

RESEARCH

Open Access



Geography, not lifestyle, explains the population structure of free-living and host-associated deep-sea hydrothermal vent snail symbionts

Michelle A. Hauer¹, Corinna Breusing¹, Elizabeth Trembath-Reichert², Julie A. Huber³ and Roxanne A. Beinart^{1*}

Abstract

Background Marine symbioses are predominantly established through horizontal acquisition of microbial symbionts from the environment. However, genetic and functional comparisons of free-living populations of symbionts to their host-associated counterparts are sparse. Here, we assembled the first genomes of the chemoautotrophic gammaproteobacterial symbionts affiliated with the deep-sea snail *Alviniconcha hessleri* from two separate hydrothermal vent fields of the Mariana Back-Arc Basin. We used phylogenomic and population genomic methods to assess sequence and gene content variation between free-living and host-associated symbionts.

Results Our phylogenomic analyses show that the free-living and host-associated symbionts of *A. hessleri* from both vent fields are populations of monophyletic strains from a single species. Furthermore, genetic structure and gene content analyses indicate that these symbiont populations are differentiated by vent field rather than by lifestyle.

Conclusion Together, this work suggests that, despite the potential influence of host-mediated acquisition and release processes on horizontally transmitted symbionts, geographic isolation and/or adaptation to local habitat conditions are important determinants of symbiont population structure and intra-host composition.

Keywords Symbiosis, Hydrothermal vents, Population genomics, *Alviniconcha*, Mariana Back-Arc, Microbial biogeography

Introduction

Mutualistic animal-microbe associations are globally significant phenomena, shaping the ecology and evolution of both host animals and microbial symbionts [1]. These

symbiotic associations are maintained by transmission of symbionts from host parent to progeny either [1] directly, for example via the germline (vertical transmission), [2] indirectly, for example through an environmental population of symbionts (hereafter referred to as “free-living” symbionts) (horizontal transmission), or [3] via a combination of both vertical and horizontal transmission (mixed mode transmission) [2].

Horizontal transmission is more commonly found in aquatic than terrestrial habitats, likely due to the ease with which microbes can be transported in water compared to air or soil [3]. However, even for marine symbioses where horizontally transmitted microbial symbionts are observed in the environment [4], it is not yet clear

*Correspondence:

Roxanne A. Beinart
rbeinart@uri.edu

¹ Graduate School of Oceanography, University of Rhode Island, Narragansett, RI, USA

² School of Earth and Space Exploration, Arizona State University, Tempe, AZ, USA

³ Department of Marine Chemistry and Geochemistry, Woods Hole Oceanographic Institution, Woods Hole, Falmouth, MA, USA



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

whether free-living, environmental populations of symbionts represent host-associated populations at the strain level, or whether their diversity and composition differs. Free-living symbiont populations may be shaped by local environmental conditions as well as the dynamic interactions with their host—for example, host animals may “seed” the environment by the release of their symbionts into the water column only upon host death [5] or via continuous release from live adults [6, 7]. In addition, ecological and evolutionary processes, such as dispersal barriers, natural selection, and genetic drift, can contribute to the diversity and biogeography of environmental symbionts [8, 9].

Deep-sea hydrothermal vents are discontinuous, island-like habitats dominated by vent-endemic invertebrates that host primarily horizontally transmitted chemoautotrophic bacterial symbionts, making them opportune natural systems for understanding the biogeography of free-living microbial symbionts. In these mutualisms, the symbiotic bacteria are either obtained during a narrow competence window in early developmental stages or throughout the lifetime of the host [10, 11] and are, in most cases, housed intracellularly within the host’s tissues, e.g., gill or trophosome. The symbionts oxidize chemical reductants (e.g., H_2S , H_2 , CH_4) in venting fluids to generate energy for the production of organic matter, thereby providing the primary food source for the host in an otherwise oligotrophic deep ocean [12] and accounting for the high ecosystem productivity characteristic of hydrothermal vents [13–15].

Despite reliance on horizontal transmission, the majority of host species from hydrothermal vents affiliate with only one or two specific endosymbiont phylotypes (i.e., species or genera based on 16S rRNA gene sequence similarity) [12], possibly as a means to reduce the acquisition of cheaters [16]. While a significant number of studies have focused on the diversity, composition and structure of the host-associated symbiont populations (e.g., [10, 17–22]), their free-living, environmental stages remain poorly investigated [4], partly due to the difficulty of detecting low abundance free-living symbionts in environmental samples. As a consequence, few free-living symbiont studies exist. Most of these studies have so far relied on investigations of particular marker genes [4, 23]; only one used an -omics approach but was limited to a single metagenome [24].

A recent shotgun metagenomic study found putative free-living symbiont populations of the provannid snail *Alviniconcha hessleri* in low-temperature diffuse venting fluids at two distinct vent fields of the Mariana Back-Arc, Northwest Pacific (15.5–18° N), [25], providing a unique opportunity to compare free-living and host-associated stages of chemosynthetic symbionts at hydrothermal

vents. *Alviniconcha hessleri* belongs to the dominant fauna at hydrothermal vents in the Mariana Back-Arc Basin, where it lives in nutritional endosymbiosis with one species of sulfur-oxidizing, environmentally acquired Gammaproteobacteria [26, 27]. Although patterns of host-symbiont phylogenetic discordance strongly support a mode of horizontal transmission for the *A. hessleri* symbiont [26, 27], the exact dynamics of symbiont uptake and release are unknown. As an endemic species to the Mariana region, *A. hessleri* is currently listed as “Vulnerable” on the International Union for Conservation of Nature Red List of Threatened Species (<https://www.iucnredlist.org>), highlighting the need to identify the factors that contribute to its limited biogeographic range, including the population structure of its obligate microbial symbiont.

In this study, we applied phylogenomic and population genomic methods to evaluate the evolutionary relationships as well as the genetic and functional variation of *Alviniconcha hessleri* symbionts based on lifestyle by comparing free-living and host-associated symbiont populations collected from the same habitats. In addition, we addressed the effect of geography by comparing populations of both host-associated and free-living symbionts between vent fields of the northern and central Mariana Back-Arc Basin that are approximately 300 km apart and differ notably in their geochemistry: the central vent sites are known to support both low-temperature diffuse flow and black smokers that emit high-temperature fluids, with high amounts of hydrogen sulfide (H_2S), whereas the northern sites only harbor diffuse flow habitats with lower concentrations of H_2S [25].

Methods

Host-associated symbiont collection, sequencing, and genome assemblies

Three *A. hessleri* specimens each were collected from snail beds at the Illium vent field (3582 m) and the Voodoo Crater-2 (VC2) location within the Hafa Adai vent field (3277 m) in the Mariana Back-Arc Basin (Fig. 1) using the remotely operated vehicle *SuBastian* on board the R/V *Falkor* in 2016. Symbiont-bearing snail gill tissues were dissected and stored in RNALater™ (Thermo Fisher Scientific, Inc.) at $-80\text{ }^\circ\text{C}$ until DNA extraction with the Zymo Quick DNA 96 Plus and ZR-96 Clean-up kits (Zymo Research, Inc.). High-throughput Illumina 150-bp paired-end libraries for all six samples were prepared and sequenced by Novogene Corporation, Inc. (Beijing, China) with an average yield of 58 million total Illumina reads (Supplementary Table 1). In addition, one gill sample from each vent field—Hafa Adai 172 (VC2) and Illium 13—was selected for long-read Nanopore sequencing on 2–3 MinION flow cells (Oxford Nanopore

