and collection sites.*

Typical procedures for natural products chemistry samples utilize bulk specimen collections for chemistry extraction (~ 30 individual ascidian lobes per extraction in the case of *S. adareanum*). Prior to this study, variation in PalA content at the individual lobe or colony level (inset, Figure 1) was unknown. Our sampling design addressed within and between colony variation at a given sampling site, as well as between site variation. The sites were constrained to the region that was

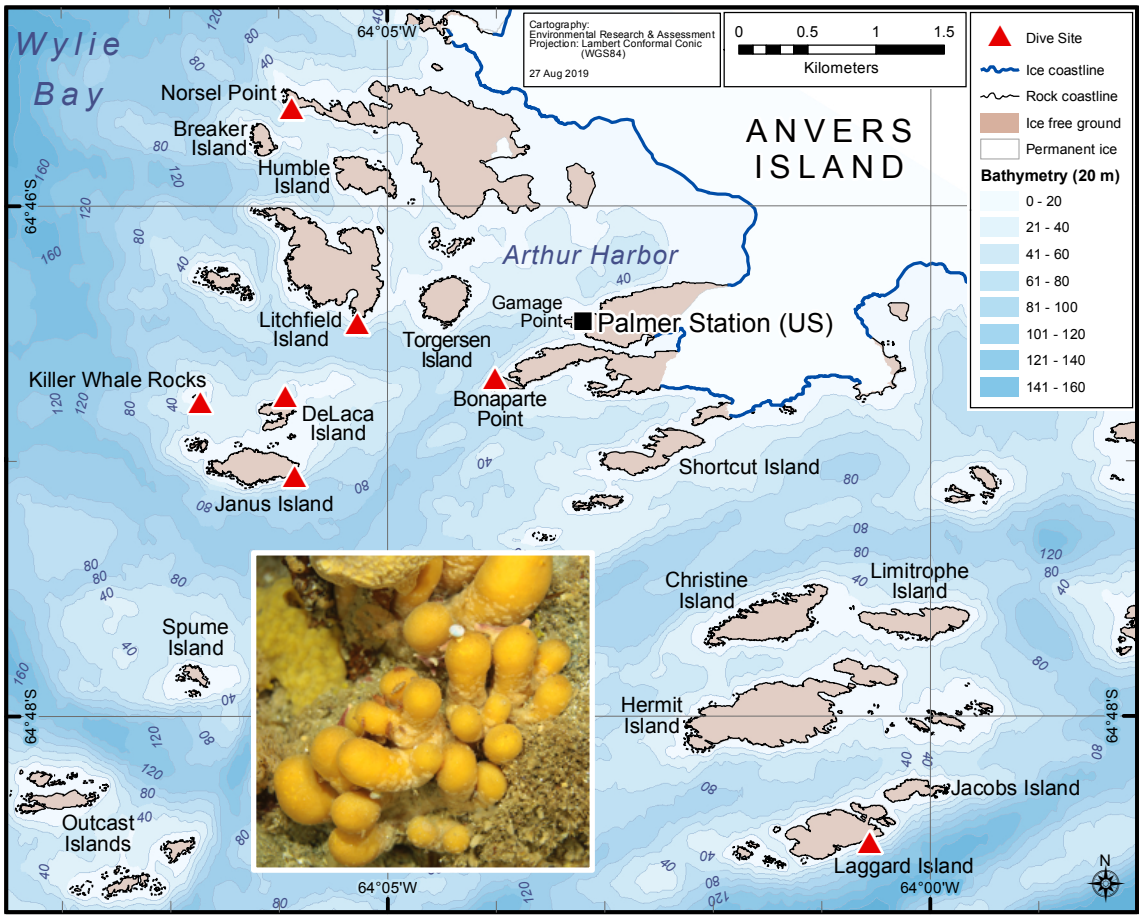

Figure 1. Bathymetric map of the study area off Anvers Island. *Synoicum adareanum* collection sites are shown with red triangles. The map was generated by Environmental Research & Assessment, Cambridge, UK, using Arthur Harbor bathymetry data from the PRIMO survey project 2004-06 (Dr. Scott Gallagher 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).

logistically accessible by zodiac in the Anvers Island Archipelago off-shore of the United States Antarctic Program (USAP)-operated Palmer Station. *S. adareanum* colonies were sampled across seven dive sites (Figure 1) in which three lobes per multi-lobed colonies were sampled at three colonies per dive site, totaling 63 lobes for PalA comparison. PalA stands out as the dominant peak in all LC/MS analyses of the dichloromethane:methanol soluble metabolite fraction of all samples analyzed (e.g., Figure 2a). The range in PalA levels varied about an order of magnitude 0.49 -4.06 mg PalA per g host dry weight across the 63 lobes surveyed. Our study design revealed lobe-to-lobe, intra-site colony level and some site-to-site differences in PalA levels ($p < 0.05$) in the archipelago (Figures 2b, S2). Within a given colony, the lobe-to-lobe variation was often high and significantly different in 17/21 colonies surveyed. Significant differences in PalA levels between colonies were also observed within some sites (Janus Island (Jan), Bonaparte Point (Bon), Laggard Island (Lag), and Litchfield Island (Lit); see Figure 1) in which at least one colony had significantly different levels compared to another colony or both. Despite this, we found differences between some of the sites. Namely Bon was significantly lower than all six other sites. This site is the closest to the largest island, Anvers Island, and Palmer Station. Samples from Killer Whale Rocks (Kil) and Lit sites were also found to have significantly higher PalA levels than Jan, Bon and Norsel Point

(Nor), although these did not appear to have a particular spatial pattern or association with sample collection depth.

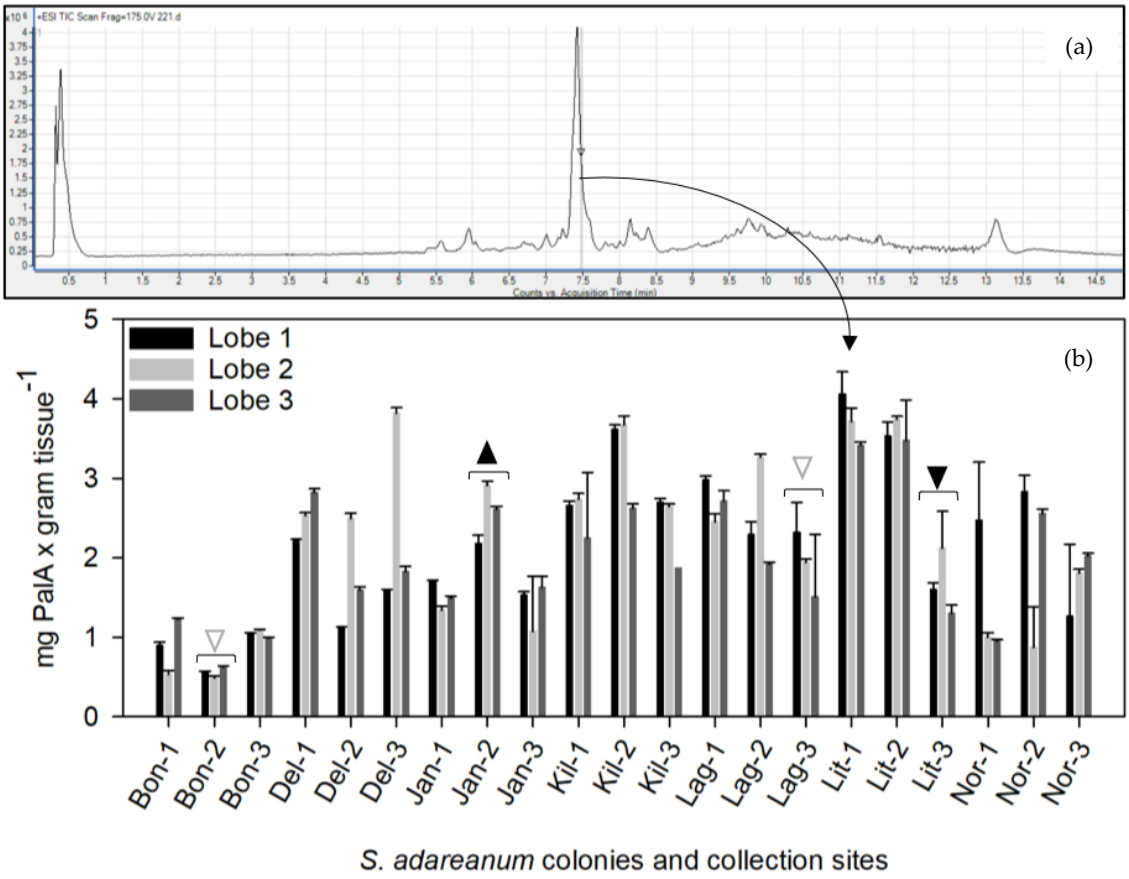

Figure 2. Palmerolide A (PalA) detection in *S. adareanum* holobionts. (a) Mass spectra derived from sample Lit-1a. The PalA peak dominates the dichloromethane-methanol fraction of the *S. adareanum* extract. (b) Levels of PalA normalized to tissue dry weight detected by mass spectrometry in *S. adareanum* 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 (Figure S2).

2.1. Characterization of host-associated cultivated bacteria.

Given our interest in identifying a PalA producing microorganism, we executed a cultivation effort with *S. adareanum* homogenate on three different marine media formulations at 10 °C. 16S rRNA gene sequencing revealed seven unique isolates (of 16 brought into pure culture) at a level of > 99% sequence identity. All but one of the isolates was affiliated with the class Gammaproteobacteria, including five different genera commonly isolated from marine enviornments (*Shewanella*, *Moritella*, *Photobacterium*, *Psychromonas* and *Pseudoalteromonas*, of which nine were highly related). In many cases, we characterized their near neighbors as marine psychrophiles, many from polar habitats (Figure S1). The exception to this was the isolation of a cultivar associated with the Alphaproteobacteria class, *Pseudovibrio* sp. str. TunPSC04-5.I4, in which its two nearest neighbors were isolated from a marine sponge and a different ascidian. This result marks the first *Pseudovibrio* sp. cultivated from high latitudes. HPLC screening results of biomass from all sixteen isolates did not reveal the presence of PalA.

2.2 *Synoicum adareanum microbiome (SaM).*

To understand the nature of conservation in the composition of the host-associated microbiome of *S. adareanum* we identified the microbiome structure and diversity (based on the V3-V4 region of the 16S rRNA gene) with sections of the 63 samples used for holobiont PalA determinations. It resulted in 461 amplicon sequence variants (ASVs) distributed over 14 bacterial phyla (Table 1). The core suite of microbes, defined as those present in > 80% of samples (referred to as the Core80), included 21 ASVs (six of which were present across all 63 samples). The Core80 ASVs represented the majority of the sequenced iTags (95% on average across all 63 samples), in which the first four ASVs dominated the sequence set (Figure 3). The ASVs present in 50-79% of samples represented the Dynamic50 category and contained 14 ASVs, which represented only 3.3% of the data set. The remaining ASVs fell into the Variable fraction, which included 426 ASVs, representing 1.7 % of the iTag sequences, yet the majority of phylogenetic richness (Table 1). Comparative statistical analyses were conducted with the complete sample set, which was subsampled to the lowest number (9987) of iTags per sample. This procedure was limited by one sample (Bon1b) that underperformed in terms of iTag sequence yield. After the elimination of this sample from the analysis, the 62-sample set had 19,003 iTags per sample with a total of 493 ASVs, the same 21 sequences in the Core80 (with seven common to all 62 samples), the same 14 Dynamic50 sequences, and a total of 458 Variable ASVs.

