Global 16S rRNA diversity of provannid snail endosymbionts from Indo-Pacific deep-sea hydrothermal vents

Corinna Breusing, Jade Castel, Yi Yang, Thomas Broquet, Jin Sun, Didier Jollivet, Pei-yuan Qian, Roxanne A Beinart

Global 16S rRNA diversity of provannid snail endosymbionts from Indo-Pacific deep-sea hydrothermal vents

Corinna Breusing[1], Jade Castel[2], Yi Yang[3], Thomas Broquet[2], Jin Sun[4], Didier Jollivet[2], Pei-Yuan Qian[3] and Roxanne A. Beinart[1*]

[1]Graduate School of Oceanography, University of Rhode Island, Narragansett, RI, USA
[2]CNRS UMR 7144 'Adaptation et Diversité en Milieux Marins' (AD2M), Team 'Dynamique de la Diversité Marine' (DyDiv), Station Biologique de Roscoff, Roscoff, France
[3]Department of Ocean Science, Division of Life Science and Hong Kong Branch of the Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), The Hong Kong University of Science and Technology, Hong Kong, China
[4]Institute of Evolution & Marine Biodiversity, Ocean University of China, Qingdao, China

*Corresponding author: rbeinart@uri.edu

Running title: Endosymbiont diversity in hydrothermal vent snails

Keywords: Chemosynthetic symbiosis, hydrothermal vents, 16S rRNA amplicon sequencing, symbiont diversity, *Alviniconcha*, *Ifremeria*

Originality-Significance Statement

Microbial symbionts are increasingly recognized as ubiquitous phenomena that are important components of host biology. Yet, our knowledge of how symbionts vary across geographic scales, habitats and host species remains limited, especially for symbioses in remote environments that are challenging to sample comprehensively. We assembled a global dataset of chemosynthetic symbionts associated with provannid gastropods from Indo-Pacific deep-sea hydrothermal vents and evaluated their diversity and biogeographic structure through 16S rRNA amplicon sequencing. With unprecedented sample size and geographic coverage included in our analyses, we found that symbiont composition within a host species is shaped by broad-scale geography, while other factors such as host size seem to be of limited importance. Furthermore, the richness of symbionts associated with a host species was not always related to sample size or biogeographic range, which indicates that there are likely additional factors shaping symbiont composition and diversity. Altogether, this work contributes to our understanding of the patterns and processes underlying symbiont biogeography in the marine environment.

Summary

Symbioses between invertebrate animals and chemosynthetic bacteria build the foundation of deep-sea hydrothermal ecosystems worldwide. Despite the importance of these symbioses for ecosystem functioning, the diversity of symbionts within and between host organisms and geographic regions is still poorly understood. In this study we used 16S rRNA amplicon sequencing to determine the diversity of gill endosymbionts in provannid snails of the genera *Alviniconcha* and *Ifremeria*, which are key species at deep-sea hydrothermal vents in the Indo-Pacific Ocean. Our analysis of 761 snail samples across the distributional range of these species confirms previous findings that symbiont lineages are strongly partitioned by host species and broad-scale geography. Less structuring was observed within geographic regions, probably due to insufficient strain-resolution of the 16S rRNA gene. Symbiont richness in individual hosts appeared to be unrelated to host size, suggesting that provannid snails might acquire their symbionts only during a permissive time window in early developmental stages in contrast to other vent mollusks that obtain their symbionts throughout their lifetime. Despite the extent of our dataset, symbiont accumulation curves did not reach saturation, highlighting the need for increased

sampling efforts to uncover the full diversity of symbionts within these and other hydrothermal vent species.

Introduction

Microbial symbioses are increasingly recognized as universal phenomena that impact virtually all levels of biological organization, from cellular to organismal to ecosystem scale (Bronstein, 2015). Growing evidence from various symbiotic partnerships suggests that microbial symbioses can expand the physiological and ecological capabilities of hosts and symbionts, which are predicted to be critical for ecosystem productivity, stability and biogeochemical cycling (Apprill, 2017; Beinart, 2019; Wilkins *et al.*, 2019). Deep-sea hydrothermal vents are probably some of the most enigmatic ecosystems that are sustained by microbial symbioses. In these systems, invertebrate animals live in association with chemoautotrophic bacteria that use chemical energy from venting fluids for the production of organic carbon, thereby providing food for their hosts (Dubilier *et al*., 2008; Sogin *et al.*, 2021). Despite decades of research on this topic and the significance of chemosynthetic symbioses for ecosystem processes at hydrothermal vents, the diversity and distribution of symbionts within and across hosts and habitats remains underexplored, especially at large biogeographic scales.

Provannid snails of the sister genera *Alviniconcha* and *Ifremeria* are dominant animals in benthic communities at deep-sea hydrothermal vents in the Indian and Western Pacific Ocean (Van Dover *et al*., 2001; Desbruyères *et al*., 2006). While the Western Pacific genus *Ifremeria* is represented by a single species, *I. nautilei*, that affiliates with methane- and/or sulfide-oxidizing gammaproteobacterial symbionts (Windoffer and Giere, 1997; Borowski *et al*., 2002; Suzuki *et al*., 2006a), the genus *Alviniconcha* comprises five Western Pacific species (*A. adamantis*, *A. boucheti*, *A. hessleri*, *A. kojimai*, *A. strummeri*) and one Indian Ocean species (*A. marisindica*) that live in symbiosis with thiotrophic Gammaproteobacteria or Campylobacteria (Suzuki *et al*., 2006b; Johnson *et al*., 2015; Breusing *et al*., 2020). In both *Alviniconcha* and *Ifremeria*, the symbionts are assumed to be horizontally acquired and are harbored intracellularly within the host's gill tissue (Suzuki *et al.*, 2006a, b). Despite an environmental pathway for symbiont transmission, host and symbiont genera or species appear to exhibit a relatively strong selectivity in their partnerships towards each other (Beinart *et al.*, 2012; Breusing *et al*., 2020), though host individuals are flexible in recruiting local strains of their specific symbiont phylotype(s) (Breusing *et al*., 2021).