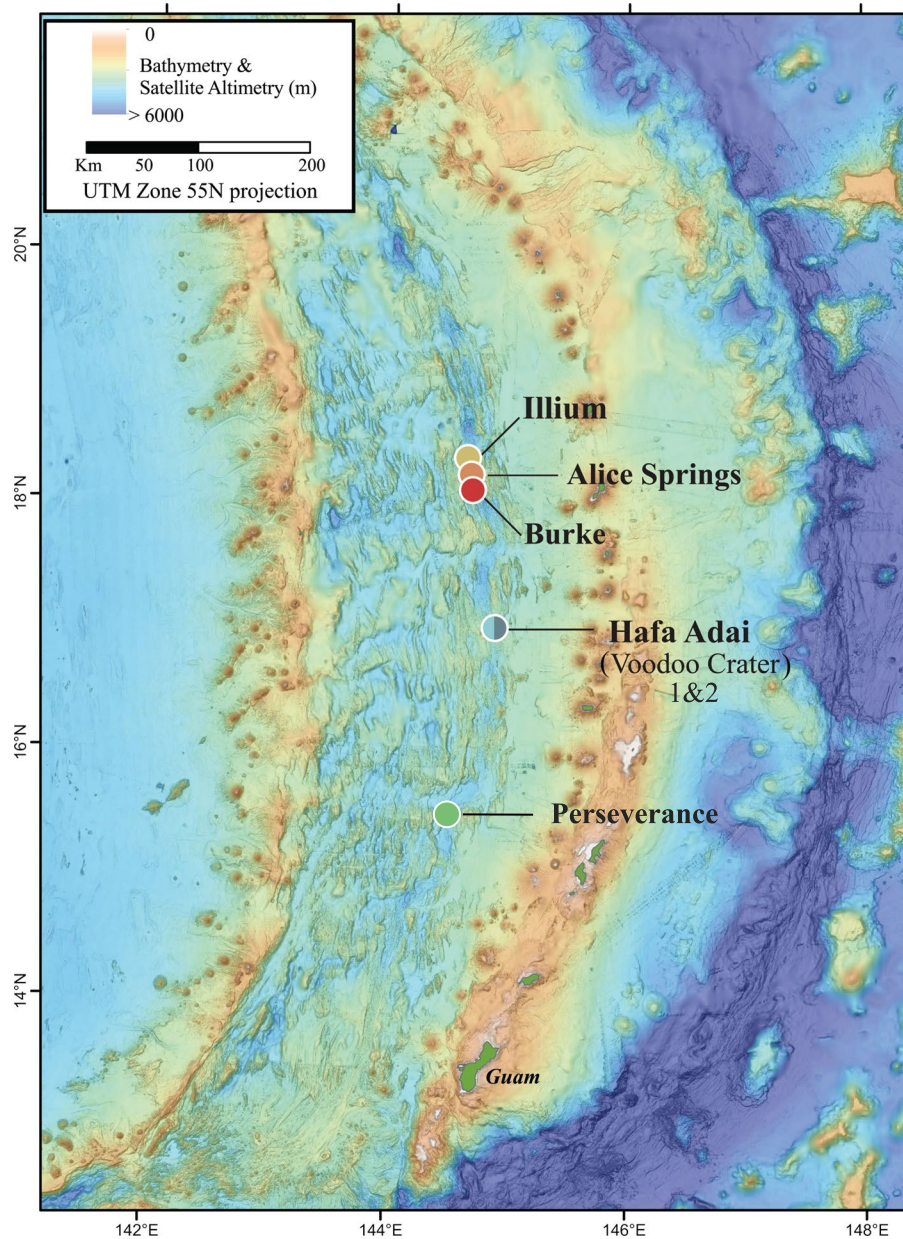


Fig. 1 Map of the Mariana Back-Arc Basin adapted from ref. [25], indicating the locations of the two main hydrothermal vent fields sampled in this study, Illium and Hafa Adai. Free-living symbiont samples were obtained at low coverage from two additional vent sites investigated in ref. [25], Alice Springs and Burke. Colors associated with vent fields are used consistently across samples. A light cyan and teal are used for Hafa Adai, representing Voodoo Crater 1 and 2, respectively

Technologies, Oxford, UK) using the SQK-LSK109 ligation kit (Supplementary Table 1). Financial constraints prevented the ability to perform long-read sequencing on all host-associated samples.

Raw Illumina reads were trimmed with Trimmomatic v0.36 [28], and common sequence contaminants were removed by mapping against the PhiX and human (GRCh38) reference genomes. Nanopore reads were

base-called with Albacore (Oxford Nanopore Technologies, Oxford, UK) and adapter-clipped with Porechop (<https://github.com/rrwick/Porechop>). For Illium, symbiont Illumina and Nanopore reads were extracted through mapping against a draft co-assembly constructed with MEGAHIT [29] and reassembled with SPAdes v3.13.1 [30] using kmers between 21 and 91 in increments of 10. The Hafa Adai 172 MEGAHIT assembly was low quality;

therefore, for Hafa Adai 172, all trimmed Illumina and Nanopore reads were assembled with metaSPAdes v3.13.1 [31] using the same parameters. Manual binning of the assemblies was conducted with GBTools [32], and contigs <200 bp and <500 bp were excluded from the Illium and Hafa Adai-VC2 assemblies, respectively. Assemblies were then scaffolded with SSPACE-Standard v3.0 [33] and SLR [34] and gapfilled with GapFiller v1-10 [35] and LR_GAPCLOSER [36]. Final assemblies were polished with Pilon [37]. Trimmed Illumina reads for the remaining samples—Illum 11, Illum 17, Hafa Adai 60, and Hafa Adai 64—were assembled with metaSPAdes v3.13.1 as described above and binned with MaxBin [38]. All metagenome-assembled genomes (MAGs) were quality checked with Quast v5.0.2 [39] and CheckM v1.0.18 [40] and taxonomically assigned through the Genome Taxonomy Database toolkit [41].

Free-living symbiont collection, sequencing, and genome assembly

All sequences from free-living samples used here were retrieved from a previous study [25], including two high-quality MAGs of environmental *A. hessleri* symbionts from the Illum (GCA_003972985.1) and Hafa Adai-VC2 (GCA_003973075.1) vent sites. The fluid samples from which these MAGs were assembled were collected in direct vicinity of snail beds where the *A. hessleri* specimens for host-associated analyses were obtained [25]. We further included a third previously assembled free-living symbiont MAG from diffuse venting fluids at the Voodoo Crater-1 (VC1) (GCA_003973045.1) location within the Hafa Adai vent field, ~5 m from Hafa Adai-VC2. Though symbiont MAGs were not previously able to be assembled from the other vent sites sampled in ref. [25]—Burke, Alice Springs, Perseverance, and Hafa Adai-Alba—we attempted again to retrieve symbiont MAGs from these samples by assembling and binning the raw reads from these sites with our methods described above, but did not produce usable symbiont MAGs.

All details of the hydrothermal fluid collection, sample storage, sample processing, sequencing, assembly, and binning of metagenome-assembled genomes can be found in ref. [25]. Information about raw sequencing reads is provided in Supplementary Table 1.

Genome similarity and phylogenomic analyses

To confirm that all symbiont MAGs belong to the same bacterial species, we calculated average nucleotide identities (ANIs) via FastANI [42]. A phylogenomic tree that included the six host-associated and the three free-living symbiont MAGs as well as reference genomes of other chemosynthetic Gammaproteobacteria (Supplementary Table 2) was then constructed with IQ-TREE2 [43] based

on 70 single-copy core genes in the Bacteria_71 collection [44]. Parameter choice for phylogenomic reconstructions followed ref. [18].

Population structure and gene content analysis

To determine symbiont population structure according to geography and lifestyle, we inferred DNA sequence polymorphisms in the free-living and host-associated samples by mapping metagenomic reads to a pangenome created with Panaroo [45] from all nine symbiont MAGs. Variants were called and filtered following the pipeline in ref. [18]. All samples from Illum and Hafa Adai met our minimum 10× coverage threshold. Free-living samples from Burke and Alice Springs mapped at 5.9× and 3.9× coverage, respectively. The population structure and gene content analyses (see below) were repeated for these lower-coverage samples. All other free-living metagenomic samples from the remaining vent sites collected in ref. [25] (i.e., Perseverance, Hafa Adai-Alba) had an insufficient number of reads mapped for further analyses. Principal coordinate analysis (PCoA) plots were created based on nucleotide counts converted to Bray–Curtis dissimilarities with the ggplot2 [46] and vegan [47] packages in Rstudio [48]. To quantify the qualitative variant calling results depicted in the PCoAs, fixation indices (F_{ST}) between individual metagenomic samples (wherein each individual gill metagenome was treated as a population) were calculated following ref. [49] and plotted with pheatmap [50]. The method from ref. [49] as well as scikit-allel (<https://github.com/cggh/scikit-allel>) were further used to calculate pairwise F_{ST} values between samples pooled by lifestyle or vent field.