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

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

ASV11 most closely matched a sequence in the Nitrospirae family from Arctic marine sediments. There were two Verrucomicrobium-affiliated sequences represented in different families (Puniceicoccaceae, SaM_ASV14, and Opitutacae, SaM_ASV15). Lastly, there were two ASVs affiliated with the phylum Bacteroidetes: one related to polar strain, *Brumimicrobium glaciale* (SaM_ASV19), and the other to a marine *Lutibacter* strain (SaM_ASV12).

Five Dynamic50 ASVs were affiliated with the marine Bacteroidetes phylum (Cryomorphaceae and Flavobacteriacae-related), in addition to six ASVs associated with the class Gammaproteobacteria (including four additional *Microbulbifer*-related sequences). There was also two additional phyla, an ASV related to a sponge-affiliated Verrucomicrobium isolate, and a Planktomycetes-related ASV (Table 1, Figure S1). A number of these ASVs were most closely related to isolates from marine sediments.

Interestingly, five sequences identified from earlier cloning and sequencing efforts with this host-associated microbiome (Figure S1; [13]) matched sequences in the Core80 and Dynamic50 data sets. Phylogenetic comparisons also revealed that the SaM isolates were distinct from the Core80

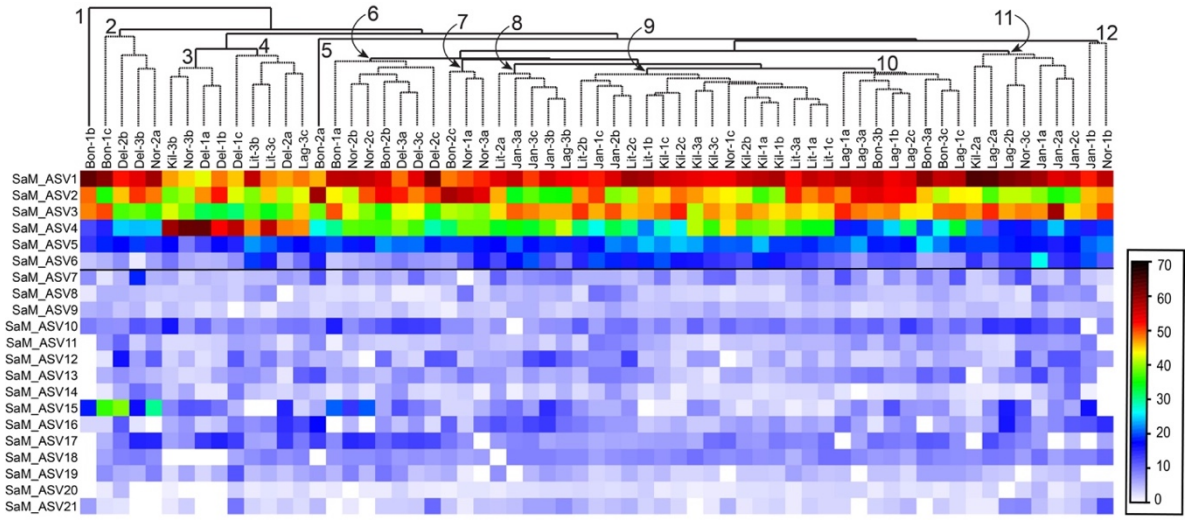

Figure 3. Heatmap of square root transformed ASV occurrence data for the core microbiome. ASVs (ranked and numbered) are shown on the y-axis, and 63-samples were hierarchically clustered, shown on the y-axis (site-based; square root transformed abundance data). Nodes with significant clusters are indicated 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.

and Dynamic50 except for the *Pseudovibrio* sp. TunPSC04-5.I4 isolate, which was present in both the Core80, and the clone and sequencing study.

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

Overall the community structures in the *S. adareanum* microbiomes across the 63 lobes surveyed had a high degree of similarity. Bray-Curtis pairwise similarity comparisons between lobes and colonies within each site were higher than 54% in all cases. When comparing the averages of pairwise similarity values within and between colonies, all sites, other than Lag, had higher similarity values within lobes in the same colony (ranging from 69.9 - 82.2%) compared to colonies within a site (66.5-81.0%; Figure S3), although the differences were small, and only Bon and Jan were significantly different ($p< 0.05$). We performed a two-dimensional tmMDS analysis based on Bray-Curtis similarity to investigate the structure of the microbiome between sites (Figure 4a). The microbiomes sampled at Kil and Lit had the highest overall degree of clustering (>75% similarity) while Kil, Lag, Lit, and Jan samples all clustered at a level of 65%. The microbiomes from DeLaca Island (Del) were the most dissimilar which was supported by SIMPER analysis in which two of the most abundant ASVs in the Core80 were lower than the average across other sites (SaM_ASV1 and 3) while others in the Core80 (SaM_ASV4, 15, and 17) were higher than the average across sites. We also performed a 2D tmMDS on the SaM fractions (Core80, Dynamic50, Variable) with and without permutational iterations which showed similar trends although partitioning of community structures between sites was more evident with the permutations (Figure S4). Site-based clustering patterns shifted to some degree in the different SaM fractions. For the Core80 alone, Jan samples clustered apart from the others. For Dynamic50, both Jan and Kil were outliers. Finally, for Variable fraction, Kil and Del samples clustered apart from the other sites. Variable fraction was more homogeneous, obscuring any site-to-site variability, while the core displayed tighter data clouds that showed a modest level of dispersion.

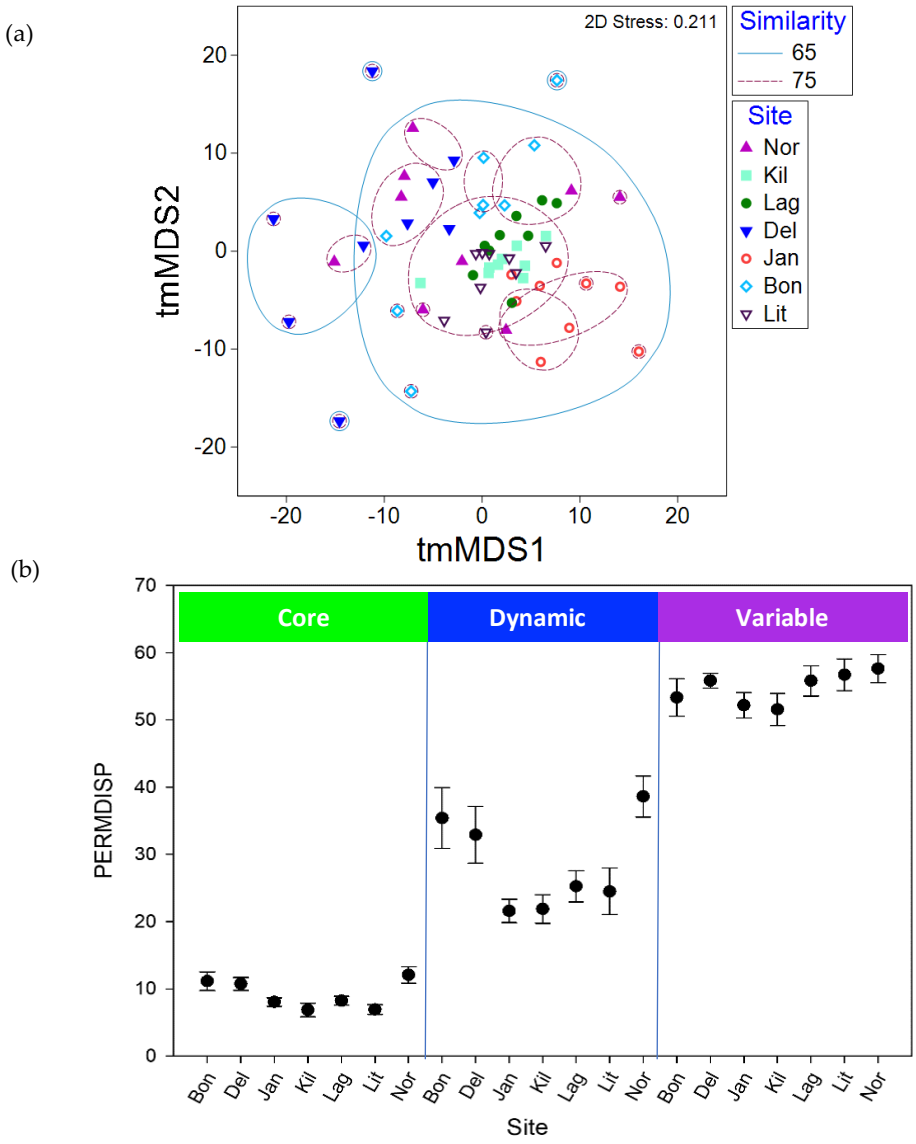

Figure 4. Similarity relationships amongst the *S. adareanum* microbiome samples in the Anvers Island archipelago. (a) tmMDS of Bray-Curtis similarities of square root transformed ASV occurrence data representing the microbiome of the 63 *S. adareanum* lobe samples using the complete ASV occurrence profiles. Microbiome samples with significant levels of similarity are shown (see legend). (b) B-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.

A PERMANOVA analysis investigated the drivers of variability in community structure in which we tested the role of colony-level, site-level, and stochastic environmental variation. When evaluating the whole community (all sites and ASVs), site-to-site differences explained 25% of the variability in the microbiome (Table S1). Colony-to-colony differences explained 28%, while the remaining 47% of the variability was unexplained and is likely attributed to stochastic environmental variation. When the SaM fractions were analyzed separately, the most significant difference ($p < 0.05$) was in the Variable fraction of the microbiome in which the site-to-site differences explained only 19.2% of the variation and the residual (stochastic) level increased to 58.4% (Table S1). Conversely, PERMDISP (Figure 4), a measure of dispersion about the centroid calculated for each site (a measure of β-diversity), revealed differences in the community structures

in each SaM fraction across sites, as well as differences between the microbiome fractions. The Core80 had a low level of dispersion (PERMDISP average of 9.1, range: 6.8-12.1), compared to the Dynamic50 (average 28.6, range: 24.5-38.6), which were moderately dispersed about the centroid, and were more variable with differences between the sites more apparent; Bon, Del, and Nor had higher values compared to the other four sites. The Variable fraction had high PERMDISP values (average 54.7, range: 51.6-57.6), representing high differences in β-diversity in which values were relatively close across the seven sites.