Most current analyses on the variation and structure of microbial symbionts within *Alviniconcha* and *Ifremeria* stem from studies in the Lau Back Arc Basin, while little is known about these patterns in populations from other spreading systems within the distributional range of these genera. Here, we compiled an extensive dataset of 761 snail samples from 10 geographic regions of the Indo-Pacific Ocean (Fig. 1), some of which were previously unexplored, to assess the global diversity of chemosynthetic gill endosymbionts within *Alviniconcha* and *Ifremeria* through identification of 16S rRNA amplicon sequence variants (ASVs). Using ordination analyses and correlative statistics, we determined the influence of host species, host size, depth and geography on symbiont composition and distribution.

Results and Discussion

Symbiont 16S rRNA diversity is partitioned by host species and geography

Our conservative analysis pipeline, which extends a previous study by Breusing *et al.* (2020) to now include seven species and 10 geographic areas, recovered 60 symbiont ASVs that were assigned to two campylobacterial (*Sulfurovum*, *Sulfurimonas*) and four gammaproteobacterial (*Ca*. Thiobios, *Methylomonas*, *Thiolapillus*, unclassified Thiomicrospiraceae) genera of provannid snail endosymbionts (Fig. 2, 3). Average pairwise identities within genera ranged from 95% to 99% (*Sulfurovum*: 95.4%; *Sulfurimonas*: 95.0%; *Ca*. Thiobios: 97.1%; *Methylomonas*: n.a.; *Thiolapillus*: 98.1%; unclassified Thiomicrospiraceae: 99.0%). In agreement with Breusing *et al.* (2020), ASVs were generally segregated by host species and broader geographic region (i.e., back-arc basin, volcanic arc or mid-ocean ridge), except for lineages within the unclassified Thiomicrospiraceae group which were shared between *A. kojimai* and *A. strummeri* (Fig. 2, 4A; Appendix 1: Fig. S1). Based on PERMANOVAs and linear decomposition models the impact of host species and geography superseded the influence of DNA preservation, extraction and sequencing method (81.17% versus 1.99% explained variation) and was significant even when corrected for confounding technical effects. In addition, there was no evident clustering of samples by methodology in multidimensional scaling, indicating that the observed patterns are true biological signals (Table 1; Appendix 1: Fig. S2).

Like *A. kojimai* and *A. strummeri*, most other host species were associated with particular lineages of thiotrophic Gammaproteobacteria. *Alviniconcha adamantis* was affiliated with symbionts of the genus *Ca.* Thiobios, whereas *A. hessleri* and *I. nautilei* hosted distinct

Thiolapillus symbiont ASVs. Many *I. nautilei* individuals further harbored a minority methanotrophic symbiont from the genus *Methylomonas*, especially at vent sites within the Eastern Lau Spreading Center (ELSC). Only *Alviniconcha boucheti* and *A. marisindica* were dominated by different region-specific campylobacterial ASVs of the genera *Sulfurimonas* or *Sulfurovum*.

Within geographic area, the gammaproteobacterial symbionts of *A. kojimai* and *A. hessleri* showed evidence for structuring by vent field (Appendix 1: Fig. S3), while no intra-regional differentiation was observed or could be tested in symbionts of any other host species that we sampled from multiple localities (data not shown). However, this finding is likely an artifact of the limited resolution of the 16S rRNA marker gene. For example, recent metagenomic analyses indicate that symbiont populations of all host taxa from the Lau Basin are partitioned between vent sites (Breusing *et al.*, 2021). In contrast to the traditional view of microbial biogeography that poses that "everything is everywhere" (Baas-Becking, 1934), geographic subdivision of microbial symbionts appears to be common in a variety of marine symbioses, often exceeding that of the corresponding host populations (Ho *et al.*, 2017; Gould and Dunlap, 2019; Davies *et al*., 2020; Breusing *et al.*, 2021; Ücker *et al*., 2021). Depending on the symbiotic system, these patterns might arise from local adaptation, contrasting dispersal limitations between hosts and symbionts, host ecological behavior and/or differences in environmental transmission mode. Given the strong oceanographic barriers among back-arc basins in the Western Pacific Ocean (Mitarai *et al*., 2016), the observed partitioning of host-specific symbiont ASVs according to broader geographic area might be largely due to decreased symbiont dispersal opportunities (though environmental differences cannot be ruled out). By contrast, symbiont structure within regions, where dispersal limitations appear to be mostly absent (Mitarai *et al*., 2016), is probably driven by additional ecological factors, such as differences in depth or vent geochemistry (Breusing *et al*., 2021). Indeed, in *A. kojimai* the observed partitioning of symbiont types by vent field was correlated with contrasting depth regimes (Appendix 2), which often aligns with gradients in fluid chemistry (Beinart *et al*., 2012). On the other hand, the strong latitudinal subdivision found for the *Thiolapillus* symbiont of *A. hessleri* might be explained by dispersal limitations as biophysical models indicate that the southern and northern parts of the Mariana Basin are largely isolated (Mitarai *et al*., 2016; Breusing *et al*., 2021).