Gene content variation among symbiont populations was determined via Pangenome-based Phylogenomic Analysis (PanPhlAn) [51] and visualized through a PCoA plot based on the Jaccard Similarity Coefficient. Genes that were uniquely associated with lifestyle and vent field, respectively, were extracted from the PanPhlAn gene presence/absence matrix. Functional predictions for these genes were either obtained from the Prokka [52] annotations created during pangenome construction or inferred by blasting the respective protein sequences against the NR database. Hypothetical and unknown proteins were further annotated via KEGG [53] and AlphaFold [54]. Differences in gene content between symbiont populations were visualized through Likert plots with the HH package [55] in RStudio.

Validation of free-living symbiont populations

To gain confidence that the symbionts detected in our environmental samples represented truly “free-living” symbiont stages as opposed to symbionts associated with host larvae or shed gill cells, we calculated the ratio of

symbiont 16S rRNA genes to host mitochondrial CO1 genes in all nine samples by mapping raw metagenomic reads from the snail gills to custom-generated *Alviniconcha* symbiont 16S rRNA and host mtCO1 gene databases. To account for false positive mappings, we created additional background databases consisting of select bacterial (SUP05 clade bacteria, *Thiomicrospira*, and *Marinomonas*) and mollusk gene sequences. Bacterial 16S rRNA genes were downloaded from SILVA [56], while all *Alviniconcha* and mollusk mtCO1 genes were downloaded from BOLD [57]. BBSplit (<https://sourceforge.net/projects/bbmap/>) was then used to separate *Alviniconcha* symbiont and host reads based on the taxon-specific and background 16S rRNA and CO1 gene databases.

Results

Free-living and host-associated symbionts belong to the same bacterial species

Our analysis included nine *A. hessleri* symbiont MAGs from the Illium and Hafa Adai vent fields: six host-associated symbiont genomes assembled in this study, and three previously published, free-living symbiont genomes from the diffuse venting fluids around *A. hessleri* beds [25] (Supplementary Table 3). All host-associated and two of the three free-living MAGs were of very high quality, with >90% completeness and <3% contamination. The third free-living MAG, Hafa Adai-VC1, had a medium quality (~67% completeness). ANI values between all MAGs were >97.7% (Supplementary Table 4), suggesting that the nine *A. hessleri* symbiont genomes belong to the same bacterial species [42] within the genus *Thiolapillus* based on the Genome Taxonomy Database, and confirming that the previously assembled free-living symbiont genomes were indeed *A. hessleri* symbionts (Supplementary Table 3). Corroborating the ANI results, the nine *A. hessleri* MAGs were monophyletic in our phylogenomic analysis relative to the gammaproteobacterial symbionts of other vent invertebrates (Fig. 2). In agreement with phylogenetic analyses of the 16S rRNA gene [27], the nearest neighbors of the *A. hessleri* symbionts were the *Ifremeria nautilei* SOX symbiont, as well as *Thiolapillus brandeum*, a microbe not known to be symbiotic [58].

Environmental samples contain free-living symbiont populations

To investigate whether the symbionts observed in the diffuse flow samples were true free-living symbionts rather than symbionts associated with *A. hessleri* larvae or shed gill tissue, we calculated the ratio of symbiont 16S rRNA gene to host mitochondrial CO1 gene reads in all nine environmental and host-associated samples (Table 1). If the environmental symbiont samples were associated

with larvae or host tissue debris/cells, we expect the ratio in the environmental and host-associated samples to be similar to one another. However, the 16S rRNA:mtCO1 ratio was consistently orders of magnitude higher in environmental samples than in host-associated samples, indicating the presence of a population of symbiont cells independent from host tissue. This finding provides evidence that our environmental samples include truly free-living *A. hessleri* symbiont populations.

A. hessleri symbiont populations are structured primarily by vent field, not lifestyle

Our genome assemblies from both host tissue and diffuse vent fluids likely represent the dominant symbiont strain in each sample, but do not reveal the full extent of strain-level population variation between samples. To determine whether *A. hessleri* symbionts form subpopulations consistent with geography or lifestyle, we created a pangenome out of the individual symbiont MAGs from the Illium and Hafa Adai vent fields that we used as reference for variant calling (Supplementary Tables 5, 6). Our variant detection method resulted in 2177 sequence polymorphisms for investigation of population genomic structure based on F_{ST} and ordination analyses (Fig. 3).

F_{ST} values were calculated pairwise between all nine populations (Fig. 3a), as well as between samples pooled by lifestyle and vent field. F_{ST} values range from 0 to 1, where an F_{ST} value of 1 indicates that samples form genetically isolated subpopulations, while an F_{ST} value of 0 indicates that the samples form a single, well-mixed population. Between individual samples, pairwise F_{ST} s showed a moderate (0.2–0.5) to strong (>0.5) differentiation, indicating that all samples represent distinct subpopulations with limited genetic exchange among each other (Fig. 3a, Supplementary Table 7). Genetic isolation among individual samples was typically stronger between (0.54–0.76) than within (0.21–0.46) vent fields (i.e., Illium vs Hafa Adai-VC2). Within vent sites, the degree of differentiation was comparable among samples independent of lifestyle at Illium, while host-associated samples were more similar to one another than to free-living samples at Hafa Adai-VC2. When samples were pooled, overall pairwise F_{ST} values were markedly higher by vent field (0.47 ± 0.03 s.d.) than by lifestyle (0.05 ± 0.01 s.d.). The dominant effect of geography on symbiont population structure was supported by PCoAs where both free-living and host-associated samples from Illium clustered distinctly from Hafa Adai (VC1 and VC2) (Fig. 3b). Despite the fact that Hafa Adai-VC1 and -VC2 differ spatially by only ~5 m, the free-living VC1 sample formed its own distinct subpopulation from both host-associated and free-living populations at VC2 (F_{ST} 0.58–0.67),