2.3 *ASV co-occurrences, and relationship to PalA.*

To further investigate the ascidian-microbiome ecology, we performed a network analysis of the ascidian holobiont with a particular emphasis on ASV co-occurrence and ASV ecological niche associated with PalA. The co-occurrence network depicted a sparse graph with 102 nodes and 64 edges (average degree 1.255; diameter 15; average path length 5.81) that associated ASVs from similar SaM fractions (assortativity coefficient of 0.41), confirming the robustness of the above discrimination. Upon inspection of the co-occurrence network, a dominant connected component contained 42.1% of ASVs, in which we identified three highly connected modules (via the application of an independent network analysis WGCNA, soft threshold = 9, referred to here as subsystems; Figure 5). The three subsystems were not significantly associated with PalA (resp. Corr= -0.23; 0.033; -0.17 for subsystems 1, 2 and 3). The network also contained several smaller systems (2-6 nodes), and 30 singletons, only one of which was in the Core80 (Gammaproteobacteria class, *Endozoicomonas*-affiliated). The three subsystems included ASVs from the Core80, Dynamic50 and Variable SaM fractions, but were unevenly distributed. Mostly driven by Core80 SaM fractions, Subsystem 1 harbored most of the highly represented ASVs, including the *Microbulbifer*-related ASVs as well as the *Pseudovibrio* ASV. Subsystem 2, interconnecting Subsystems 1 and 3, is mostly driven by Variable fraction ASVs and includes lower relative abundant Core80 *Hoeflea* and the Opitutaceae ASVs, as well as several Bacteroidetes taxa. Finally, Subsystem 3 was smaller. It included several diverse taxa dominated by Dynamic50 fraction, but still including two representatives from the Core80, *Lentimonas*, and Cryomorphaceae-affiliated ASVs. The overall role of Variable ASVs in the network is surprising, as they appear to be critical nodes linking the Subsystems, via 13 edges that lie between Subsystems 2 and 3. Another, perhaps noteworthy smaller connected component is in the lower left of the graph in which three Core80 ASVs, including the chemoautotrophic ammonia and nitrite oxidizers, *Nitrosomonas* and *Nitrospira*, were linked along with an uncharacterized Rhodospiralles-related ASV, and a Variable ASV.

To address the first-order question as to whether there was a relationship between the PalA concentration levels detected in the LC/MS analysis and the semi-quantitative ASV occurrences of *S. adareanum* ASVs, we performed three complementary analyses: correlation analysis, weighted co-occurrence networks analysis (WGCNA) and niche robust optima with PalA concentration as a variable. Pearson correlations between ASV occurrences and PalA concentrations ranged from -0.33 to 0.33 at the highest (24 of which were significant, ≤ 0.05; although none of those with a significant relationship were part of the core microbiome and were present in < 50% of the samples with a high occurrence of 24 sequences), suggesting little relationship at the gross level of dry weight normalized PalA and ASV relative abundance. This indicated that relative abundance of ASV is not a good predictor of PalA. A complementary WGCNA result showed no significant association between subsystem co-occurrence topology and PalA, pointing to the lack of relationship between microbial community structure and PalA. Finally, another aspect of the relationship between ASV occurrence and PalA levels was explored using the robust optimum method, which estimates the ecological optimum and tolerance range for ASVs about PalA. In this case we calculated the PalA niche optimum for each ASV and ranked them based on the median (Figure S5). Core80 ASVs showed a consistent PalA niche range. Furthermore, the median optima values of the Dynamic50 ASVs lie, for the most part, in lower or higher PalA optima compared to the Core80, with a substantial niche overlap between Core80. The Variable ASVs collectively lie at the lower and higher extremes of the optimum and tolerance range. Altogether, these complementary analyses

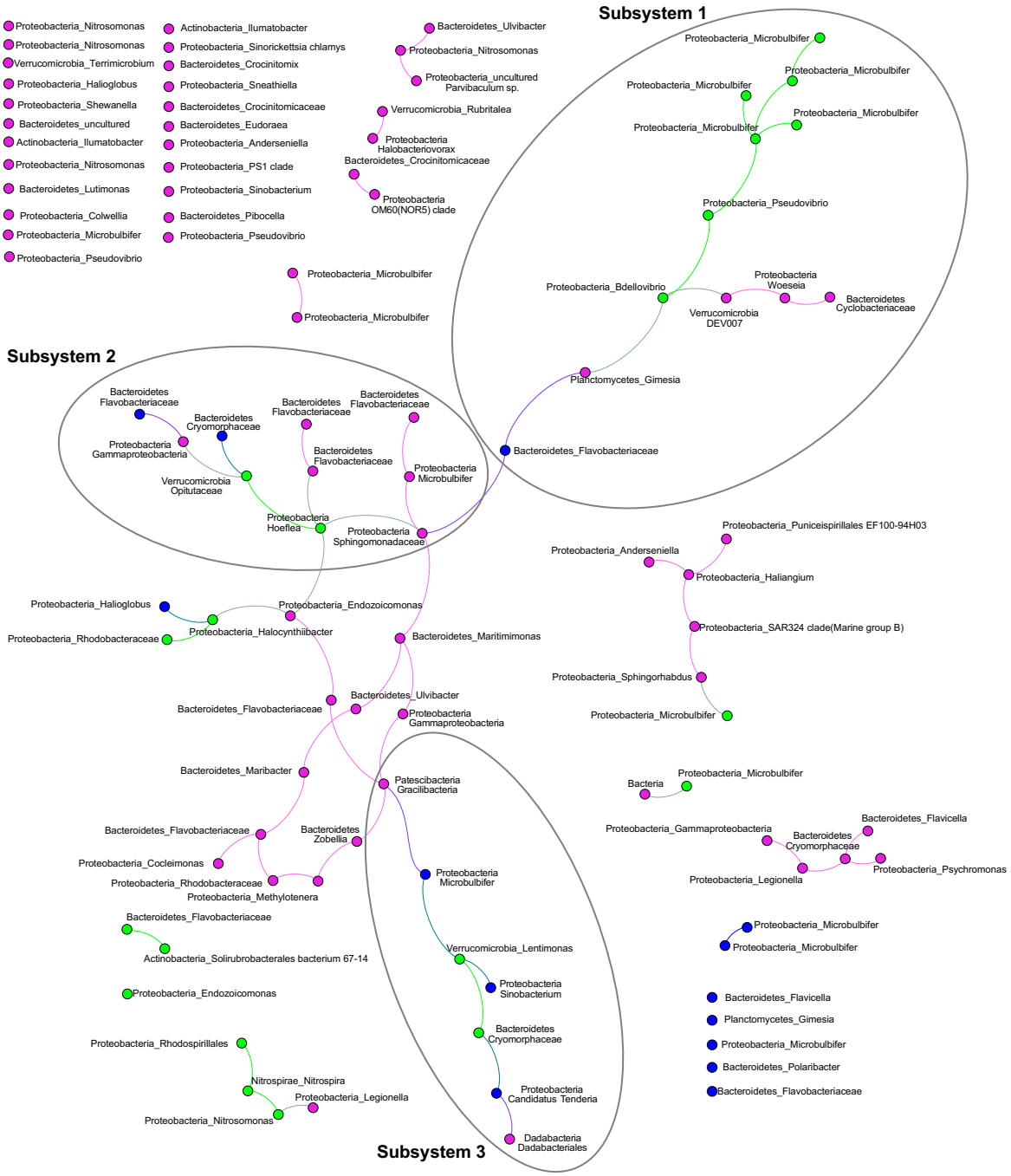

Figure 5. ASV co-occurrence network. The largest connected component of the co-occurrence network (seeded with ASVs found in at least 5 samples, 102 in total) identified three subsystems. Node colors represent the microbiome fractions (Core80, green; Dynamic50, blue; Variable, pink). Taxonomic identities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

advocated for not considering an individual nor a community effect on PalA, but rather, acclimation to the high levels of PalA observed that likely rely on unknown metabolic or environmental controls.

2.4 *Culture collection:microbiome and bacterioplankton comparisons.*

To address whether the composition of the *Synoicum adareanum*-associated bacterial cultivars were also present in the SaM or in the free-living bacterioplankton (< 2.5-µm fraction), we considered the overlap in membership between the isolate 16S rRNA gene sequences and these other two data sets. Representation of isolates in the 16S rRNA gene ASV data set was estimated by

317 comparing the two sequence data sets, albeit different ascidian samples were used for the culture
318 collection and the *S. adareanum* survey. Excepting *Pseudovibrio* sp. TunPSC04-5I4 that was present in
319 the Core80 with a 100% sequence match, two other isolates also had 100% matches to sequences in
320 the Variable SaM (BONS.1.10.24 and BOMB.9.10.19). Three other isolate 16S rRNA sequences
321 (BOMB.3.2.14, BOMB.9.10.16, BOMB.9.10.21) were found to match sequences in the Variable
322 microbiome at a level of 97% or higher. Only the *Pseudoalteromonas*-related isolate and the
323 *Shewanella*-related isolate sequences were not detected with relatives at a level of at least 97%
324 identity to the SaM ASVs; which could be explained by under-sampling. The bacterioplankton
325 composition was dominated by Gammaproteobacteria ASVs (47.35% of all ASVs; Table S2) in
326 which six of the isolates (BOMB.9.10.21, BONS.1.10.24, BOMB.9.10.16, BOMB.3.2.20, BOMB.3.2.14,
327 BONSW.4.10.32) matched sequences in the bacterioplankton data set at > 99.2% identity; even at a
328 level of 95% sequence identity the remaining ten isolates did not match sequences in the plankton,
329 including the *Pseudovibrio* sp. str. TunPSC04-5I4 isolate.

330 2.4 *Microbiome:bacterioplankton comparisons.*

331 Although at a high taxonomic level Proteobacteria and Bacteroidetes phyla dominated the
332 microbiome and bacterioplankton (Table S2), the relative proportions varied and the taxa
333 represented were quite different. Membership between SaM and the bacterioplankton data set
334 indicated a low-level overlap at 100% identity (Figure S6), with 39 of 604 perfectly matched ASVs.
335 At 100% ASV sequence identity, the results indicate a single, Core80 SaM ASV was a perfect match
336 with the bacterioplankton data set – the *Microbulbifer*-associated sequence that is the most abundant
337 across all 63 SaM data sets. Interestingly, this sequence was only identified in one bacterioplankton
338 sample (IPY-225-9) at a low occurrence (14 of > 1.18 million tags distributed across 604 ASVs). There
339 were three (of 14) Dynamic50 ASVs that were perfect matches with the bacterioplankton ASVs.
340 These were affiliated with a poorly classified Bacteroidetes family Flavobacteraceae ASV (77% of
341 SaM samples), a Gammaproteobacteria-associated *Sinobacterium* (73% of SaM samples), and a
342 second Gammaproteobacteria-associated ASV that is associated with *Candidatus* Tenderia (71% of
343 SaM samples). The remaining 35 perfect match ASVs between the two data sets were classified as
344 part of the Variable microbiome. These sequences fell across four phyla and nine classes, 13 of
345 which were well-distributed across the Bacterioplankton data set samples (>50%).
346 At a level of 97% ASV sequence identity, there were three additional matches between the
347 Bacterioplankton data set and the Core80. These included an Alphaproteobacteria-related *Hoeflea*
348 and *Halocynthibacter*-related sequences, and a *Nitrospira*-related sequence. There were also two other
349 Dynamic50-related ASVs: these were both related to unclassified Flavobacteriaceae. The rest (79
350 ASVs) of the matches at >97% were affiliated with the Variable fraction of the microbiome.

351 **3. Discussion**

352 This study reports our growing understanding of the microbiome composition of PalA-
353 containing *S. adareanum*. To enhance our understanding of the ecology of the PalA-containing
354 ascidian, *S. adareanum*, we investigated the ascidian colony microbiome and PalA chemistry levels
355 at an individual lobe level, and compared the ascidian microbiome to the plankton. This
356 comparison allowed us to address questions regarding variability in PalA content and whether a
357 conserved core microbiome occurs across these PalA-containing Antarctic ascidians, thereby
358 supporting the logic that if a microbial producer synthesizes PalA, the producing organism should
359 be present in all PalA-containing *S. adareanum* samples. Before this study, however, we did not have
360 quantitative data at the level of the individual ascidian lobe that forms the pedunculated *S.*
361 *adareanum* colonies (Figure 1). This discussion focuses on the core microbiome, then takes a broader
362 look at secondary metabolite distributions about the microbiome and in other marine invertebrates
363 as well as the biosynthetic potential of core membership, and concludes with information gained
364 from our initial cultivation effort.