Our data suggest that other factors, such as host size, have a comparatively small influence on the diversity and composition of symbiont ASVs within host individuals. Despite significant

associations of symbiont richness with host size, correlation coefficients were low, suggesting limited biological relevance of this factor on intra-host symbiont diversity (Appendix 1: Fig. S4). These results were consistent independent of whether analyses were carried out across or within individual host species. For intra-species analyses only correlations for *A. kojimai* and *A. boucheti* were significant, though weak ($p \leqq 0.05$; $R^2 \leqq 0.09$). In most cases individuals contained only one symbiont ASV in accordance with Sanger sequence analyses (Beinart *et al*., 2012, 2015), though in some individuals up to six ASVs were observed. Although our study lacks data from settling larvae and juveniles, these findings could indicate that symbiont acquisition in provannid snails follows a different process than in bathymodiolin mussels and is more similar to that in vestimentiferan tubeworms. Hydrothermal vent mussels remain competent for symbiont acquisition throughout their lifetime (Wentrup *et al*., 2014; Ansorge *et al.*, 2019), which should favor increased symbiont diversity in older individuals as well as newly infected juveniles where symbiont sorting has not yet been completed. By contrast, vestimentiferan tubeworms obtain their symbionts exclusively in a narrow window after settlement during post-larva metamorphosis (Nussbaumer *et al*., 2006). Symbiont diversity can thus be expected to be highest at that developmental stage, with little effect of host size on symbiont richness during later stages. Alternatively, our observations may indicate that 16S rRNA amplicon sequences do not provide enough strain-level resolution to observe shifts in symbiont composition across development stages, and that metagenomic analyses of symbiont populations are necessary instead.

Symbiont richness differs between host species and individuals

Despite low impact of host size, *Alviniconcha* and *Ifremeria* exhibited notable variability in symbiont diversity, both among individuals and species (Fig. 4B). These patterns could result from differences in the availability and composition of free-living symbiont lineages at the time of infection, subsequent mutations inside the host and/or host selection on particular strains. Among host taxa, *A. adamantis* and *A. marisindica* showed the lowest symbiont diversity, which is probably due to the fact that these species were each sampled from only a single vent site and were represented by relatively few individuals (Fig. 4B). Interestingly, *A. hessleri* displayed some of the highest alpha diversities, with up to six ASVs within single host individuals, despite its restricted geographic distribution and small sample size compared to some of the other *Alviniconcha* species included in our analyses. Maybe the wide variation of geochemical

conditions in the Mariana Back-Arc Basin (Trembath-Reichert *et al*., 2019) allows for a greater range of micro-niches, which could promote diversity in the free-living symbiont pool. In this case, symbionts within this host species might have a higher functional diversity that could favor co-existence of multiple strains, as has recently been reported for bathymodiolin mussels, where hosts can carry up to 16 symbiont strains due to variation in metabolic gene content (Ansorge *et al*., 2019). Alternatively, some of the observed variation might reflect intra-host mutations of a single or a few symbiont phylotypes post-infection. In the absence of genomic data, this explanation seems likely as all *A. hessleri* symbiont ASVs were very similar to each other, with an average of 99.4% pairwise sequence identity.

Symbiont richness is not saturated

Although we analyzed symbiont 16S rRNA composition in over 700 snail individuals, symbiont discovery did not reach saturation in our dataset (Fig. 5). The number of ASVs within *A. hessleri* and *I. nautilei*, which both host symbionts of the genus *Thiolapillus*, was closest to reaching a plateau, while ASV accumulation curves for all other species showed a steady increase (Fig. 5). This is an interesting finding given that *A. hessleri* and *I. nautilei* were sampled across a relatively restricted area compared to some of the other species (Appendix 1: Table S1). For other taxa that were represented by few individuals and geographic locations (e.g., *A. adamantis*, *A. marisindica*), but also those with widespread distributions (e.g., *A. kojimai*, *A. boucheti*), increased sampling efforts will probably reveal a currently hidden diversity of symbiont ASVs in the future. Consequently, while our dataset does not allow comparisons of diversification between symbiont genera or species at this time, more ASVs especially for some of the gammaproteobacterial taxa (e.g., unclassified Thiomicrospiraceae, *Ca*. Thiobios) will likely be recovered given the prevalence of gammaproteobacterial symbioses in provannid snails and other vent invertebrates (Dubilier *et al.*, 2008).

Conclusions

Here, we characterized the global diversity of chemosynthetic gill endosymbionts associated with species within the genera *Alviniconcha* and *Ifremeria*. As predicted by previous work, we found that each host species harbored 1–2 species- or genus-level symbiont phylotypes. However, we

were able to further assess strain-level symbiont composition and diversity within and between individual snails by employing amplicon analysis of the 16S rRNA gene. In all host species, ASV accumulation curves indicated that the full diversity of symbionts associated with *Alviniconcha* and *Ifremeria* remains to be characterized. In most cases, symbiont ASV composition and richness was related to geographic range, with most ASVs detected in species where we sampled a large number of individuals across >10 geographically distant vent fields (e.g., *A. kojimai* and *A. boucheti*). An exception to this was *A. hessleri*, which had high symbiont richness and inter-region symbiont structure despite a smaller sample size and much more modest geographic range, suggesting that these are not the only factors dictating symbiont composition and diversity. A more complete appraisal of the taxonomic and functional diversity of symbionts associated with *Alviniconcha* and *Ifremeria* will be critical to our understanding of the ecology and evolution of these genera, which have been assessed as "Endangered" or "Vulnerable" on the IUCN Red List (https://www.iucnredlist.org) due to imminent risks from deep-seabed mining activities at hydrothermal vents in the Indian and Pacific oceans.