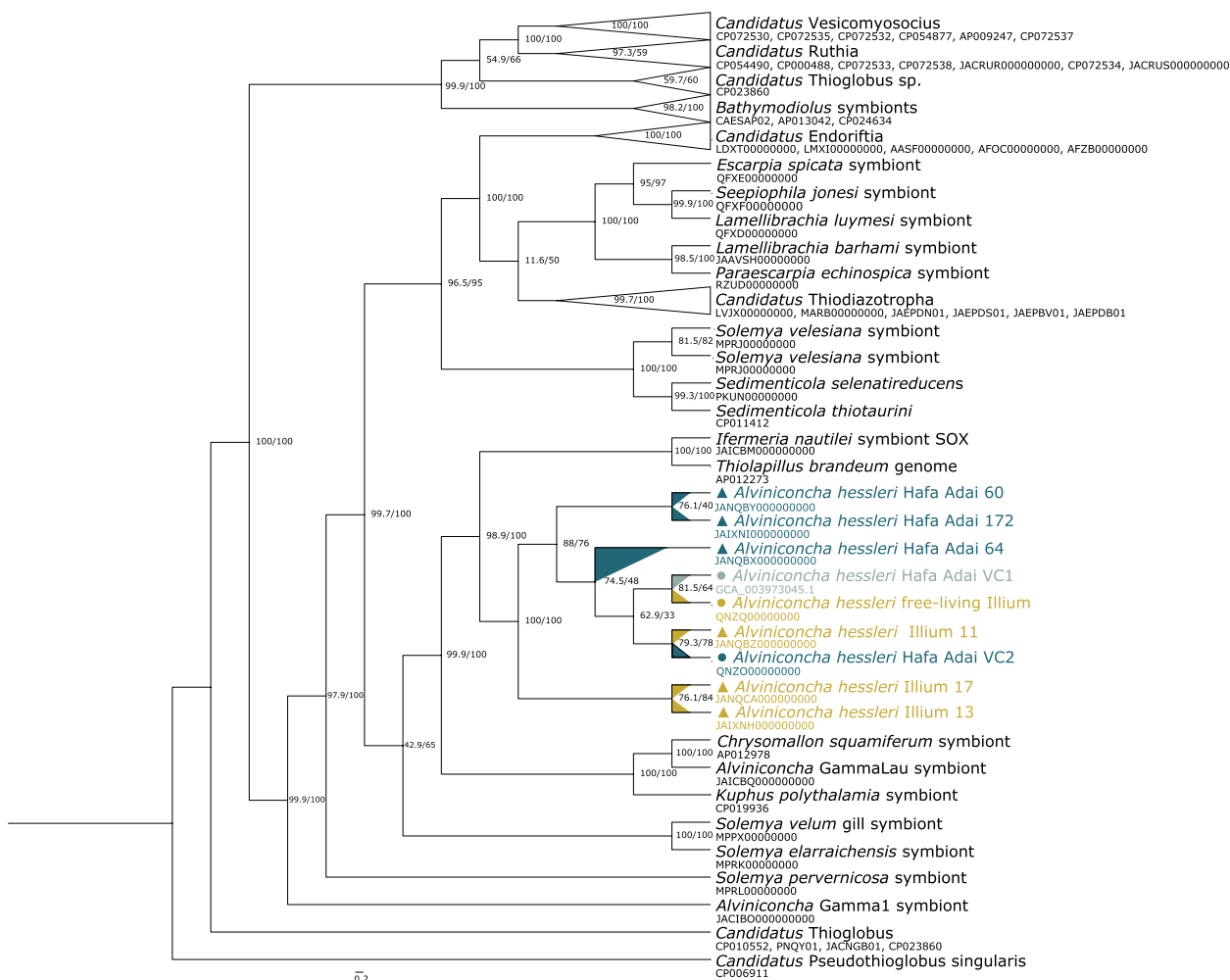


Fig. 2 Cladogram branch transformed phylogenomic tree of chemosynthetic Gammaproteobacteria based on 70 single-copy core genes. Genome accession numbers for all genomes included in the phylogenomic analysis can be found in Supplementary Table 1. *A. hessleri* symbionts are colored by vent field (Hafa Adai VC2: teal, Hafa Adai VC1: light cyan, Illium: yellow) with dots and triangles indicating free-living or host-associated lifestyle. *Candidatus Pseudothioglobus singularis* was used as outgroup for tree rooting

suggesting very fine-scale geographic or environmental structuring.

These patterns were consistent in analyses based on 1271 and 793 variant sites that included the free-living, low-coverage symbiont samples from Burke and Alice Springs, respectively (Supplementary Fig. 1, 4; Supplementary Tables 8, 9). Burke represented the most divergent population, reaching F_{ST} values > 0.8 in all pairwise comparisons. Although Alice Springs clustered closely with free-living and host-associated symbionts from Illium in the PCoAs, F_{ST} values indicated a high degree of genetic isolation for this population (F_{ST} > 0.7). Analyses with samples pooled by vent field confirmed patterns of strong genetic differentiation between geographic

locations without evidence for isolation-by-distance (Supplementary Table 10).

A. hessleri symbiont gene content differs by vent field, not lifestyle

Gene content variation between symbiont populations was assessed based on lifestyle and geography. Similar to the population structure analyses, PCoA plots based on gene content variation across all nine host-associated and free-living populations revealed clustering by vent field but not by lifestyle: symbiont populations from Hafa Adai-VC1 and Hafa Adai-VC2 were more similar to one another than to Illium (Fig. 4), although Hafa Adai-VC1 clustered as an independent population from all other samples.

Table 1 Percentage of unambiguous reads in environmental samples mapping to *Alviniconcha* symbiont 16S rRNA and mitochondrial CO1 genes, as well as the ratio of 16S rRNA to mtCO1

Sample	<i>Alviniconcha</i> 16S rRNA	<i>Alviniconcha</i> mtCO1	16S rRNA:CO1 Origin
Host-associated Hafa Adai 172	0.019%	0.0032%	5.83 Gill tissue
Host-associated Hafa Adai 60	0.016%	0.0111%	1.46 Gill tissue
Host-associated Hafa Adai 64	0.017%	0.0102%	1.67 Gill tissue
Host-associated Illium 11	0.011%	0.0084%	1.32 Gill tissue
Host-associated Illium 13	0.020%	0.0080%	2.50 Gill tissue
Host-associated Illium 17	0.014%	0.0072%	1.99 Gill tissue
Free-living Illium	0.043%	0.0022%	19.73 Diffuse fluids
Free-living Hafa Adai-VC2	0.084%	0.0019%	43.54 Diffuse fluids
Free-living Hafa Adai-VC1	0.065%	0.0003%	241.70 Diffuse fluids

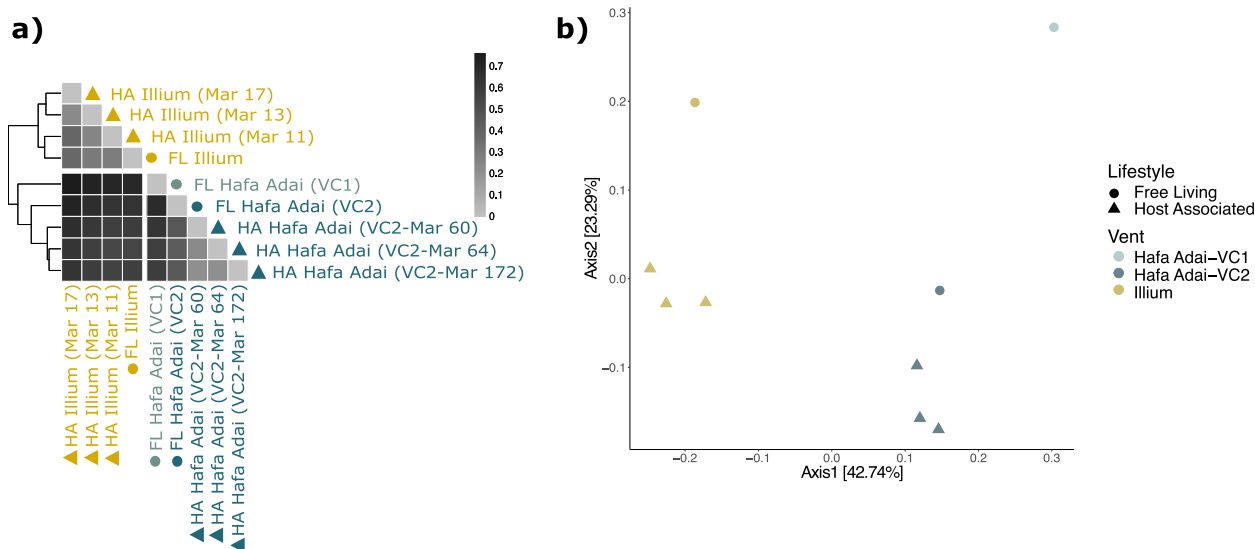


Fig. 3 Symbiont population structure based on single nucleotide polymorphisms. **A** Heatmap of genome-wide, pairwise fixation indices (F_{ST}) created using the pheatmap package in RStudio. F_{ST} values range from 0 to 1, where a value of 0 indicates no genetic differentiation, while a value of 1 indicates complete isolation among populations. “FL” and “HA” indicate free-living and host-associated symbionts, respectively. **B** PCoA plot based on Bray–Curtis distances illustrating the population structure of *A. hessleri* symbionts of different lifestyles and vent fields

Gene content differed more substantially by geography than by lifestyle: the Illium symbionts had 44 unique gene clusters, and the Hafa Adai (VC1 and VC2) symbionts had 26 (Fig. 5a, Supplementary Table 11), while only three total gene clusters were unique by lifestyle (group_681 for host-associated; group_2104 and group_2131 for free-living). However, these genes could not be characterized by any database we used for functional annotations. For all unique gene clusters across both biogeography and lifestyle, hypothetical and unknown proteins based on Prokka and the NR database were also assessed via KEGG and AlphaFold, but yielded low-confidence results. Of the successfully annotated genes unique to the Illium symbionts, most were predicted to be involved in the

mobilome and DNA metabolism, followed by membrane transport, virulence, disease, defense; RNA metabolism; sulfur metabolism; cell signaling and regulation; conjugation; iron metabolism; glycolysis and gluconeogenesis; and detoxification and stress response. Genes unique to the Hafa Adai (VC1 & VC2) symbionts were predominantly associated with the mobilome, followed by membrane transport; RNA metabolism; motility and chemotaxis; DNA metabolism; virulence, disease and defense; and glycolysis and gluconeogenesis.