365

366 3.1 *Core microbiome.*

367 Ascidian (host)-microbiome specificity is an active area of research. Compared to sponges and
368 corals, for example, ascidian microbiomes are less-well characterized. To get a broader perspective
369 on the microbiome we also used cultivation-independent approaches. We found that the Antarctic
370 ascidian *S. adareanum* has a persistent core microbiome across the Anvers Island archipelago that is
371 distinct from the plankton. This dissimilarity between ascidian host-associated microorganisms and
372 bacterioplankton appears to be a consistent observation across the global ocean (e.g., [16-19]). The
373 Core80 is comprised of ASVs that numerically dominate the community, as well as those
374 representing only a fraction of a percent of the sequences surveyed. Although ascidian symbioses
375 have not yet been systematically studied in the Antarctic, better-studied lower latitude ascidian
376 microbiomes provide several examples for comparison. The overall trend across ascidian
377 microbiome studies to date suggests that there is a high degree of both geographical as well as host-
378 species level specificity of microbiome composition (e.g., [16, 17, 20]). The same appears to be true
379 of *S. adareanum.* Though this study was restricted to a small geographical region, we identified a
380 conserved core of 21 16S rRNA gene sequence types across 63 individual pedunculate lobes
381 studied. We attribute the detection of this high degree of persistent members in part to the uniform
382 homogenization, extraction, and sequencing methodological pipeline applied. Microbiome analysis
383 is sensitive to sequencing depth, quality parameter choices, and algorithmic differences in data-
384 processing pipelines (amplicon sequence variants vs. cluster-derived operational taxonomic units),
385 which can impact direct comparisons between studies. Along these lines, the numerous highly
386 related Microbulbifer ASVs would have fallen into a single OTU (97% sequence identity), resulting
387 in a core with 14 members. These limitations aside, our findings are in line with several other
388 ascidian microbiome studies from lower latitudes in terms of the relative size of core membership
389 (where core definitions vary to some degree between studies). For example, *Styela plicata*, a solitary
390 ascidian, was reported to have a core membership of 10 [21] to 16 OTUs [22]. Other solitary
391 ascidians, including Herdmania momus had a core of 17 OTUs [21], while two *Ciona* species ranged
392 from 8-9 OTUs [23]. Temperate colonial ascidians *Botryliodes leachi* and *Botryllus schlosseri* ranged
393 from 10-11 members in their core microbiomes [20]. Also, an extensive survey of 10 different
394 ascidian microbiomes (representing both solitary and colonial forms) conducted on the Great
395 Barrier Reef reported core memberships ranging from 2 to 35 OTUs [16], while the numbers of
396 individuals surveyed in each case were only 2-3. Note that a few other studies reported much
397 higher numbers of shared OTUs ranging from 93-238 [18,19]; the scale of sequencing was higher in
398 these later studies. Further, as others have reported [24], the membership of these core ascidian
399 microbiomes is distinct, and in the case of SaM, the core microbiome diversity appears to be unique
400 at the ASV level, although several taxa are in common with other ascidian-associated microbes at
401 the genus level including *Microbulbifer* associated with *Cystodytes sp.* [25], *Pseudovibrio* with *Polycitor*
402 *proliferus* [26] and an *Endozoicomonas* specific-clade was identified in a survey of a number of
403 ascidians [27].

404 Predicted metabolic abilities of the Core80 taxa suggest aerobic heterotrophy (aerobic
405 respiration – organic carbon is the carbon and energy source), microaerophily (growth in low
406 oxygen conditions) and chemoautotrophy (CO_2 fixation provides carbon and reduced chemicals
407 provide energy, e.g., NH_4^+ and NO_2^-) are themes amongst the Core80, in which the most abundant
408 ASVs are high molecular weight carbon degraders. The *Microbulbifer* genus has members known to
409 degrade cellulose [28], and perhaps non-coincidently, ascidians are the only known invertebrate
410 capable of cellulose biosynthesis in the marine environment (e.g., [29, 30]). From this, we could
411 speculate that the *Microbulbifer* strains associated with *S. adareanum* could occupy a commensal, if
412 not somewhat antagonistic relationship [31]. In support of this possibility is the fact that the only
413 overlapping sequence between the Core80 and the bacterioplankton was a *Microbulbifer* sequence,
414 which was a rare sequence in the plankton – suggesting it may be an opportunistic member of the
415 *S. adareanum* microbiome. Besides, free-living and sponge-associated isolates from the *Microbulbifer*
416 genus have been found to produce bioactive compounds including pelagiomicins [32] and parabens

[33], respectively. This observation, in the least, suggests that the *Microbulbifer*(s) is/are likely well-adapted to their ascidian host and might be considered a potential PalA producing organism.

The NH_4^+-oxidizing *Nitrosopumulis*-related Thaumarchaeota have been commonly detected in ascidian microbiomes [16, 21, 22, 34], which contrasts phylogenetically, but not in terms of overall function with the NH_4^+-oxidizing *Nitrosomonas*-ASV that was part of the SaM core. The niche however has been reported to be different for the archaeal and bacterial ammonia oxidizers in which the archaea tend to be found in oligotrophic systems, while the bacteria (e.g., *Nitrosomonas*) can tolerate high levels of dissolved ammonia [reviewed by 35]. This result might reflect both the environment, and in situ *S. adareanum* tissue ammonia levels where it may accumulate, as several studies have reported on the high levels of oxidizing ammonia Thaumarchaeota in the coastal waters of the Anvers Island archipelago. This group, however, is numerous only in winter to early spring waters [36-38], while this study was conducted with samples collected in Fall when the ammonia-oxidizing Thaumarchaeota are not abundant in the coastal seawater [36], advocating for the comparisons between the SaM with bacterioplankton collected in both summer and winter periods.

In a similar vein, although we did not intentionally conduct a temporal study, the data from samples collected in 2007 and 2011, appear to suggest that a number of the core microorganisms are stable over time. We found several ASVs in 2011 samples that matched (at 100% sequence identity) cloned sequences from samples collected in 2007. Stability of the ascidian microbiome over time has been reported in a few studies [17, 24, 34]. Studying the persistence of the core membership over the annual cycle would be interesting (and provide compelling evidence for stable relationships) in this high latitude environment where light, carbon production, and sea ice cover are highly variable.

The co-occurrence analysis indicated three subsystems of ASVs that co-occur within *S. adareanum*. A small side-network included the two taxa involved with the 2-step nitrification process, including the Nitrosomonas ASV mentioned above and a Nitrospira ASV. Even though at present, the functional underpinnings of the host-microbial system have not been studied, the co-occurrence relationships provide fodder for hypothesis testing in the future. One interaction network that warrants mentioning here is the ASVs in Subsystem 1, which harbor several Core80 *Microbulbifer* ASVs, and the *Pseudovibrio* ASV are linked to a *Bdellovibrio* ASV that is also a member of the Core80. Members of the Bdellovibrio genus are obligate bacterial predators [32] that penetrate the outer membrane and cell wall of their prey. The linkage position in the subsystem is compelling in the sense that the *Bdellovibrio* could potentially control the abundance of the "downstream" less connected members of the network. Lastly, the positions of a couple of Dynamic and several the Variable ASVs in the network, as links between the subsystems, was unexpected. The central positions of these ASVs suggest that they may not be merely stochastic members of the microbiome; that they could play opportunistic, adaptive or ecological roles in the functionality of the microbiome subsystem(s) which potentially participate in different aspects of the holobiont system in particular, by promoting the switch between different ecological modes supported by different subsystems. Such roles were proposed for dynamic members of the *Styela plicata* microbiome [21].

The culturing effort succeeded in isolating a *Pseudovibrio* strain that is a crucial member of the Core80. In addition, several other Gammaproteobacteria-affiliated strains which matched sequences in the Variable SaM and the bacterioplankton were cultivated, however the cultivated diversity using the approaches applied here reared a collection of limited diversity. It is likely that additional media types and isolation strategies could result in additional cultivated diversity as there are a number of taxa with aerobic heterotrophic lifestyles in the Core80 that have been brought into pure culture (e.g., *Microbulbifer*, *Hoeflea*). One challenge we experienced using the nutrient replete media was overgrowth of plates, even at 10 °C.

3.2 *Secondary metabolite distributions and bioaccumulation in marine biota.*

Although the results of the archipelago spatial ascidian survey did not support a direct relationship between PalA levels and the relative abundance of microbiome ASVs, the results of the PalA niche analysis suggests that the Core80 ASVs occur in a preferred optimum and tolerance

range of PalA levels. The lack of specific ASV-PalA patterns may not be entirely surprising, as secondary metabolites result from a complex combination of metabolic reactions that require a fine-tuning to environmental conditions and further metabolic modeling for the sake of understanding. Furthermore, these metabolites have been found to accumulate in the tissues in several different marine invertebrates. The Optimal Defense Theory can be applied to marine invertebrates and reflects the hypothesis that secondary metabolites are distributed in specific tissues based on exposure and anatomic susceptibility for predation [40]. For example, nudibranchs sequester toxic compounds, which have been biosynthesized by the gastropod or acquired from their prey. The toxins are concentrated in the anatomical space of their mantles, the most vulnerable portion of their soft, exposed bodies [40-42]. Bioaccumulation of secondary metabolites in invertebrates with less anatomical differentiation is also known to occur. In the phylum Porifera, different cell types and layers have been studied to determine spatial and anatomical differences in secondary metabolite concentrations [43-45]. Compounds have been found to be concentrated spatially on the surface (e.g., [46]) or apical parts of the sponge [47] in some cases. Sponges may be able to differentially bioaccumulate secondary cytotoxic metabolites based on tissues more susceptible to predation [48]. Metabolite distribution investigations that are ascidian-specific are less well documented; however, there is also evidence of ascidian secondary cytotoxic metabolite bioaccumulation. The patellazoles, marine macrolides from the ascidian *Lissoclinum patella*, bioaccumulate in the ascidian tissues to concentrations up to seven orders of magnitude higher than their cytotoxic dose in mammalian cell lines [49, 50]. Additionally, there are other instances in which bioaccumulation in ascidian host tissues suggests metabolic cooperation of producer and host as well as compound translocation from producer to host [15, 51, 52]. Although the PalA levels were normalized to grams of dry lobe weight, tissue-specific spatial localization is a potentially confounding factor in the statistical analyses investigating the ASV:PalA relationship.