Experimental Procedures

Sample collection and amplicon library preparation

Animal samples were obtained with remotely or human operated vehicles from 23 Indo-Pacific vent localities that encompassed the global distributional range of species within the genera *Alviniconcha* and *Ifremeria* (Appendix 2; Fig. 1). Upon recovery of the samples, endosymbiont-bearing gill tissue was dissected and frozen or stored in RNALater™ (Thermo Fisher Scientific, Inc., Waltham, MA, USA) at –80°C. DNA was purified with the Zymo Quick DNA 96 Plus and ZR-96 Clean-up kits (Zymo Research, Inc., Irvine, CA, USA) or the Qiagen DNeasy Blood & Tissue kit (Qiagen, Inc., Hilden, Germany). 2x250 bp paired-end amplicon libraries for the 16S rRNA V4-V5 region were constructed with the 515F/926R primer pair (Walters et al. 2015) and sequenced to an average of 34844 total reads on Illumina MiSeq and NovaSeq platforms at the Argonne National Laboratory (Lemont, IL, USA) and Novogene Co. (Beijing, China), respectively (Appendix 2). Host species were identified through shell morphology (Laming *et al.*, 2020) and subsequent sequencing of the mitochondrial *COI* gene with universal primers (Folmer *et al*., 1994; Geller *et al*., 2013).

Identification of amplicon sequence variants

We used the USEARCH v11 denoising pipeline (Edgar, 2010) to decompose merged, adapter-clipped paired end reads into ASVs, imposing a merge length of 300–400 bp, a maximum error rate of 0.001 and a minimum base quality of 20. The taxonomic identity of each variant was determined in QIIME2 (https://qiime2.org) with a Naïve Bayes classifier trained against the V4-V5 region extracted from the SILVA 132 99% reference database as well as through BLAST+ searches against the NR database (Camacho *et al*., 2009). Only ASVs that had a match to a previously verified *Alviniconcha* or *Ifremeria* gill endosymbiont sequence were considered for further analysis. To assess potentially unrecovered variation in the symbiont dataset we applied the OLIGOTYPING v2.0 method (Eren *et al*., 2013). ASVs with less than 2.37% abundance in a sample were excluded to account for sample cross-contamination (Minich *et al*., 2019). Phylogenetic relationships among ASVs were determined with the IQTREE (Minh *et al*., 2020) plugin for QIIME2 based on 10 independent runs with each 5000 ultrafast bootstrap samples. Ultrafast bootstrap trees were optimized through the nearest neighbor interchange procedure with a perturbation strength of 0.2 and a stopping criterium of 200 trees.

16S rRNA diversity analyses

We used the PHYLOSEQ package in R v4.0.3 (McMurdie and Holmes, 2013; R Core Team, 2020) to assess symbiont 16S rRNA variation within and between hosts and geographic regions, excluding samples with less than 1000 reads to ensure statistical robustness. For alpha and beta diversity analyses symbiont abundances were normalized to proportions (McKnight *et al*., 2018). Metric and non-metric multidimensional scaling plots were constructed based on weighted UniFrac distances. To verify that the distribution of ASV diversity is representative of real biological patterns and not technical artifacts from differences in methodology, we performed linear decomposition models (LDMs) and a modified version of PERMANOVA with the LDM package in R, as these methods have been shown to be relatively robust to variance in group dispersion (Hu and Satten, 2020). Analyses were run on both the full dataset and a data subset including only samples of *Alviniconcha* from the ELSC which were processed with a mixture of methods. PERMANOVAs and LDMs were conducted with 1000 and 10000 maximum permutations, respectively, with methodology included as either confounding variable or main

explanatory factor. Relationships between number of ASVs and host size were determined based on Spearman rank correlations with the GGPUBR package (Kassambara, 2020).

Data availability

All bioinformatic scripts and final files for analysis are available on GitHub under https://github.com/cbreusing/Provannid_16S_SSU_meta-analysis. Raw 16S rRNA amplicon reads have been deposited in the Sequence Read Archive under BioProjects PRJNA473256, PRJNA473257, PRJNA610289, PRJNA610290, PRJNA763784 and PRJNA767887, while host *COI* sequences are available in GenBank under accession numbers listed in Appendix 2.

Acknowledgements

We thank the captains, crews and pilots of the R/V *Falkor* (ROV *Ropos* and *SuBastian*), R/V *L'Atalante* (ROV *Victor*) and R/V *Xiangyanghong 9* (HOV *Jiaolong*) for supporting the sample collections that made this study possible. Verena Tunnicliffe and Amanda Bates are gratefully acknowledged for providing snail specimen from the Mariana Arc and Back-Arc. We further thank the technical staff at the Argonne National Laboratory and Novogene Co. for preparing and sequencing the Illumina 16S rRNA amplicon libraries. This study was supported by the Key Special Project for Introduced Talents Team of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (GML2019ZD0409) (to P.Y.Q.), the General Research Fund of Hong Kong SAR (16101219) (to P.Y.Q.), the Schmidt Ocean Institute, a University of Rhode Island Proposal Development Grant award (to R.A.B.), the U.S. National Science Foundation (grant numbers OCE-1536331, 1819530 and 1736932 to R.A.B. and EPSCoR Cooperative Agreement OIA-#1655221), and the Agence Nationale de la Recherche Cerberus project (ANR-17-CE02-0003-01, coord. S. Hourdez).

Conflict of Interest

The authors declare no conflict of interest.

References

Ansorge, R., Romano, S., Sayavedra, L., Porras, M.Á.G., Kupczok, A., Tegetmeyer, H.E., Dubilier, N., *et al.* (2019) Functional diversity enables multiple symbiont strains to coexist in deep-sea mussels. *Nat Microbiol* **4**: 2487–2497.

Apprill, A. (2017) Marine animal microbiomes: toward understanding host–microbiome interactions in a changing ocean. *Front Mar Sci* **4**: 222.