Given the small-scale geographic structuring found between VC1 and VC2 at Hafa Adai, and given that VC2 has a larger sample size to represent its subpopulation, we also compared the unique genes between Illium and

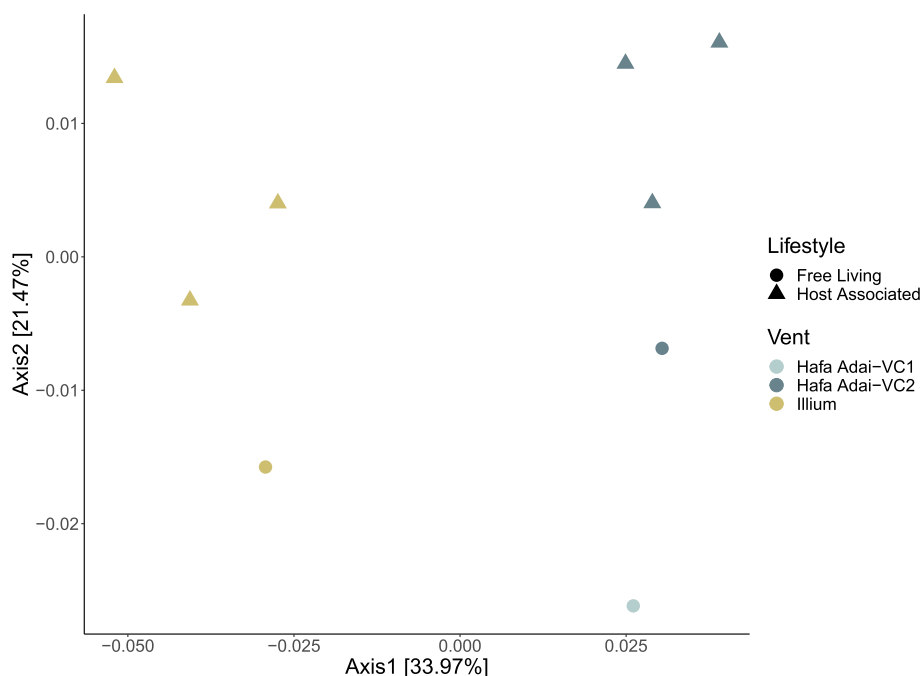


Fig. 4 PCoA plot based on Jaccard distances illustrating the difference in gene content between *A. hessleri* symbionts based on both lifestyle and vent field

Hafa Adai VC2 symbionts alone (i.e., without VC1) (Supplementary Table 12, Fig. 5b). In this case, there were 62 unique gene clusters for symbionts from Illium and 28 unique gene clusters for symbionts from Hafa Adai-VC2 (Fig. 5b), i.e., two additional as compared to VC-1 and VC-2 combined. Only one of the genes unique to the Hafa Adai symbionts could be annotated and fell under the larger subcategory of “Virulence, Disease and Defense,” whereas unique genes of the Illium symbionts spanned a variety of metabolic functions.

Analyses that included symbiont reads from Alice Springs and Burke (Supplementary Figs. 2, 3, 5; Supplementary Tables 13, 14) further supported the effect of geography over lifestyle on gene content variation in the *A. hessleri* symbionts. The population at Burke harbored a single unique, uncharacterized gene (Supplementary Table 13). When pooled with Illium as a “northern site,” additional genes unique to DNA metabolism and membrane transport were found, followed by genes involved in the mobilome, RNA metabolism, virulence, glycolysis and gluconeogenesis, cell signaling, conjugation, and stress response (Supplementary Fig. 3; Supplementary Table 13).

Alice Springs harbored three uncharacterized or hypothetical genes. When all three northern sites (Alice Springs, Illium, and Burke) were pooled together, seven unique genes were found. Four of these were related to DNA metabolism, virulence, conjugation, and

transposition (Supplementary Table 14). Since Alice Springs and Illium are more geochemically similar to one another than either vent is to Burke [25], we also investigated the unique genes shared by these two vent fields alone: four unique genes were found, one of which fell under the functional category of virulence, disease, and defense.

Discussion

Here, we compared free-living and host-associated symbiont populations of *Alviniconcha hessleri* from two vent fields in the Mariana Back-Arc. Based on ANI and taxonomic assignments, our nine representative, medium- to high-quality MAGs can be considered to represent a single species within the genus *Thiolapillus* [58]. Our results provide strong evidence that diffuse fluid flow microbial communities include populations of free-living symbionts, further supporting an expected model of horizontal transmission in *Alviniconcha* species [18, 59].

Both population structure and gene content analyses suggest that *A. hessleri* symbionts form subpopulations that segregate by geography more strongly than by lifestyle. These patterns agree with previous studies of non-symbiotic hydrothermal vent microbial communities, which show that microbes are shaped by their local environment [60], as well as of host-associated *A. hessleri* symbiont biogeography at the 16S rRNA gene level [27] and other horizontally transmitted associations from

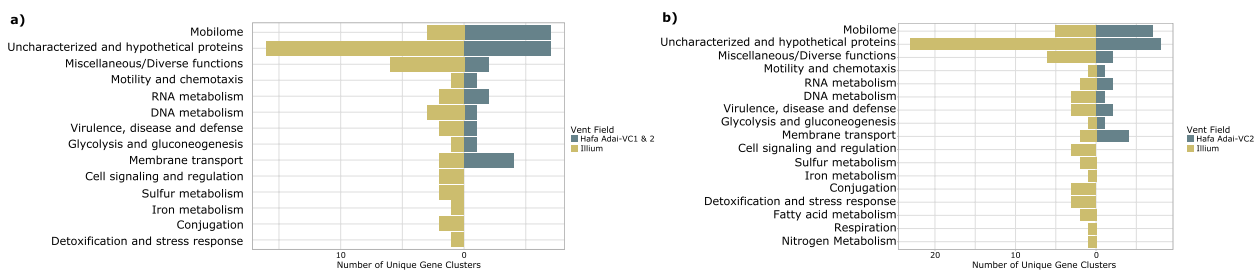


Fig. 5 Likert plots showing **A** the number of unique genes between Ilium and Hafa Adai (including both VC1 and VC2), and **B** the number of unique genes between Ilium and Hafa Adai (VC2 only)

hydrothermal vents, such as bathymodiolin mussels [49, 61, 62] and provannid snails [18], that have been shown to partner with habitat-specific symbiont strains. These results, therefore, provide further evidence for horizontal transmission in the *A. hessleri* symbiont system. Such uptake of environmental symbiont strains bears a risk of infection to the host by cheaters [16], but also enhances an animal's ability to flexibly associate with locally available symbiont strains and, therefore, to maximize the habitat range in which they can settle [2, 18]. Furthermore, since hydrothermal vents are ephemeral and geochemically dynamic habitats that harbor microbial communities shaped by local environmental conditions [60], it may be ecologically and evolutionarily advantageous for vent animals to acquire symbiont strains that are likely locally adapted [63].