3.3 *Biosynthetic potential of the core.*

We investigated the natural product biosynthetic potential of the nine genera associated with 15 of 21 Core80 ASVs using antiSMASH (Table 2, Table S3). From this, it appears that all genera had at least one relative at the genus level with biosynthetic capacity for either polyketide or nonribosomal peptide biosynthesis or both. Even though the number of genomes available to survey were highly uneven, there is quite a disparity of biosynthetic capacity between the genera analyzed, thus, it appears that *Pseudovibrio*, *Nitrosomonas*, *Microbulbifer*, *Nitrospira* have the greatest capacities (in that order). Likewise, *Microbulbifer*, *Pseudovibrio*, *Hoeflea* and Opitutaceae might be prioritized as candidate PalA producers based solely on relative abundance ranking (Table 2; [24]). Although we did not conduct this analysis for the six ASVs that were classified at best at the family or order level, a few of these might be worth considering as potential producers considering their higher-level relationships with marine natural product producing lineages. For example, marine actinobacteria are classically associated with the production of numerous bioactive natural products (e.g., [53, 54]), although speculation is difficult with actinobacteria SaM_ASV20 in the core as it is only distantly related to known natural product producers. Likewise, Opitutaceae-related SaM_ASV15 is ranked 7 in terms of average relative abundance and falls in the same family the ascidian-associated *Candidatus* Didemnitutus mandela, which harbors the biosynthetic gene cluster predicted to produce mandelalide, a glycosylated polyketide [55]. From this we might prioritize the *Microbulbifer*, *Pseudovibrio*, and Opitutaceae ASVs for downstream investigation, with the lower relative abundance *Nitrosomonas* and *Nitrospira* ASVs also holding some potential given the perhaps surprising abundance of biosynthetic gene cluster content in these chemoautotrophic, and generally small genome size taxa. Although this analysis focused on predicted pathway characterizations across genera detected, the distributions of predicted pathways varied substantially across the taxa analyzed. The potential for these new Antarctic ascidian-associated strains to harbor secondary metabolite pathways remains speculative as they are amongst the most variable component of a bacterium's genome.

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

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			

4. Materials and Methods

4.1 *Cultivation-dependent effort.*

S. adareanum samples collected by SCUBA in 2004 and 2007 were used for cultivation (**Table S4**). The 2004 specimen were archived in 20% glycerol at -80 °C until processing by manual homogenization using sterilized mortar and pestle prior to plating a suspension onto marine agar 2216 plates in 2006. The 2007 samples were homogenized immediately following collection using sterilized mortar and pestle, and suspensions were prepared in 1X marine broth or filter-sterilized seawater then transferred at 4 °C to DRI. Shortly after (within 1 month of collection), the 2007

isolates were cultivated on three types of media in which suspensions initially stored in marine broth were plated onto marine agar (2216), while homogenate preparations stored in seawater were plated onto VNSS agar media [56] and amended seawater plates (3 grams yeast extract (Difco), 5 g peptone (Difco), and 0.2 grams casein hydroslyate (Difco) per liter). Colonies were selected from initial plates, and purified through three rounds of growth on the same media they were isolated on.

4.2 *Field sample collections for cultivation-independent efforts.*

Next, a spatial survey of *Synoicum adareanum* was executed in which samples were collected by SCUBA in austral fall between 23 March and 3 April 2011. Seven sampling sites (depths 24.7 - 31 m; Table S5) around the region accessible by Zodiac boat from Palmer Station were selected in which we sampled in a nested design where three multi-lobed colonies were selected from each site, and three lobes per colony were sampled (Figure 1). Underwater video [57] was taken at each site, then video footage was observed to note general ecosystem characteristics (% cover of major benthic species and algae). In total, 63 *S. adareanum* lobes were sampled (9 from each site). Samples were transported to Palmer Station on ice, and frozen at -80 °C until processing at DRI and USF. Frozen *S. adareanum* lobes were cut longitudinally in half for parallel processing through DNA and palmerolide detection pipelines.

Then, to address whether the composition of the SaM was distinct from the free-living bacterioplankton (< 2.5-µm fraction), we considered the overlap in membership between the SaM and bacterioplankton in the water column. To accomplish this, we used a reference seawater data set represented by samples that had been collected in February-March 2008 (five samples) and in August-September 2008 (nine samples) from LTER Station B near Anvers Island (east of the Bonaparte Point dive site), and at a few other locations in the region (Table S6). Seawater samples were collected by a submersible pump and acid washed silicone tubing at 10 m at Station B, and using a rosette equipped with 12 L Niskin bottles for the offshore samples at 10 m and 500 m (2 samples each depth). The February-March seawater samples were processed using in-line filtration with a 2.5 µm filter (Polygard, Millipore) to screen larger organisms and bacterioplankton were concentrated using a tangential flow filtration system and the cells were harvested on 25 mm 0.2 µm Supor filters (Millipore). The August-September seawater samples were processed using inline filtration with a 3.0 µm filter (Versapor, Millipore), and then bacterioplankton was collected onto 0.2 µm Sterivex (Millipore) filters using a multichannel peristaltic pump (Masterflex). All filters were immersed in sucrose:Tris:EDTA buffer [58] and stored frozen at -80 °C until extraction.

4.3 *Palmerolide A screening.*

Sixty-three frozen *Synoicum adareanum* lobes were cut in half. Half lobes were lyophilized and then exhaustively extracted using dichloromethane for three days, followed by methanol for three days. The extracts were combined and dried on a rotary evaporator. The extracts were filtered, dried down, and reconstituted at 1.0 mg/mL to ensure the injected concentration was consistent. The residue was subjected to Liquid Chromatography-Mass Spectrometry (LC-MS) analysis using a H_2O:ACN gradient with constant 0.05% formic acid. The high-resolution mass spectra were recorded on an Agilent Technologies 6230 electrospray ionization Time-of-Flight (ESI-ToF) spectrometer. LC-MS was performed using a C-18 Kinetex analytical column (50 x 2.1 mm; Phenomenex). The presence of PalA was verified using MS/MS on an Agilent Technologies 6540 UHD Accurate-Mass QTOF LC-MS.

The *S. adareanum*-associated microbial culture collection was screened for the presence of PalA. Isolates were cultivated in 30 mL volumes, and the resulting biomass was lyophilized and screened by HPLC. Each sample was analyzed in ESI-SIM mode targeting masses of 585 amu and 607 amu. The reversed-phase chromatographic analysis consisted of a 0.7 mL/min solvent gradient from 80% H_2O:CH_3CN to 100% CH_3CN with constant 0.05% formic acid using the same C-18 column as above. The analysis was conducted for over 18 min. PalA had a retention time of approximately 14 min. The twenty-four sample sequence was followed by a PalA standard to confirm its analytical characteristics.

4.4 *S. adareanum-associated microbial cell preparation.*

The outer 1-2 mm of the ascidian tissue, which could contain surface-associated microorganisms, was removed using a sterilized scalpel before sectioning tissue subsamples (~0.2 g) of frozen (-80 °C) *S. adareanum* (½ lobe sections). Tissue samples were diced with a scalpel before homogenization in sterile autoclaved and filtered seawater (1 mL) in 2 mL tubes. Each sample was homogenized (MiniLys, Bertin Instruments) using sterile CK28 beads (Precellys, Bertin Instruments) 3X at 5000 rpm for 20 sec, samples were placed on ice in between each homogenization. Homogenates were centrifuged at 500 × g at 4 °C for 5 min to pellet tissue debris. The supernatant was removed to a new tube for a second spin at the same conditions. This supernatant was decanted and the cell suspension centrifuged at 12,000 × g at 4 °C for 5 min to collect the microbial cells. Suspensions were stored on ice, then entered an extraction pipeline in which 12 samples were processed in parallel on the QIAvac 24 Plus Manifold (Qiagen).

4.5 *DNA extractions.*

S. adareanum-associated microbial cell preparations were extracted with Powerlyzer DNEasy extraction (Qiagen) following manufacturer's instructions starting at with the addition of the lysis solution. Samples were processed in parallel in batches of twelve at a time using the QiaVac 24 Plus Vacuum Manifold (Qiagen). The lysis step (2X at 5000 rpm for 60 seconds each with incubation on ice was performed on the MiniLys (Precellys, Bertin Instruments) using the 0.1 mm glass beads that come with the Powerlyzer kit. DNA concentrations of final preparations were estimated using Quant-iT Picogreen dsDNA Assay Kit (Invitrogen) fluorescence detection on a Spectramax Gemini (Molecular Devices).

DNA from bacterioplankton samples was extracted following [58], and DNA from bacterial cultures was extracted using the DNeasy Blood and Tissue kit (Qiagen) following manufacturer's instructions. All DNA concentrations were estimated using Picogreen.

4.6 *16S rRNA gene sequencing.*

Illumina tag sequencing for the *S. adareanum* microbiome (SaM) targeted the V3-V4 region of the 16S rRNA gene using primers 341F CCTACGGGNBGCASCAG and 806R GGACTACHVGGGTWTCTAAT. The first round of PCR amplified the V3-V4 region using HIFI HotStart Ready Mix (Kapa Biosystems). The first round of PCR used a denaturation temperature of 95 °C for 3 min, 20 cycles of 95 °C for 30 sec, 55 °C for 30 sec and 72 °C for 30 sec and followed by an extension of 72 °C for 5 minutes before holding at 4 °C. The second round of PCR added Illumina-specific sequencing adapter sequences and unique indexes, permitting multiplexing, using the Nextera XT Index Kit v2 (Illumina) and HIFI HotStart Ready Mix (Kapa Biosystems). The second round of PCR used a denaturation temperature of 95 °C for 3 minutes, 8 cycles of 95 ºC for 30 seconds, 55 °C for 30 seconds and 72 °C for 30 seconds and followed by an extension of 72 °C for 5 minutes before holding at 4 °C. Amplicons were cleaned up using AMPure XP beads (Beckman Coulter). A no-template control was processed but did not show a band in the V3-V4 amplicon region and was sequenced for confirmation. A Qubit dsDNA HS Assay (ThermoFisher Scientific) was used for DNA concentration estimates. The average size of the library was determined by the High Sensitivity DNA Kit (Agilent) and the Library Quantification Kit – Illumina/Universal Kit (KAPA Biosystems) quantified the prepared libraries. The amplicon pool sequenced on Illumina MiSeq generated paired end 301 bp reads was demultiplexed using Illumina's bcl2fastq.

The bacterioplankton samples sent to the Joint Genome Institute (JGI) for library preparation and paired end (2 x 250 bp) MiSeq Illumina sequencing of the variable region 4 (V4) using primers 515F (5′-GTGCCAGCMGCCGCGGTAA-3′) and 806R (5′-GGACTACHVGGGTWTCTAAT-3′) [59]. Sequence processing included removal of PhiX contaminants and Illumina adapters at JGI.

The identity of the cultivated isolates was confirmed by16S rRNA gene sequencing using Bact27F and Bact1492R primers either by directly sequencing agarose-gel purified PCR products (Qiagen), or TA cloning (Invitrogen) of PCR fragments into *E. coli* following manufacturer's instructions in which three clones were sequenced for each library, plasmids were purified (Qiagen)

638 at the Nevada Genomics Center, where Sanger sequencing was conducted on an ABI3700 (Applied
639 Biosystems). Sequences were trimmed and quality checked using Sequencer, v. 5.1.