Baas-Becking, L.G.M. (1934) Geobiologie of Inleiding Tot De Milieukunde. The Hague, The Netherlands: W.P. van Stockum & Zoon.

Beinart, R.A. (2019) The significance of microbial symbionts in ecosystem processes. *mSystems* **4**: e00127-19.

Beinart, R.A., Gartman, A., Sanders, J.G., Luther, G.W., and Girguis, P.R. (2015) The uptake and excretion of partially oxidized sulfur expands the repertoire of energy resources metabolized by hydrothermal vent symbioses. *Proc R Soc B* **282**: 20142811.

Beinart, R.A., Sanders, J.G., Faure, B., Sylva, S.P., Lee, R.W., Becker, E.L., Gartman, A., *et al.* (2012) Evidence for the role of endosymbionts in regional-scale habitat partitioning by hydrothermal vent symbioses. *Proc Natl Acad Sci USA* **109**: E3241–E3250.

Borowski, C., Giere, O., Krieger, J., Amann, R., and Dubilier, N. (2002) New aspects of the symbiosis in the provannid snail *Ifremeria nautilei* from the North Fiji Back Arc Basin. *Cah Biol Mar* **43**: 321–324.

Breusing, C., Genetti, M., Russell, S.L., Corbett-Detig, R.B., and Beinart, R.A. (2021) Host-symbiont population genomics provide insights into partner fidelity, transmission mode and habitat adaptation in deep-sea hydrothermal vent snails. *bioRxiv* https://www.biorxiv.org/content/10.1101/2021.07.13.452231v1.

Breusing, C., Johnson, S.B., Mitarai, S., Beinart, R.A., and Tunnicliffe, V. (2021) Differential patterns of connectivity in Western Pacific hydrothermal vent metapopulations: A comparison of biophysical and genetic models. *Evol Appl* DOI: 10.1111/eva.13326.

Breusing, C., Johnson, S.B., Tunnicliffe, V., Clague, D.A., Vrijenhoek, R.C., and Beinart, R.A. (2020) Allopatric and sympatric drivers of speciation in *Alviniconcha* hydrothermal vent snails. *Mol Biol Evol* **37**: 3469–3484.

Bronstein, J.L. ed. (2015) Mutualism, 1st edition. Oxford, United Kingdom: Oxford University Press.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., and Madden, T.L. (2009) BLAST+: architecture and applications. *BMC Bioinformatics* **10**: 421.

Davies, S.W., Moreland, K.N., Wham, D.C., Kanke, M.R., and Matz, M.V. (2020) *Cladocopium* community divergence in two *Acropora* coral hosts across multiple spatial scales. *Mol Ecol* **29**: 4559–4572.

Desbruyères, D., Hashimoto, J., and Fabri, M.-C. (2006) Composition and biogeography of hydrothermal vent communities in Western Pacific Back-Arc Basins. In *Geophysical Monograph Series*. Christie, D.M., Fisher, C.R., Lee, S.-M., and Givens, S. (eds). Washington, D.C.: American Geophysical Union, pp. 215–234.

Dubilier, N., Bergin, C., and Lott, C. (2008) Symbiotic diversity in marine animals: the art of harnessing chemosynthesis. *Nat Rev Microbiol* **6**: 725–740.

Edgar, R.C. (2010) Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* **26**: 2460–2461.

Eren, A.M., Maignien, L., Sul, W.J., Murphy, L.G., Grim, S.L., Morrison, H.G., and Sogin, M.L. (2013) Oligotyping: differentiating between closely related microbial taxa using 16S rRNA gene data. *Methods Ecol Evol* **4**: 1111–1119.

Folmer, O., Black, M., Hoeh, W., Lutz, R., and Vrijenhoek, R. (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol* **3**: 294–299.

Geller, J., Meyer, C., Parker, M., and Hawk, H. (2013) Redesign of PCR primers for mitochondrial cytochrome *c* oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. *Mol Ecol Resour* **13**: 851–861.

Gould, A.L. and Dunlap, P.V. (2019) Shedding light on specificity: population genomic structure of a symbiosis between a coral reef fish and luminous bacterium. *Front Microbiol* **10**: 2670.

Ho, P.-T., Park, E., Hong, S.G., Kim, E.-H., Kim, K., Jang, S.-J., Vrijenhoek, R.C., *et al.* (2017) Geographical structure of endosymbiotic bacteria hosted by *Bathymodiolus* mussels at eastern Pacific hydrothermal vents. *BMC Evol Biol* **17**: 121.

Hu, Y.J., and Satten, G.A. (2020) Testing hypotheses about the microbiome using the linear decomposition model (LDM). *Bioinformatics* **36(14)**: 4106–4115.

Johnson, S.B., Warén, A., Tunnicliffe, V., Dover, C.V., Wheat, C.G., Schultz, T.F., and Vrijenhoek, R.C. (2015) Molecular taxonomy and naming of five cryptic species of *Alviniconcha* snails (Gastropoda: Abyssochrysoidea) from hydrothermal vents. *Syst Biodivers* **13**: 278–295.

Kassambara, A. (2020) ggpubr: "ggplot2" Based Publication Ready Plots. https://CRAN.R-project.org/package=ggpubr.

Laming, S.R., Hourdez, S., Cambon-Bonavita, M.A., Pradillon, F. (2020) Classical and computed tomographic anatomical analyses in a not-so-cryptic *Alviniconcha* species complex from hydrothermal vents in the SW Pacific. *Front Zool* **17**: 12.

McKnight, D.T., Huerlimann, R., Bower, D.S., Schwarzkopf, L., Alford, R.A., and Zenger, K.R. (2019) Methods for normalizing microbiome data: An ecological perspective. *Methods Ecol Evol* **10**: 389–400.