The dynamics of microbial interaction with the host during acquisition and release processes can have significant impacts on the population structure and composition of horizontally transmitted symbionts. It is not known whether *A. hessleri* can replenish or recycle its symbionts, or if symbiont acquisition occurs only once upon settlement. For example, hydrothermal vent tube-worms seed the environment with their symbionts only upon death [5], *Bathymodiolus* mussels can acquire and release their symbionts throughout their lifetime [11, 64], and *Vibrio fischeri* symbionts are expelled every morning by their sepiolid squid host [65]. In *V. fischeri*, it is well established that evolution in the free-living stage—for example, via horizontal gene transfer—impacts the evolution of host-microbe interactions, though the role of novel mutations remains unclear [65]. Although *A. hessleri* symbionts were overall more strongly partitioned by geography than by lifestyle, all symbiont samples were genetically distinct from each other and formed separate free-living or host-associated subpopulations. These findings suggest that symbiont exchanges between host and environment throughout the lifetime of the host are limited but might occur occasionally via symbiont uptake or release [49], thereby leading to mixing of host-associated

and free-living symbiont pools. Periodic switching of symbiont strains could increase shared genetic variation among intra- and extra-host-symbiont populations, while maintaining geographic differentiation in the presence of dispersal barriers and/or environmental selection. All samples from Ilium showed a comparably small degree of differentiation from each other, while samples from Hafa Adai were notably divergent between free-living and host-associated lifestyles. These patterns could arise from differences in the sampling locations of the free-living symbiont populations (e.g., distance from the snail beds) and/or the age of the *Alviniconcha* host individuals. Although we do not have size-related data for the collected specimens, it is possible that the snail individuals from Hafa Adai were older than those from Ilium, giving host-associated symbiont populations more time to diverge from their free-living counterparts. Strong genetic differentiation between host-associated and free-living symbiont populations can be expected if hosts take up similar symbiont strains that have limited exchange with the environment post-infection, while the free-living symbiont population experiences more turnover.

The high genetic isolation of symbiont populations observed between vent fields may reflect the influence of both neutral (e.g., dispersal barriers and isolation-by-distance) and selective processes (e.g., adaptation to habitat differences between vent fields) on symbiont biogeography. Ilium, Burke, and Alice Springs are all northern vent fields within the Mariana Back-Arc Basin that are characterized by sites of low-temperature diffuse fluid flow, while Hafa Adai is located further south and contains high-temperature black smokers [25]. Ilium and Alice Springs are similar geochemically, notably in that they are both low in H₂S concentrations, while Burke and Hafa Adai exhibit elevated H₂S concentrations [25]. The close proximity (~360 m) and overlap in geochemical characteristics between Alice Springs and Ilium may explain why these vent fields clustered together in our population structure analyses. By contrast, Burke's distinct geochemical signature might contribute to the high

genetic isolation seen for this vent field, despite its relative proximity to Alice Springs and Illium (~4 km). Overall, however, no clear pattern of isolation-by-distance was observed, indicating that ecological factors might play a more important role than dispersal barriers in shaping symbiont population structure, in agreement with the oceanographic connectivity between the northern and central Mariana Back-Arc Basin [66].

Interestingly, Hafa Adai-VC1—while more similar to Hafa Adai-VC2 than any other vent site—represented its own symbiont subpopulation, suggesting small-scale population structuring of symbionts within vent fields. Local patchiness of symbionts, as observed in our study, mirrors patterns found for host-associated symbionts of cold-seep vestimentiferan tubeworms [67] and *Acropora* corals [68]. Although Hafa Adai-VC1 and -VC2 were only ~5 m apart, it is possible that *Alviniconcha hessleri* symbionts have extremely low dispersal potential that could be further reduced by small-scale circulation within vent sites due to physical structuring in the seafloor [69, 70]. Alternatively, micro-niche adaptation driven by locally fluctuating environmental conditions might contribute to these patterns.

Among the identified differences in gene content, symbionts from Illium uniquely harbored genes related to iron and sulfur metabolism. As iron and sulfur concentrations appear to be reduced at northern Mariana Back-Arc vents [25, 71] and are typically lower in diffuse flow than black smoker fluids such as those found at Hafa Adai, it is possible that symbionts at Illium harbor high affinity sulfur and Fe^{2+} transporters to efficiently obtain this essential element for their metabolism. All symbiont populations, including the low-coverage samples from Alice Springs and Burke, showed differences in the presence of genes related to the mobilome and virulence, disease, and defense. This suggests that each vent field supports distinct viral communities that may uniquely infect and interact with the symbionts, as hydrothermal vent viruses have restricted bacterial and archaeal host ranges, and viral communities are typically endemic to a given vent site due to limited dispersal or environmental selection [72, 73]. The high number of unique genes related to the mobilome may be a consequence of integrated phage-derived genetic material that reflect the local, free-living viral communities.

Conclusions

Our research demonstrates that *Alviniconcha hessleri* symbiont populations are primarily structured by geography rather than by their host-associated or free-living lifestyle. Future work using population genomic approaches should help clarify the predominant force(s)

shaping the geographic population structure, as recent analyses of the symbionts associated with other *Alviniconcha* species suggest that both genetic drift and local adaptation play a role in symbiont biogeography [18]. Although our analyses indicate a weak effect of lifestyle on symbiont genetic structure, it is possible that free-living and host-associated populations are characterized by differences in gene expression. A comparison of gene expression between lifestyles may provide additional clarity on the extent to which these symbiont subpopulations differ functionally. Our work also strengthens previous evidence for horizontal symbiont transmission in *Alviniconcha* species [18, 59], despite the fact that almost nothing is currently known about the dynamics of symbiont acquisition and release in these species. Given that *A. hessleri* has been classified as “Vulnerable” on the IUCN Red List (<https://www.iucnredlist.org>) and is a dominant species at vents that are part of the Marianas Trench Marine National Monument, it is critical for future conservation and management that we understand the genetic connectivity of the symbiotic microbes that support this foundation species.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s40168-023-01493-2>.

Additional file 1.

Additional file 2: Supplementary Figure 1. Population structure analyses based on 1271 single nucleotide polymorphisms. Includes all nine samples from Illium and Hafa Adai plus symbiont reads from Burke. A) PCoA plot based on Bray-Curtis distances exhibiting population genetic structuring by geography and lifestyle. B) Heatmap of genome-wide, pairwise fixation indices (F_{ST}) created using the pheatmap package in RStudio. F_{ST} values range from 0 to 1, where a value of 0 indicates no genetic differentiation, while a value of 1 indicates complete isolation among populations. “FL” and “HA” indicate free-living and host-associated symbionts, respectively.

Additional file 3: Supplementary Figure 2. PCoA plot based on Jaccard distances illustrating the difference in gene content between *A. hessleri* symbionts based on both lifestyle and vent field, including all nine samples from Illium and Hafa Adai in addition to the free-living symbiont population from the Burke vent field.

Additional file 4: Supplementary Figure 3. Barplot showing the number of unique gene clusters per functional category between vent fields, including a “northern” site category which combines the Illium and Burke vent fields. Created using ggplot in RStudio.

Additional file 5: Supplementary Figure 4. Population structure analyses based on 793 single nucleotide polymorphisms. Includes all nine samples from Illium and Hafa Adai, plus symbiont reads from both the Burke and Alice Springs vent fields. A) PCoA plot based on Bray-Curtis distances exhibiting population genetic structuring by geography and lifestyle. B) Heatmap of genome-wide, pairwise fixation indices (F_{ST}) created using the pheatmap package in RStudio. F_{ST} values range from 0 to 1, where a value of 0 indicates no genetic differentiation, while a value of 1 indicates complete isolation among populations. “FL” and “HA” indicate free-living and host-associated symbionts, respectively.

Additional file 6: Supplementary Figure 5. PCoA plot based on Jaccard distances illustrating the difference in gene content between *A. hessleri*

symbionts based on both lifestyle and vent field, including all nine samples from Illium and Hafa Adai in addition to free-living symbiont samples from the Burke and Alice Springs vent fields.