640 4.7 *16S rRNA gene sequencing.*

641 We employed a QIIME2 pipeline [60] using the DADA2 plug-in [61] to de-noise the data and
642 generate amplicon sequence variant (ASV) occurrence matrices for the SaM and bacterioplankton
643 samples. The rigor of ASV determination was used in this instance given the increased ability to
644 uncover variability in the limited geographic study area, interest in uncovering patterns of host-
645 specificity, and ultimately in identifying the conserved, core members of the microbiome, at least
646 one of which may be capable of PalA biosynthesis. Sequence data sets were initially imported into
647 QIIME2 working format and the quality of forward and reverse were checked. Default trimming
648 parameters included trimming all bases after the first quality score of 2, in addition, the first 10
649 bases were trimmed, and reads shorter than 250 bases were discarded. Next the DADA2 algorithm
650 was used to de-noise the reads (corrects substitution and insertion/deletion errors and infers
651 sequence variants). After de-noising, reads were merged. The ASVs were constructed by grouping
652 the unique full de-noised sequences (the equivalent of 100% OTUs, operational taxonomic units).
653 The ASVs were further curated in the QIIME2-DADA2 pipeline by removing chimeras in each
654 sample individually if they can be exactly reconstructed by combining a left-segment and a right-
655 segment from two more abundant "parent" sequences. A pre-trained SILVA 132 99% 16S rRNA
656 Naive Bayes classifier (https://data.qiime2.org/2019.1/common/silva-132-99-nb-classifier.qza) was
657 used to perform the taxonomic classification. Compositions of the SaM and the bacterioplankton
658 ASVs were summarized by proportion at different taxonomy levels, including genus, family, order,
659 class, and phylum ranks. In order to retain all samples for diversity analysis, we set lowest reads
660 frequency per sample (n = 62 samples at 19003 reads; n = 63 samples at 9987 reads) as rarefaction
661 depth to normalize the data for differences in sequence count. ASVs assigned to Eukarya or with
662 unassigned taxa (suspected contaminants) were removed from the final occurrence matrix such that
663 the final matrix read counts were slightly uneven with the lowest number of reads per sample with
664 9961 reads.
665 The SaM ASVs were binned into Core (highly persistent) if present in ≥80% of samples
666 (Core80), Dynamic if present in 50-79% of samples (Dynamic50) and those that comprise the
667 naturally fluctuating microbiome, or Variable fraction that was defined as those ASVs present in
668 <50% of the samples [2, 3]. We used these conservative groupings of the core microbiome due to the
669 low depth of sequencing in our study [3].
670 ASV identities between the SaM, the *S. adareanum* bacterial isolates and the bacterioplankton
671 data sets were compared using CD-HIT (cd-hit-est-2d; http://cd-hit.org). The larger SaM data set
672 which included 19,003 sequences per sample was used for these comparisons to maximize the
673 ability to identify matches; note that this set does exclude one sample, Bon1b which had half as
674 many ASVs, though overall this larger data set includes nearly 200 additional sequences in the
675 Variable fraction for comparison. ASVs with 100 and 97% identity between the pairwise
676 comparisons were summarized in terms of their membership in the Core, Dynamic or Variable
677 fractions of the SaM. Likewise, CD-HIT was used to dereplicate the isolate sequences at a level of
678 99% sequence identity, and then the dereplicated set was compared against the bacterioplankton
679 iTag data set.
680 Phylogenetic analysis of the SaM ASVs, *S. adareanum* bacterial isolates, and 16S rRNA gene
681 cloned sequences from Riesenfeld et al. [13] was conducted with respect to neighboring sequences
682 identified in the Ribosomal Database Project and SILVA and an archaea outgroup using MEGA v.7
683 [62]. Two maximum likelihood trees were constructed; the first with the Core80 ASVs, and the
684 second with both Core80 and Dynamic50 ASVs. A total of 369 aligned positions were used in both
685 trees. A total of 1000 bootstrap replicates were run in both instances, in which the percentage (≥ 50
686 %) of trees in which the associated taxa clustered together are shown next to the branches.

687

688

689 4.8 *Statistical analyses.*

690 T-tests were run in Statistica (v. 13) to determine significance ($p < 0.05$) of site-to-site, within
691 and between colony variation in PalA concentrations. Similarity-matrix and hierarchical clustering
692 analyses were performed using PRIMER v.7 and PERMANOVA+ (PRIMER-e). Analyses were
693 performed on the complete microbiome as well as the three microbiome fractions in most cases. The
694 ASV occurrence data was square root transformed for all analyses. A heat map based on the Core80
695 ASV occurrence was generated with the transformed data, and hierarchical clustering with group
696 average parameter was employed, which was integrated with SIMPROF confidence using 9,999
697 permutations and a 5% significance level. Bray-Curtis resemblance matrixes were created using
698 ASV occurrences without the use of a dummy variable. To determine basic patterns in community
699 structure within vs. between colony and site differences the significance was determined by t-test
700 using Statistica v. 13. To compare within colony (n = 9) vs. between colony differences (n = 27), nine
701 "between colony pairwise similarity values were randomly sampled in order to compare equal
702 sample sizes, checking that the homogeneity of variance was similar between them. Then threshold
703 metric Multi-Dimensional Scaling (tmMDS) was conducted based on Kruskal fit scheme 1,
704 including 500 iterations, and a minimum stress of 0.001. Similarity profile testing through SIMPROF
705 were performed based on a null hypothesis that no groups would demonstrate differences in ASV
706 occurrences. This clustering algorithm was also used to generate confidence levels on the MDS plot,
707 which were set to 65% and 75%. In addition, 95% bootstrap regions were calculated with 43
708 bootstraps per group, set to ensure a minimum rho of 0.99. In order to assess the contribution of
709 each factor to the variance of the microbial community in this nested experimental design,
710 Site(Colony), permutational multivariate analysis of variance (PERMANOVA) was used. Site-based
711 centroids were calculated and the PERMDISP algorithm was used to determine the degree of
712 dispersion around the centroid for each site. Overall site-to-site difference in dispersion was
713 determined and pairwise comparisons were also calculated with 9,999 permutations used to
714 determine significance ($P(perm) < 0.05$). Exploratory analysis of the major ASV contributors to
715 similarity was performed using the SIMPER procedure based on sites and colonies, with a cut off
716 for low contributions set to 70%.
717 Co-occurrence networks were constructed using filtered ASV occurrence data sets in which the
718 ASV were filtered to only those that were present in at least five samples resulting in a 102 ASV
719 data set. The 102 x 63 matrix was provided as input to FlashWeave v1.0 [63] using default
720 parameters, and visualized in Gephi v. 0.9.2 [64]. Then to consider whether the ASVs in the Core80,
721 Dynamic50, or Variable fractions of the SaM were affiliated with particular levels of PalA in the
722 ascidian lobes, PalA niche robust optimum and range were computed using the occurrence and dry
723 weight normalized contextual data [65]. Weighted gene correlation network analysis (WGCNA
724 package in R [66]) was used to identify modules and their correlation with PalA levels. The matrix
725 was total-sum normalized [67], and WGCNA was used in signed mode. There were few modules
726 detected, although they were not correlated with PalA. Modules were projected on the FlashWeave
727 co-occurrence network and called subsystems.

728 4.9 *Biosynthetic gene cluster analysis.*

729 Subsequently, in order to predict the likelihood of Core80 ASV lineages harboring the potential
730 for natural product biosynthesis we designed a meta-analysis of neighboring genomes found at the
731 Integrated Microbial Genomes (IMG) database [68]. The analysis was conducted only for ASVs in
732 which confidence of taxonomic assignment was at the genus level. Therefore, genomes were
733 harvested from IMG that were associated with a total of 9 genera (*Microbulbifer* (16 genomes),
734 *Pseudovibrio* (24 genomes), *Endozoicomonas* (11 genomes), *Nitrosomonas* (19 of 68 total genomes in
735 this genus), *Nitrospira* (14 genomes), *Hoeflea* (7 genomes), *Lutibacter* (12 genomes), *Halocynthilibacter*
736 (2 genomes), and in the case of *Lentimonas*, since no genomes were found, we harvested 8 genomes
737 from the Puniciococcaceae family). This results in 113 genomes that were submitted to antiSMASH
738 [69] for analysis. The genomes and counts of biosynthetic gene clusters assigned to nonribosomal
739 peptide synthase, polyketide synthase, or hybrid of the two classes were tabulated (Table S2).

740 4.10 *Data availability*.

741 *Synoicum adareanum* microbiome Illumina sequence information and associated metadata are
742 described under NCBI BioProject PRJNA597083 (https://www.ncbi.nlm.nih.gov/bioproject/?
743 term=PRJNA597083), and the *S. adareanum* culture collection 16S rRNA gene sequences were
744 deposited in GenBank under MN960541- MN960556. The bacterioplankton sequence information
745 and associated metadata are described under NCBI BioProject PRJNA602715
746 (https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA602715). Records for these Antarctic
747 metadata and associated sequences depositions are also reflected in the mARS database
748 (http://mars.biodiversity.aq).

749 **5. Conclusions**

750 This work has advanced our understanding of the Antarctic ascidian *S. adareanum*, PalA
751 distributions, and microbiome in several ways. First, we found PalA to be a dominant product
752 across all 63 samples, with some variation but no coherent trends with the site, sample, or
753 microbiome ASV. The results point to a conserved, core, microbiome represented by 21 ASVs, 20 of
754 which appear to be distinct from the bacterioplankton. The phylogenetic distribution of these taxa
755 was diverse, and distinct from other ascidian microbiomes in which organisms with both
756 heterotrophic and chemosynthetic lifestyles are predicted. The co-occurrence analysis suggested the
757 potential for ecologically interacting microbial networks that may improve our understanding of
758 this ascidian-microbiome-natural product system. Likewise, based on the occurrence of natural
759 product biosynthetic gene clusters, there are several potential PalA producers. These results
760 advance the long-term goal of *Synoicum adareanum*-palmerolide-microbiome research which is
761 compelled by the fact that by identifying the producer, genome sequencing could then provide
762 information on PalA biosynthesis which could lead to the development of a potential therapeutic
763 agent to fight melanoma.

764 **Supplementary Materials:** The following are available online at www.mdpi.com/xxx/s1, Table S1:
765 PERMANOVA estimators of drivers of variability, Table S2: Taxonomic distribution of bacterioplankton ASVs,
766 Table S3 Biosynthetic gene clusters in bacterial genomes related to Core80 SaM genera, Table S4: *Synoicum*
767 *adareanum* collections and preparations for microbiome cultivation, Table S5 *Synoicum adareanum* collections for
768 palmerolide A and microbiome characterization by V3-4 rRNA gene tag sequencing, Table S6: Bacterioplankton
769 collections used in v4 rRNA gene tag sequencing. Figure S1: Maximum likelihood 16S rRNA gene phylogenetic
770 tree, Figure S2: Results of pairwise t-tests of PalA levels determined by mass spectrometry, Figure S3: Average
771 pairwise similarity within and between *S. adareanum* microbiome community structures, Figure S4: tmMDS plots
772 representing the microbiome of the 63 *S. adareanum* samples, Figure S5. PalA niche optimum for *S. adareanum*
773 microbiome ASVs.

774 **Author Contributions:** This work was the result of a team effort in which the following contributions are
775 recognized: conceptualization, A.E.M., P.S.G.C, and B.J.B. methodology and experimentation, A.E.M., N.E.A.,
776 L.B., D. E., M.L.H., S.K., C.S.R. and R.M.Y.; validation, K.D., M.L.H., and B.J.B. formal analysis, A.E.M., N.E.A.,
777 E.D., D.E., S.K., and C-C.L.; data curation, M.L.H. and C-C.L.; writing—original draft preparation, A.E.M., B.J.B.,
778 N.E.A., D.E., S.K., and C-C.L.; writing—review and editing, A.E.M., N.E.A, B.J.B., L.B., P.S.G.C., A.E.K.D., D.E.,
779 S.K., C.S.R., and R.M.Y.; supervision, A.E.M., B.J.B., P.S.G.C., K.W.D., A.E.K.D., and C.S.R.; project
780 administration, K.W.D.; funding acquisition, A.E.M., P.S.G.C., and B.J.B. All authors have read and agreed to
781 the published version of the manuscript.