McMurdie, P.J. and Holmes, S. (2013) phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. *PLoS ONE* **8**: e61217.

Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020) IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* **37**: 1530–1534.

Minich, J.J., Sanders, J.G., Amir, A., Humphrey, G., Gilbert, J.A., and Knight, R. (2019) Quantifying and understanding well-to-well contamination in microbiome research. *mSystems* **4**: 4.

Mitarai, S., Watanabe, H., Nakajima, Y., Shchepetkin, A.F., and McWilliams, J.C. (2016) Quantifying dispersal from hydrothermal vent fields in the western Pacific Ocean. *Proc Natl Acad Sci USA* **113**: 2976–2981.

Nussbaumer, A.D., Fisher, C.R., and Bright, M. (2006) Horizontal endosymbiont transmission in hydrothermal vent tubeworms. *Nature* **441**: 345–348.

R Core Team (2020) R: A language and environment for statistical computing, Vienna, Austria: R Foundation for Statistical Computing.

Sogin, E.M., Kleiner, M., Borowski, C., Gruber-Vodicka, H.R., and Dubilier, N. (2021) Life in the dark: phylogenetic and physiological diversity of chemosynthetic symbioses. *Annu Rev Microbiol* **75**: 1.

Suzuki, Y., Kojima, S., Watanabe, H., Suzuki, M., Tsuchida, S., Nunoura, T., Hirayama, H., *et al.* (2006a) Single host and symbiont lineages of hydrothermal-vent gastropods *Ifremeria nautilei* (Provannidae): biogeography and evolution. *Mar Ecol Prog Ser* **315**: 167–175.

Suzuki, Y., Kojima, S., Sasaki, T., Suzuki, M., Utsumi, T., Watanabe, H., Urakawa, H., *et al.* (2006b) Host-symbiont relationships in hydrothermal vent gastropods of the genus *Alviniconcha* from the Southwest Pacific. *Appl Environ Microbiol* **72**: 1388–1393.

Trembath-Reichert, E., Butterfield, D.A., and Huber, J.A. (2019) Active subseafloor microbial communities from Mariana back-arc venting fluids share metabolic strategies across different thermal niches and taxa. *ISME J* **13**: 2264–2279.

Ücker, M., Ansorge, R., Sato, Y., Sayavedra, L., Breusing, C., and Dubilier, N. (2021) Deep-sea mussels from a hybrid zone on the Mid-Atlantic Ridge host genetically indistinguishable symbionts. *ISME J* **15**: 3076–3083.

Van Dover, C.L., Humphris, S.E., Fornari, D., Cavanaugh, C.M., Collier, R., Goffredi, S.K., Hashimoto, J., *et al.* (2001) Biogeography and ecological setting of Indian Ocean hydrothermal vents. *Science* **294**: 818–823.

Walters, W., Hyde, E.R., Berg-Lyons, D., Ackermann, G., Humphrey, G., Parada, A., Gilbert, J.A., *et al.* (2015) Improved bacterial 16S rRNA gene (V4 and V4-5) and fungal internal transcribed spacer marker gene primers for microbial community surveys. *mSystems* **1**: e00009-15.

Wentrup, C., Wendeberg, A., Schimak, M., Borowski, C., and Dubilier, N. (2014) Forever competent: deep-sea bivalves are colonized by their chemosynthetic symbionts throughout their lifetime. *Environ Microbiol* **16**: 3699–3713.

Wilkins, L., Leray, M., O'Dea, A., Yuen, B., Peixoto, R. S., Pereira, T. J., Bik, H. M., *et al.* (2019) Host-associated microbiomes drive structure and function of marine ecosystems. *PLoS Biol* **17(11)**: e3000533.

Windoffer, R. and Giere, O. (1997) Symbiosis of the hydrothermal vent gastropod *Ifremeria nautilei* (Provannidae) with endobacteria-structural analyses and ecological considerations. *Biol Bull* **193**: 381–392.

Figure Legends

Fig. 1 Locations for *Alviniconcha* and *Ifremeria* species sampled in this study.

Fig. 2 Fractional abundance plot of symbiont ASVs within individual snails according to *Alviniconcha* and *Ifremeria* species.

Fig. 3 Mid-point rooted IQTREE consensus phylogeny of ASVs within symbiont genera. Node labels indicate ultra-fast bootstrap support values.

Fig. 4 (A) Principal coordinate analysis plot based on weighted UniFrac distances. Data were normalized to proportions before analysis. Numbers in brackets indicate sample sizes for each host taxon. (B) Alpha diversity within host species based on Shannon's and Simpson's diversity index.

Fig. 5 Symbiont ASV accumulation curves.

Table 1 Results for linear decomposition models (LDM) and PERMANOVAs based on weighted UniFrac distances. Three different models were run to assess the effects of DNA preservation, extraction and sequencing method on patterns of symbiont diversity: 1) Model including the complete dataset and controlling for effects of methodology, 2) Model restricted to *A. boucheti*, *A. kojimai* and *A. strummeri* from the ELSC and controlling for effects of methodology, 3) Model restricted to *A. boucheti*, *A. kojimai* and *A. strummeri* from the ELSC and including methodology as main explanatory factor. Sources of variation are shown in sequential order tested in the model. Significant sources of variation are indicated in bold. df = degrees of freedom, F = F statistic, VE = explained variation, p = p value.