Acknowledgements

We thank the Schmidt Ocean Institute, the captain, crew and pilots of the R/V *Falkor* and ROV *SuBastian*, as well as Bill Chadwick, David Butterfield, Verena Tunnicliffe, and Amanda Bates for their support in the sample collections. The data collected in this study includes work supported by the Schmidt Ocean Institute during the R/V *Falkor* cruise FK161129.

Authors' contributions

M.H., C.B., R.A.B., and J.H. designed the study. J.H. collected the samples. E.T.R. and R.A.B. did laboratory work for the free-living and host-associated metagenomic samples, respectively. M.H., C.B., and E.T.R. performed bioinformatic analyses. M.H. drafted the manuscript. All authors edited, reviewed, and approved the text.

Funding

This work was funded by the National Science Foundation (grant number OCE-1736932 to RAB and Graduate Research Fellowship award number 1747454 to MAH). JAH was funded by the NOAA Ocean Exploration and Research (OER) Program and the NSF Center for Dark Energy Biosphere Investigations (C-DEBI) (OCE-0939564). ETR was supported by the NASA Postdoctoral Fellowship with the NASA Astrobiology Institute and the L'Oréal USA For Women in Science Fellowship. This is C-DEBI Contribution no. 614.

Availability of data and materials

The datasets supporting the conclusions of this article are available in the National Center for Biotechnology Information repository, under BioProject number PRJNA763533. The previously published, free-living raw sequencing reads and corresponding MAGs are available at the National Center for Biotechnology Information under BioProject PRJNA454888.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 20 August 2022 Accepted: 11 February 2023

Published online: 16 May 2023

References

- McFall-Ngai M, Hadfield MG, Bosch TCG, Carey HV, Domazet-Lošo T, Douglas AE, et al. Animals in a bacterial world, a new imperative for the life sciences. *Proc Natl Acad Sci U S A*. 2013;110(9):3229–36.
- Vrijenhoek RC. Genetics and evolution of deep-sea chemosynthetic bacteria and their invertebrate hosts. In: Kiel S, editor. *The Vent and Seep Biota: Aspects from Microbes to Ecosystems*. Dordrecht, Netherlands: Springer; 2010. p. 15–49.
- Russell SL. Transmission mode is associated with environment type and taxa across bacteria-eukaryote symbioses: a systematic review and meta-analysis. *FEMS Microbiol Lett*. 2019;366(3):1–13.
- Harmer TL, Rotjan RD, Nussbaumer AD, Bright M, Ng AW, DeChaine EG, et al. Free-living tubeworm endosymbionts found at deep-sea vents. *Appl Environ Microbiol*. 2008;74(12):3895–8.
- Klose J, Polz MF, Wagner M, Schimak MP, Gollner S, Bright M. Endosymbionts escape dead hydrothermal vent tubeworms to enrich the free-living population. *Proc Natl Acad Sci*. 2015;112(36):11300–5.
- Doi JA, McFall-Ngai MJ. A transient exposure to symbiosis-competent bacteria induces light organ morphogenesis in the host squid. *Biol Bull*. 1995;189(3):347–55.
- Nyholm SV, McFall-Ngai MJ. Sampling the light-organ microenvironment of *Euprymna scolopes*: description of a population of host cells in association with the bacterial symbiont *Vibrio fischeri*. *Biol Bull*. 1998;195(2):89–97.
- Martiny JBH, Bohannan BJM, Brown JH, Colwell RK, Fuhrman JA, Green JL, et al. Microbial biogeography: Putting microorganisms on the map. *Nat Rev Microbiol*. 2006;4(2):102–12.
- Dickey JR, Sweney RA, Turner SC, Winfrey CC, Yaffar D, Padukone A, et al. The utility of macroecological rules for microbial biogeography. *Front Ecol Evol*. 2021;9:633155.
- Nussbaumer AD, Fisher CR, Bright M. Horizontal endosymbiont transmission in hydrothermal vent tubeworms. *Nature*. 2006;441(7091):345–8.
- Wentrup C, Wendeberg A, Schimak M, Borowski C, Dubilier N. Forever competent: Deep-sea bivalves are colonized by their chemosynthetic symbionts throughout their lifetime. *Environ Microbiol*. 2014;16(12):3699–713.
- Dubilier N, Bergin C, Lott C. Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. *Nat Rev Microbiol*. 2008;6:725–40.
- Olins HC, Rogers DR, Frank KL, Vidoudez C, Girguis PR. Assessing the influence of physical, geochemical and biological factors on anaerobic microbial primary productivity within hydrothermal vent chimneys. *Geobiology*. 2013;11(3):279–93.
- McNichol J, Stryhanyuk H, Sylva SP, Thomas F, Musat N, Seewald JS, et al. Primary productivity below the seafloor at deep-sea hot springs. *Proc Natl Acad Sci U S A*. 2018;115(26):6756–61.
- Beinart RA, Gartman A, Sanders JG, Luther GW, Girguis PR. The uptake and excretion of partially oxidized sulfur expands the repertoire of energy resources metabolized by hydrothermal vent symbioses. *Proc R Soc B Biol Sci*. 2015;282:20142811.
- Douglas AE. Conflict, cheats and the persistence of symbioses. *New Phytol*. 2008;177(4):849–58.
- Lan Y, Sun J, Chen C, Sun Y, Zhou Y, Yang Y, et al. Hologenome analysis reveals dual symbiosis in the deep-sea hydrothermal vent snail *Gigantopelta aegis*. *Nat Commun*. 2021;12(1):1–15.
- Breusing C, Genetti M, Russell SL, Corbett-Detig RB, Beinart RA. Horizontal transmission enables flexible associations with locally adapted symbiont strains in deep-sea hydrothermal vent symbioses. *Proc Natl Acad Sci USA*. 2022;119(14):1–11.
- Beinart RA, Sanders JG, Faure B, Sylva SP, Lee RW, Becker EL, et al. Evidence for the role of endosymbionts in regional-scale habitat partitioning by hydrothermal vent symbioses. *Proc Natl Acad Sci U S A*. 2012;109(47):1–10.
- Beinart RA, Luo C, Constantinidis KT, Stewart FJ, Girguis PR. The bacterial symbionts of closely related hydrothermal vent snails with distinct geochemical habitats show broad similarity in chemoautotrophic gene content. *Front Microbiol*. 2019;10:1818.
- Breusing C, Johnson SB, Tunnicliffe V, Vrijenhoek RC. Population structure and connectivity in Indo-Pacific deep-sea mussels of the *Bathymodiolus septemdiarium* complex. *Conserv Genet*. 2015;16(6):1415–30.
- Hinzke T, Kleiner M, Meister M, Schlüter R, Hentschker C, Pané-Farré J, et al. Bacterial symbiont subpopulations have different roles in a deep-sea symbiosis. *Elife*. 2021;10:1–45.
- Ücker M. Metagenomic analyses of a deep-sea mussel symbiosis. Dissertation, University of Bremen. 2021. <https://doi.org/10.26092/elib/672>.
- Polzin J, Arevalo P, Nussbaumer T, Polz MF, Bright M. Polyclonal symbiont populations in hydrothermal vent tubeworms and the environment. *Proc R Soc B Biol Sci*. 2019;286:20181281.
- Trembath-Reichert E, Butterfield DA, Huber JA. Active subseafloor microbial communities from Mariana back-arc venting fluids share metabolic strategies across different thermal niches and taxa. *ISME J*. 2019;13(9):2264–79.
- Endow K, Ohta S. The symbiotic relationship between bacteria and a mesogastropod snail, *Alviniconcha hessleri*, collected from hydrothermal vents of the Mariana Back-Arc Basin. *Bull Japanese Soc Microb Ecol*. 1989;3(2):73–82.
- Breusing C, Johnson SB, Tunnicliffe V, Clague DA, Vrijenhoek RC, Beinart RA. Allopatric and sympatric drivers of speciation in *Alviniconcha* hydrothermal vent snails. *Mol Biol Evol*. 2020;37(12):3469–84.

28. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114–20.
29. Li D, Liu CM, Luo R, Sadakane K, Lam TW. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics*. 2015;31(10):1674–6.
30. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19(5):455–77.
31. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. MetaSPAdes: A new versatile metagenomic assembler. *Genome Res*. 2017;27(5):824–34.
32. Seah BKB, Gruber-Vodicka HR. gbtools: Interactive visualization of metagenome bins in R. *Front Microbiol*. 2015;6:1451.
33. Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics*. 2011;27(4):578–9.
34. Luo J, Lyu M, Chen R, Zhang X, Luo H, Yan C. SLR: a scaffolding algorithm based on long reads and contig classification. *BMC Bioinformatics*. 2019;20(1):1–11.
35. Boetzer M, Pirovano W. Toward almost closed genomes with GapFiller. *Genome Biol*. 2012;13(6):R56.
36. Xu GC, Xu TJ, Zhu R, Zhang Y, Li SQ, Wang HW, et al. LR-GapCloser: A tiling path-based gap closer that uses long reads to complete genome assembly. *Gigascience*. 2018;8(1):1–14.
37. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS ONE*. 2014;9(11):e112963.
38. Wu Y-W, Tang Y-H, Tringie SG, Simmons BA, Singer SW. MaxBin: an automated binning method to recover individual genomes from metagenomes using. *Microbiome*. 2014;2(26):4904–9.
39. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*. 2013;29(8):1072–5.
40. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res*. 2015;25(7):1043–55.
41. Chaumeil PA, Mussig AJ, Hugenholtz P, Parks DH. GTDB-Tk: A toolkit to classify genomes with the genome taxonomy database. *Bioinformatics*. 2020;36(6):1925–7.
42. Jain C, Rodriguez-R LM, Phillippy AM, Konstantinidis KT, Aluru S. High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat Commun*. 2018;9(1):5114.
43. Minh BQ, Schmidt HA, Chernomor O, Schrempf D, Woodhams MD, Von Haeseler A, et al. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol*. 2020;37(5):1530–4.
44. Eren AM, Esen OC, Quince C, Vineis JH, Morrison HG, Sogin ML, et al. Anvi'o: An advanced analysis and visualization platform for 'omics data. *PeerJ*. 2015;2015(10):1–29.
45. Tonkin-Hill G, MacAlasdair N, Ruis C, Weimann A, Horesh G, Lees JA, et al. Producing polished prokaryotic pangenomes with the Panaroo pipeline. *Genome Biol*. 2020;21(1):1–21.
46. Villanueva RAM, Chen ZJ. ggplot2: Elegant Graphics for Data Analysis (2nd ed.). *Meas Interdiscip Res Perspect*. 2019;17(3):160–7.
47. Oksanen AJ, Blanchet FG, Kindt R, Legendre P, Minchin PR, Hara RBO, et al. *vegan*: Community Ecology Package. 2018. <https://CRAN.R-project.org/package=vegan>.
48. RStudio Team. RStudio: Integrated Development Environment for R. Boston, MA: RStudio, PBC; 2021.
49. Romero Picazo D, Dagan T, Ansorge R, Petersen JM, Dubilier N, Kupczok A. Horizontally transmitted symbiont populations in deep-sea mussels are genetically isolated. *ISME J*. 2019;13(12):2954–68.
50. Kolde R. pheatmap: Pretty Heatmaps. 2019. <https://CRAN.R-project.org/package=pheatmap>.
51. Beghini F, McIver LJ, Blanco-Míguez A, Dubois L, Asnicar F, Maharjan S, et al. Integrating taxonomic, functional, and strain-level profiling of diverse microbial communities with biobakery 3. *Elife*. 2021;10:1–42.
52. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30(14):2068–9.
53. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res*. 2016;44(D1):D457–462.
54. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. *Nature*. 2021;596(7873):583–9.
55. Heiberger RM, Holland B. *Statistical Analysis and Data Display: An Intermediate Course with Examples in R*. 2nd ed. New York: Springer; 2015.
56. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*. 2013;41(D1):590–6.
57. Ratnasingham S, Hebert PDN. BOLD: The Barcode of Life Data System. *Mol Ecol Notes*. 2007;7(3):355–64.
58. Nunoura T, Takaki Y, Kazama H, Kakuta J, Shimamura S, Makita H. Physiological and genomic features of a novel sulfur-oxidizing Gammaproteobacterium belonging to a previously uncultivated symbiotic lineage isolated from a hydrothermal vent. *PLoS ONE*. 2014;9(8):e104959.
59. Suzuki Y, Kojima S, Sasaki T, Suzuki M, Utsumi T, Watanabe H, et al. Host-symbiont relationships in hydrothermal vent gastropods of the genus *Alviniconcha* from the Southwest Pacific. *Appl Environ Microbiol*. 2006;72(2):1388–93.
60. Dick GJ. The microbiomes of deep-sea hydrothermal vents: distributed globally, shaped locally. *Nat Rev Microbiol*. 2019;17(5):271–83.
61. Ücker M, Ansorge R, Sato Y, Sayavedra L, Breusing C, Dubilier N. Deep-sea mussels from a hybrid zone on the Mid-Atlantic Ridge host genetically indistinguishable symbionts. *ISME J*. 2021;15(10):3076–83.
62. Brzechffa C, Goffredi SK. Contrasting influences on bacterial symbiont specificity by co-occurring deep-sea mussels and tubeworms. *Environ Microbiol Rep*. 2021;13(2):104–11.
63. Vrijenhoek RC. Genetic diversity and connectivity of deep-sea hydrothermal vent metapopulations. *Mol Ecol*. 2010;19:4391–411.
64. Kádár E, Bettencourt R, Costa V, Santos RS, Lobo-da-Cunha A, Dando P. Experimentally induced endosymbiont loss and re-acquirement in the hydrothermal vent bivalve *Bathymodiolus azoricus*. *J Exp Mar Bio Ecol*. 2005;318(1):99–110.
65. Soto W, Travisano M, Tolleson AR, Nishiguchi MK. Symbiont evolution during the free-living phase can improve host colonization. *Microbiology (United Kingdom)*. 2019;165(2):174–87.
66. Breusing C, Johnson SB, Mitarai S, Beinart RA, Tunnicliffe V. Differential patterns of connectivity in Western Pacific hydrothermal vent metapopulations: a comparison of biophysical and genetic models. *Evol Appl*. 2021. <https://doi.org/10.1111/eva.13326>.
67. Vrijenhoek RC, Duhaime M, Jones WJ. Subtype variation among bacterial endosymbionts of tubeworms (Annelida: Siboglinidae) from the Gulf of California. *Biol Bull*. 2007;212(3):180–4.
68. Davies SW, Moreland KN, Wham DC, Kanke MR, Matz MV. *Cladocypium* community divergence in two *Acropora* coral hosts across multiple spatial scales. *Mol Ecol*. 2020;29(23):4559–72.
69. Stewart LC, Algar CK, Fortunato CS, Larson BI, Vallino JJ, Huber JA, et al. Fluid geochemistry, local hydrology, and metabolic activity define methanogen community size and composition in deep-sea hydrothermal vents. *ISME J*. 2019;13(7):1711–21.
70. Fortunato CS, Larson B, Butterfield DA, Huber JA. Spatially distinct, temporally stable microbial populations mediate biogeochemical cycling at and below the seafloor in hydrothermal vent fluids. *Environ Microbiol*. 2018;20(2):769–84.
71. Ishibashi J, Ichiro, Tsunogai U, Toki T, Ebina N, Gamo T, Sano Y, et al. Chemical composition of hydrothermal fluids in the central and southern Mariana Trough backarc basin. *Deep Res Part II Top Stud Oceanogr*. 2015;121:126–136.
72. Paez-Espino D, Eloë-Fadrosch EA, Pavlopoulos GA, Thomas AD, Hunte-mann M, Mikhailova N, et al. Uncovering Earth's virome. *Nature*. 2016;536(7617):425–30.
73. Thomas E, Anderson RE, Li V, Rogan LJ, Huber JA. Diverse viruses in deep-sea hydrothermal vent fluids have restricted dispersal across ocean basins. *mSystems*. 2021;6(3):1–18.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.