782 **Funding:** Support for this research was provided in part by the National Institute of Health award (CA205932)
783 with additional support from National Science Foundation awards (OPP-0442857, ANT-0838776, PLR-1341339
784 to B.J.B. and ANT-0632389 to A.E.M. and Postdoctoral Research Fellowship award DBI-0532893 to C.S.R.).
785 Support for the sequencing of the bacterioplankton was provided as part of the Joint Genome Institute's
786 Community Sequencing Program (JGI-634 to A.E.M.).

787 **Acknowledgments:** The assistance of several collaborators and students is also acknowledged including Charles
788 Amsler, Margaret Amsler, Jason Cuce, Bill Dent, Alex Dussaq, Cheryl Gleasner, Alan Maschek, Robert W. Read,
789 Andrew Shilling, Santana Thomas, and the Palmer Station science support staff.

790 **Conflicts of Interest:** The authors declare no conflict of interest.

References

1. Taylor, M. W.; Tsai, P.; Simister, R. L.; Deines, P.; Botte, E.; Ericson, G.; Schmitt, S.; Webster, N. S., 'Sponge-specific' bacteria are widespread (but rare) in diverse marine environments. *ISME J* **2013**, *7*, 438-443.
2. Ainsworth, T. D.; Krause, L.; Bridge, T.; Torda, G.; Raina, J. B.; Zakrzewski, M.; Gates, R. D.; Padilla-Gamino, J. L.; Spalding, H. L.; Smith, C.; Woolsey, E. S.; Bourne, D. G.; Bongaerts, P.; Hoegh-Guldberg, O.; Leggat, W., The coral core microbiome identifies rare bacterial taxa as ubiquitous endosymbionts. *ISME J* **2015**, *9*, 2261-2274.
3. Hernandez-Agreda, A.; Leggat, W.; Bongaerts, P.; Ainsworth, T. D., The microbial signature provides insight into the mechanistic basis of coral success across reef habitats. *mBio* **2016**, *7*, e00560-16.
4. Burgsdorf, I.; Erwin, P. M.; Lopez-Legentil, S.; Cerrano, C.; Haber, M.; Frenk, S.; Steindler, L., Biogeography rather than association with cyanobacteria structures symbiotic microbial communities in the marine sponge *Petrosia ficiformis*. *Front. Microbiol.* **2014**, *5*, 529.
5. Kelly, L. W.; Williams, G. J.; Barott, K. L.; Carlson, C. A.; Dinsdale, E. A.; Edwards, R. A.; Haas, A. F.; Haynes, M.; Lim, Y. W.; McDole, T.; Nelson, C. E.; Sala, E.; Sandin, S. A.; Smith, J. E.; Vermeij, M. J. A.; Youle, M.; Rohwer, F., Local genomic adaptation of coral reef-associated microbiomes to gradients of natural variability and anthropogenic stressors. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 10227-10232.
6. Pantos, O.; Bongaerts, P.; Dennis, P. G.; Tyson, G. W.; Hoegh-Guldberg, O., Habitat-specific environmental conditions primarily control the microbiomes of the coral *Seriatopora hystrix*. *ISME J* **2015**, *9*, 1916-1927.
7. Lo Guidice, A.; Azzaro, M.; Schiaparelli, S., Microbial symbionts of Antarctic marine benthic invertebrates. In The ecological role of microorganisms in the Antarctic environment, Castro-Sowinski, S., Ed. Springer Polar Sciences, Springer Nature: Switzerland, 2019; pp 277-296.
8. Cardenas, C. A.; Gonzalez-Aravena, M.; Font, A.; Hestetun, J. T.; Hajdu, E.; Trefault, N.; Malmbergg, M.; Bongcarn-Rudloff, E., High similarity in the microbiota of cold-water sponges of the Genus *Mycale* from two different geographical areas. *PeerJ* **2018**, *6*, 10.7717/peerj.4935.
9. Webster, N. S.; Negri, A. P.; Munro, M.; Battershill, C. N., Diverse microbial communities inhabit Antarctic sponges. *Environ. Microbiol.* **2004**, *6*, 288-300.
10. Steinert, G.; Wemheuer, B.; Janussen, D.; Erpenbeck, D.; Daniel, R.; Simon, M.; Brinkhoff, T.; Schupp, P. J., Prokaryotic diversity and community patterns in Antarctic continental shelf sponges. *Front. Mar. Sci.* **2019**, *6*, 297.
11. Webster, N. S.; Bourne, D., Bacterial community structure associated with the Antarctic soft coral, *Alcyonium antarcticum*. *FEMS Microbiol. Ecol.* **2007**, *59*, 81-94.
12. Murray, A. E.; Rack, F. R.; Zook, R.; Williams, M. J. M.; Higham, M. L.; Broe, M.; Kaufmann, R. S.; Daly, M., Microbiome composition and diversity of the ice-dwelling sea anemone, *Edwardsiella andrillae*. *Integr. Comp. Biol.* **2016**, *56*, 542-555.
13. Riesenfeld, C. S.; Murray, A. E.; Baker, B. J., Characterization of the microbial community and polyketide biosynthetic potential in the Palmerolide-producing tunicate, *Synoicum adareanum*. *J. Nat. Prod.* **2008**, *71*, 1812-1818.
14. Palermo, J. A.; Brasco, M. F. R.; Spagnuolo, C.; Seldes, A. M., Illudalane sesquiterpenoids from the soft coral *Alcyonium paessleri*: The first natural nitrate esters. *J. Org. Chem.* **2000**, *65*, 4482-4486.
15. Diyabalanage, T.; Amsler, C. D.; McClintock, J. B.; Baker, B. J., Palmerolide A, a cytotoxic macrolide from the Antarctic tunicate Synoicum adareanum. *J. Am. Chem. Soc.* **2006**, *128*, 5630-5631.
16. Erwin, P. M.; Pineda, M. C.; Webster, N.; Turon, X.; Lopez-Legentil, S., Down under the tunic: bacterial biodiversity hotspots and widespread ammonia-oxidizing archaea in coral reef ascidians. *ISME J* **2014**, *8*, (3), 575-588.
17. Lopez-Legentil, S.; Turon, X.; Espluga, R.; Erwin, P. M., Temporal stability of bacterial symbionts in a temperate ascidian. *Front. Microbiol.* **2015**, *6*, 1022.
18. Evans, J. S.; Erwin, P. M.; Shenkar, N.; Lopez-Legentil, S., Introduced ascidians harbor highly diverse and host-specific symbiotic microbial assemblages. *Sci. Rep.* **2017**, *7*, 11033.
19. Evans, J. S.; Erwin, P. M.; Shenkar, N.; Lopez-Legentil, S., A comparison of prokaryotic symbiont communities in nonnative and native ascidians from reef and harbor habitats. *FEMS Microbiol. Ecol.* **2018**, *94*, fiy139.
20. Cahill, P. L.; Fidler, A. E.; Hopkins, G. A.; Wood, S. A., Geographically conserved microbiomes of four temperate water tunicates. *Environ. Microbiol. Rep.* **2016**, *8*, 470-478.

21. Dror, H.; Novak, L.; Evans, J. S.; Lopez-Legentil, S.; Shenkar, N., Core and dynamic microbial communities of two invasive ascidians: can host-symbiont dynamics plasticity affect invasion capacity? *Microb. Ecol.* **2019**, *78*, 170-184.
22. Erwin, P. M.; Pineda, M. C.; Webster, N.; Turon, X.; Lopez-Legentil, S., Small core communities and high variability in bacteria associated with the introduced ascidian *Styela plicata*. *Symbiosis* **2013**, *59*, 35-46.
23. Blasiak, L. C.; Zinder, S. H.; Buckley, D. H.; Hill, R. T., Bacterial diversity associated with the tunic of the model chordate *Ciona intestinalis*. *ISME J* **2014,** *8*, 309-20.
24. Tianero, M. D. B.; Kwan, J. C.; Wyche, T. P.; Presson, A. P.; Koch, M.; Barrows, L. R.; Bugni, T. S.; Schmidt, E. W., Species specificity of symbiosis and secondary metabolism in ascidians. *ISME J* **2015**, *9*, 615-628.
25. Peng, X.; Adachi, K.; Chen, C. Y.; Kasai, H.; Kanoh, K.; Shizuri, Y.; Misawa, N., Discovery of a marine bacterium producing 4-hydroxybenzoate and its alkyl esters, parabens. *Appl. Environ. Microbiol.* **2006**, *72*, 5556-5561.
26. Fukunaga, Y.; Kurahashi, M.; Tanaka, K.; Yanagi, K.; Yokota, A.; Harayama, S., *Pseudovibrio ascidiaceicola* sp. nov., isolated from ascidians (sea squirts). *Int. J. Syst. Evol. Microbiol.* **2006**, *56*, 343-347.
27. Schreiber, L.; Kjeldsen, K. U.; Funch, P.; Jensen, J.; Obst, M.; Lopez-Legentil, S.; Schramm, A., *Endozoicomonas* are specific, facultative symbionts of sea squirts. *Front. Microbiol.* **2016**, *7*, 1042.
28. Gonzalez, J. M.; Mayer, F.; Moran, M. A.; Hodson, R. E.; Whitman, W. B., *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.* **1997**, *47*, 369-376.
29. Schmidt, E. W.; Donia, M. S., Life in cellulose houses: symbiotic bacterial biosynthesis of ascidian drugs and drug leads. *Curr. Opin. Biotechnol.* 2010, *21*, 827-833.
30. Dou, X.; Dong, B., Origins and bioactivities of natural compounds derived from marine ascidians and their symbionts. *Mar. Drugs* **2019**, *17*, 670.
31. Little, A. E.; Robinson, C. J.; Peterson, S. B.; Raffa, K. F.; Handelsman, J., Rules of engagement: interspecies interactions that regulate microbial communities. *Annu. Rev. Microbiol.* **2008**, *62*, 375-401.
32. Imamura, N.; Nishuma, M.; Takader, T.; Adachi, K.; Sakai, M.; Sano, H., New anticancer antibiotics pelagiomicins, produced by a new marine bacterium *Pelagiobacter variabilis*. *J. Antibiot.* **1997**, *50*, 8-12.
33. Quevrain, E.; Domart-Coulon, I.; Pernice, M.; Bourguet-Kondracki, M. L., Novel natural parabens produced by a *Microbulbifer* bacterium in its calcareous sponge host *Leuconia nivea*. *Environ. Microbiol.* **2009**, *11*, 1527-1539.
34. Martínez-García, M.; Stief, P.; Díaz-Valdés, M.; Wanner, G.; Ramos-Esplá, A.; Dubilier, N.; Antón, J., Ammonia-oxidizing Crenarchaeota and nitrification inside the tissue of a colonial ascidian. *Environ. Microbiol.* **2008**, *10*, 2991-3001.
35. Hatzenpichler, R., Diversity, physiology, and niche differentiation of ammonia-oxidizing archaea. *Appl. Environ. Microbiol.* **2012**, *78*, 7501-10.
36. Murray, A. E.; Preston, C. M.; Massana, R.; Taylor, L. T.; Blakis, A.; Wu, K.; DeLong, E. F., Seasonal and spatial variability of bacterial and archaeal assemblages in the coastal waters off Anvers Island, Antarctica. *Appl. Environ. Microbiol.* **1998**, *64*, 2585-2595.
37. Murray, A. E.; Grzymski, J. J., Diversity and genomics of Antarctic marine micro-organisms. *Phil. Trans. Roy. Soc. B-Bio. Sci.* **2007**, *362*, 2259-2271.
38. Grzymski, J. J.; Riesenfeld, C. S.; Williams, T. J.; Dussaq, A. M.; Ducklow, H.; Erickson, M.; Cavicchioli, R.; Murray, A. E., A metagenomic assessment of winter and summer bacterioplankton from Antarctica Peninsula coastal surface waters. *ISME J* **2012**, *6*, 1901-1915.
39. Rhoades, D., Evolution of plant chemical defense against herbivores. Herbivores: their interaction with secondary plant metabolites. Academic Press: New York, 1979.
40. McPhail, K. L.; Davies-Coleman, M. T.; Starmer, J., Sequestered chemistry of the Arminacean nudibranch *Leminda millecra* in Algoa Bay, South Africa. *J. Nat. Prod.* **2001**, *64*, 1183-1190.
41. Carbone, M.; Gavagnin, M.; Haber, M.; Guo, Y. W.; Fontana, A.; Manzo, E.; Genta-Jouve, G.; Tsoukatou, M.; Rudman, W. B.; Cimino, G.; Ghiselin, M. T.; Mollo, E., Packaging and delivery of chemical weapons: a defensive trojan horse stratagem in chromodorid nudibranchs. *PLoS One* **2013**, *8*, e62075.
42. Winters, A. E.; White, A. M.; Dewi, A. S.; Mudianta, I. W.; Wilson, N. G.; Forster, L. C.; Garson, M. J.; Cheney, K. L., Distribution of defensive metabolites in nudibranch molluscs. *J. Chem. Ecol.* **2018**, *44*, 384-396.