	LDM				PERMANOVA	
Source of variation	**df**	**F**	**VE [%]**	***p***	**F**	***p***
Model 1						
Geographic region	8	2.2861	16.49	**0.0001**	324.710	**0.0010**
Host	3	1.3426	25.82	**0.0001**	704.879	**0.0010**
Model 2						
Vent	2	3.5363	30.39	**0.0001**	4959.805	**0.0010**
Host	2	6.1004	52.42	**0.0001**	6424.714	**0.0010**
Model 3						
Methodology	1	0.1179	1.99	**0.0001**	264.664	**0.0010**
Vent	2	3.5363	29.79	**0.0001**	9919.611	**0.0010**
Host	2	6.1004	51.38	**0.0001**	12849.428	**0.0010**

Figure 1

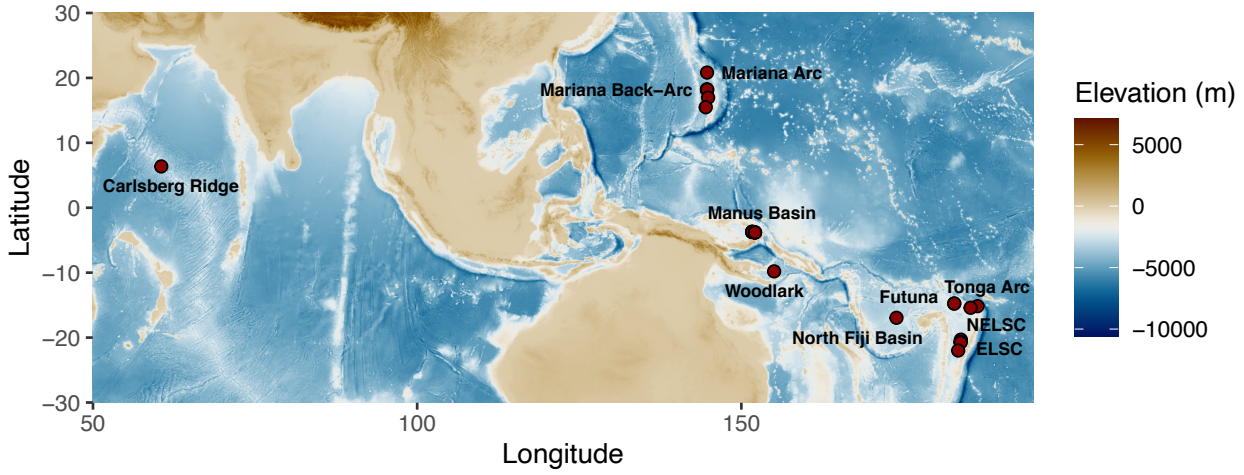
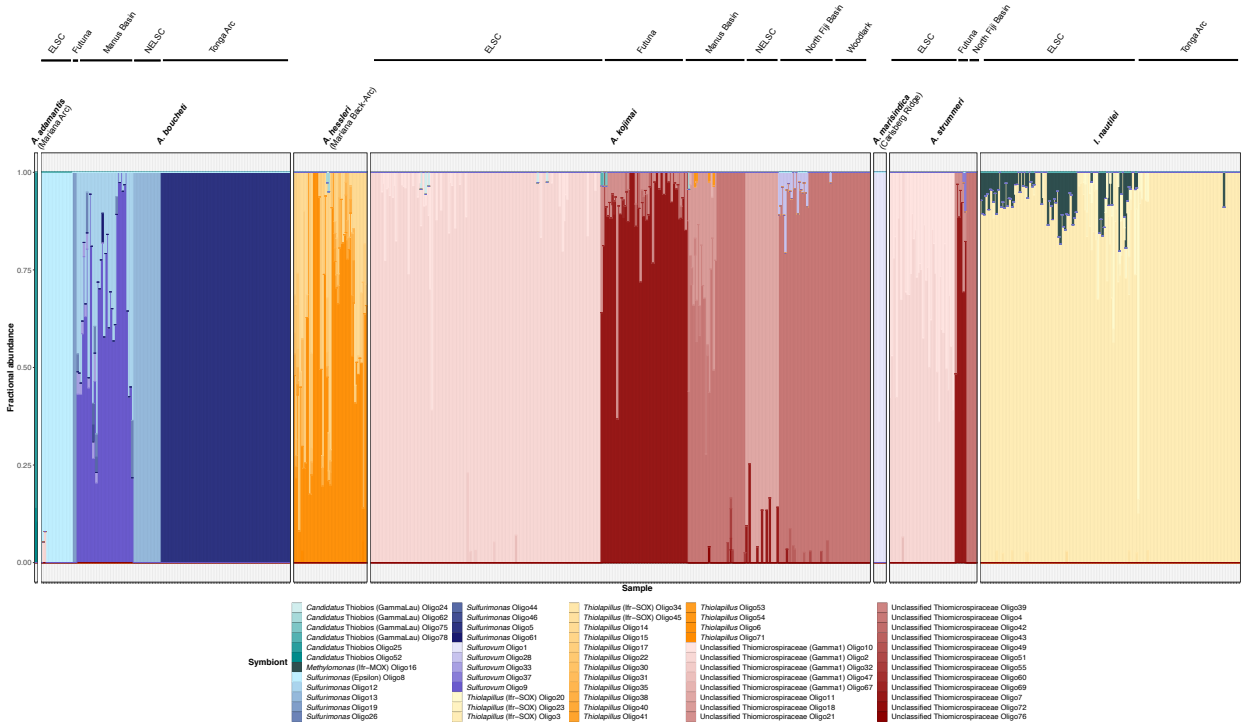
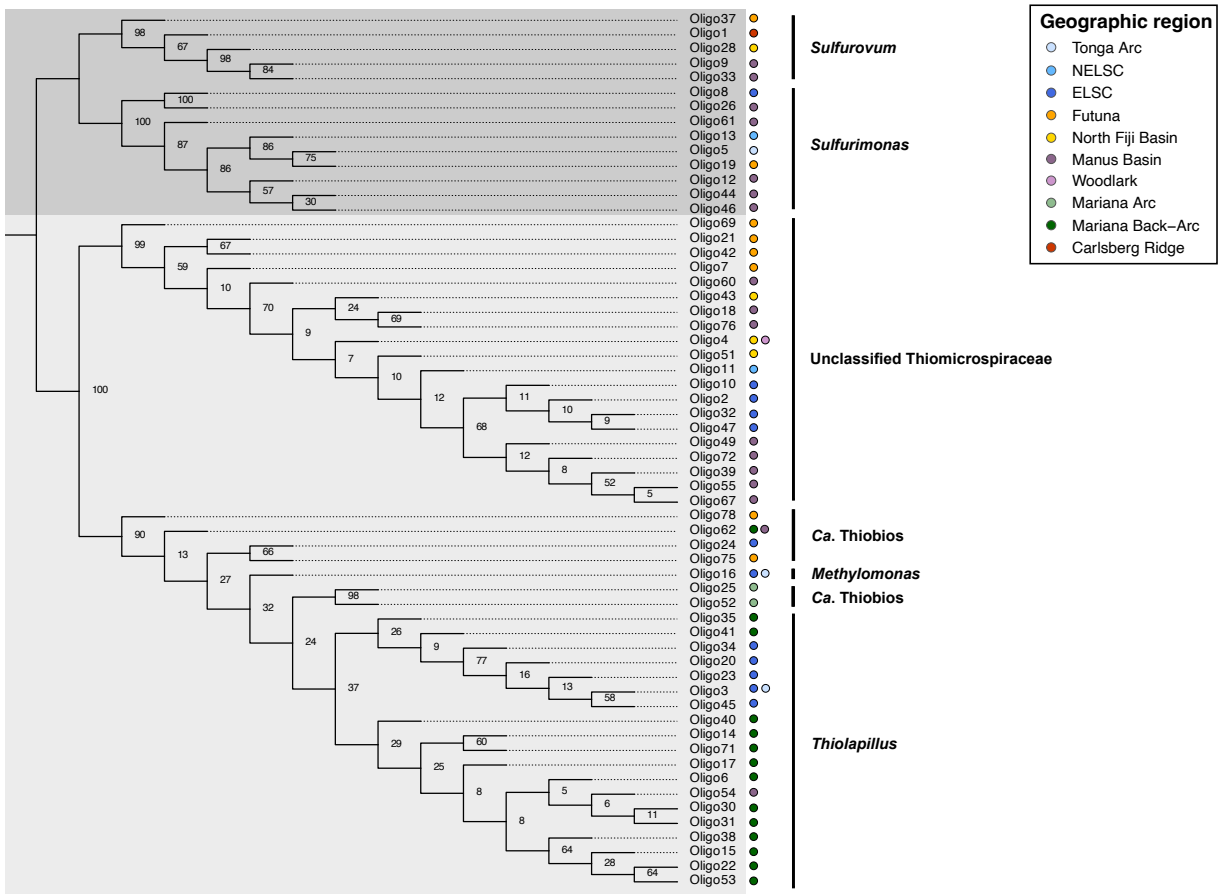