43. Schupp, P.; Eder, C.; Paul, V.; Proksch, P., Distribution of secondary metabolites in the sponge *Oceanapia sp*. and its ecological implications. *Mar. Biol.* **1999**, *135*, 573-580.
44. Freeman, C. J.; Gleason, D. F., Chemical defenses, nutritional quality, and structural components in three sponge species: *Ircinia felix*, *I. campana*, and *Aplysina fulva*. *Mar. Biol.* **2010**, *157*, 1083-1093.
45. Roue, M.; Domart-Coulon, I.; Ereskovsky, A.; Dejediat, C.; Pererz, T.; Bourguet-Kondracki, M.-L., Cellular localization of clathridimine, an antimicrobial 2-aminoimidazole alkaloid produced by the Mediterranean calcareous sponge *Clathrina clathrus*. *J. Nat. Prod.* **2010**, *73*, 1277-1282.
46. Furrow, F. B.; Amsler, C. D.; McClintock, J. B.; Baker, B. J., Surface sequestration of chemical feeding deterrents in the Antarctic sponge *Latrunculia apicalis* as an optimal defense against sea star spongivory. *Mar. Biol*. **2003**, *143*, 443-449.
47. Becerro, M. A.; Paul, V. J.; Starmer, J., Intracolonial variation in chemical defenses of the sponge *Cacospongia* sp. and its consequences on generalist fish predators and the specialist nudibranch predator *Glossodoris pallida*. Mar. Ecol. Prog. Ser 1998, 168, 187-196.
48. Siriak, T.; Intaraksa, N.; Kaewsuwan, S.; Yuenyongsawad, S.; Suwanborirux, K.; Plubrukarn, A., Intracolonial allocation of tisoxazole macrolides in the sponge *Pachastrissa* nux. *Chem. Biodivers.* **2011**, *8*, 238-246.
49. Richardson, A. D.; Aalbersberg, W.; Ireland, C. M., The patellazoles inhibit protein synthesis at nanomolar concentrations in human colon tumor cells. *Anticancer. Drugs* **2008**, *16*, 533-541.
50. Kwan, J. C.; Donia, M. S.; Han, A. W.; Hirose, E.; Haygood, M. G.; Schmidt, E. W., Genome streamlining and chemical defense in a coral reef symbiosis. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 20655-20660.
51. Gouiffes, D.; Juge, M.; Grimaud, N.; Welin, L.; Sauviat, M.; Barbin, Y.; Laurent, D.; Roussakis, C.; Henichart, J.; Verbist, J., Bistramide A, a new toxin from the urochordata *Lissoclinum bistratum* Sluiter: isolation and preliminary characterization. *Toxicon* **1988**, *26*, 1129-1136.
52. Schmidt, E. W., The secret to a successful relationship: lasting chemistry between ascidians and their symbiotic bacteria. *Invertebr*. Biol. **2015**, *134*, 88-102.
53. Subramani, R.; Aalbersberg, W., Culturable rare Actinomycetes: diversity, isolation and marine natural product discovery. *Appl. Microbiol. Biotechnol.* **2013**, *97*, 9291-321.
54. Ziemert, N.; Lechner, A.; Wietz, M.; Millan-Aguinaga, N.; Chavarria, K. L.; Jensen, P. R., Diversity and evolution of secondary metabolism in the marine actinomycete genus *Salinispora*. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, E1130-E1139.
55. Lopera, J.; Miller, I. J.; McPhail, K. L.; Kwan, J. C., Increased biosynthetic gene dosage in a genome-reduced defensive bacterial symbiont. *mSystems* **2017**, *2*, e00096-17.
56. Holmstrom, C.; James, S.; Neilan, B. A.; White, D. C.; Kjelleberg, S., *Pseudoalteromonas tunicata* sp. nov., a bacterium that produces antifouling agents. *Int. J. Syst. Bacteriol.* **1998**, *48*, 1205-1212.
57. Baker, B. J., *Synoicum adareanum* sampling u/w video Mar 2011 Palmer Station Antarctica. Dryad, Ed. *Dryad Data*, **2020**. *10.5061/dryad.gxd2547gw*.
58. Massana, R.; Murray, A. E.; Preston, C. M.; DeLong, E. F., Vertical distribution and phylogenetic characterization of marine planktonic Archaea in the Santa Barbara Channel. *Appl. Environ. Microbiol.* **1997**, *63*, 50-56.
59. Caporaso, J. G.; Lauber, C.; Walters, W.; Berg-Lyons, D.; Lozupone, C.; Turnbaugh, P.; Fierer, N.; Knight, R., Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 4516-4522.
60. Bolyen, E.; Rideout, J. R.; Dillon, M. R.; Bokulich, N.; Abnet, C. C.; Al-Ghalith, G. A.; Alexander, H.; Alm, E. J.; Arumugam, M.; Asnicar, F.; Bai, Y.; Bisanz, J. E.; Bittinger, K.; Brejnrod, A.; Brislawn, C. J.; Brown, C. T.; Callahan, B. J.; Caraballo-Rodriguez, A. M.; Chase, J.; Cope, E. K.; Da Silva, R.; Diener, C.; Dorrestein, P. C.; Douglas, G. M.; Durall, D. M.; Duvallet, C.; Edwardson, C. F.; Ernst, M.; Estaki, M.; Fouquier, J.; Gauglitz, J. M.; Gibbons, S. M.; Gibson, D. L.; Gonzalez, A.; Gorlick, K.; Guo, J. R.; Hillmann, B.; Holmes, S.; Holste, H.; Huttenhower, C.; Huttley, G. A.; Janssen, S.; Jarmusch, A. K.; Jiang, L. J.; Kaehler, B. D.; Bin Kang, K.; Keefe, C. R.; Keim, P.; Kelley, S. T.; Knights, D.; Koester, I.; Kosciolek, T.; Kreps, J.; Langille, M. G. I.; Lee, J.; Ley, R.; Liu, Y. X.; Loftfield, E.; Lozupone, C.; Maher, M.; Marotz, C.; Martin, B. D.; McDonald, D.; McIver, L. J.; Melnik, A. V.; Metcalf, J. L.; Morgan, S. C.; Morton, J. T.; Naimey, A. T.; Navas-Molina, J. A.; Nothias, L. F.; Orchanian, S. B.; Pearson, T.; Peoples, S. L.; Petras, D.; Preuss, M. L.; Pruesse, E.; Rasmussen, L. B.; Rivers, A.; Robeson, M. S.; Rosenthal, P.; Segata, N.; Shaffer, M.; Shiffer, A.; Sinha, R.; Song, S. J.; Spear, J. R.; Swafford, A. D.; Thompson, L. R.; Torres, P. J.; Trinh, P.; Tripathi, A.; Turnbaugh, P.

J.; Ul-Hasan, S.; vander Hooft, J. J. J.; Vargas, F.; Vazquez-Baeza, Y.; Vogtmann, E.; von Hippel, M.; Walters, W.; Wan, Y. H.; Wang, M. X.; Warren, J.; Weber, K. C.; Williamson, C. H. D.; Willis, A. D.; Xu, Z. Z.; Zaneveld, J. R.; Zhang, Y. L.; Zhu, Q. Y.; Knight, R.; Caporaso, J. G., Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **2019**, *37*, 852-857.

61. Callahan, B. J.; McMurdie, P. J.; Rosen, M. J.; Han, A. W.; Johnson, A. J. A.; Holmes, S. P., DADA2: High-resolution sample inference from Illumina amplicon data. *Nat. Methods* **2016**, 13, 581-583.
62. Kumar, S.; Stecher, G.; Tamura, K., MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger data sets. *Mol. Biol. Evol.* **2016**, *33*, 1870-1874.
63. Tackmann, J.; Matias Rodrigues, J. F.; von Mering, C., Rapid inference of direct interactions in large-scale ecological networks from heterogeneous microbial sequencing data. *Cell Syst.* **2019**, *9*, 286-296 e8.
64. Bastian, M.; Heymann, S.; Jacomy, M. In Gephi: an open source software for exploring and manipulating networks. , International AAAI ICWSM Conference on weblogs and social media, San Jose, California, USA, 2009; San Jose, California, USA.
65. Cristóbal, E.; Ayuso, S. V.; Justel, A.; Toro, M., Robust optima and tolerance ranges of biological indicators: a new method to identify sentinels of global warming. *Ecol. Res.* **2013**, *29*, 55-68.
66. Langfelder, P.; Horvath, S., WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **2008**, *9*, 559.
67. Paulson, J. N.; Stine, O. C.; Bravo, H. C.; Pop, M., Differential abundance analysis for microbial marker-gene surveys. *Nat. Methods* **2013**, *10*, 1200-2.
68. Markowitz, V. M.; Chen, I.-M. A.; Palaniappan, K.; Chu, K.; Szeto, E.; Grechkin, Y.; Ratner, A.; Jacob, B.; Huang, J.; Williams, P.; Huntemann, M.; Anderson, I.; Mavromatis, K.; Ivanova, N. N.; Kyrpides, N. C., IMG: the integrated microbial genomes database and comparative analysis system. *Nucleic Acids Res.* **2012**, *40*, D115-22.
69. Blin, K.; Medema, M. H.; Kottmann, R.; Lee, S. Y.; Weber, T., The antiSMASH database, a comprehensive database of microbial secondary metabolite biosynthetic gene clusters. *Nucleic Acids Res.* **2017**, *45*, D555-D559.