Figure 2


ELSC
Futuna
Manus Basin
NELSC
Tonga Arc
North Fiji Basin
Woodlark
A. adenensis (Manus Arc)
A. bouchet
A. hessleri (Futuna Back-Arc)
A. kojimai
A. marsaglia (Caldera Ridge)
A. strummeri
I. nautilei
Fractional abundance
1.00
0.75
0.50
0.25
0.00
Sample
Symbiont
Candidatus Thiobios (GammaLau) Oligo24
Candidatus Thiobios (GammaLau) Oligo62
Candidatus Thiobios (GammaLau) Oligo75
Candidatus Thiobios (GammaLau) Oligo78
Candidatus Thiobios Oligo25
Candidatus Thiobios Oligo52
Methylomonas (Ifr-MOX) Oligo16
Sulfurimonas (Epsilon) Oligo8
Sulfurimonas Oligo12
Sulfurimonas Oligo13
Sulfurimonas Oligo19
Sulfurimonas Oligo26
Sulfurimonas Oligo44
Sulfurimonas Oligo46
Sulfurimonas Oligo5
Sulfurimonas Oligo61
Sulfurovum Oligo1
Sulfurovum Oligo28
Sulfurovum Oligo33
Sulfurovum Oligo37
Sulfurovum Oligo9
Thiolapillus (Ifr-SOX) Oligo20
Thiolapillus (Ifr-SOX) Oligo23
Thiolapillus (Ifr-SOX) Oligo3
Thiolapillus (Ifr-SOX) Oligo34
Thiolapillus (Ifr-SOX) Oligo45
Thiolapillus Oligo14
Thiolapillus Oligo15
Thiolapillus Oligo17
Thiolapillus Oligo22
Thiolapillus Oligo30
Thiolapillus Oligo31
Thiolapillus Oligo35
Thiolapillus Oligo38
Thiolapillus Oligo40
Thiolapillus Oligo41
Thiolapillus Oligo53
Thiolapillus Oligo54
Thiolapillus Oligo6
Thiolapillus Oligo71
Unclassified Thiomicrospiraceae (Gamma1) Oligo10
Unclassified Thiomicrospiraceae (Gamma1) Oligo2
Unclassified Thiomicrospiraceae (Gamma1) Oligo32
Unclassified Thiomicrospiraceae (Gamma1) Oligo47
Unclassified Thiomicrospiraceae (Gamma1) Oligo67
Unclassified Thiomicrospiraceae Oligo11
Unclassified Thiomicrospiraceae Oligo18
Unclassified Thiomicrospiraceae Oligo21
Unclassified Thiomicrospiraceae Oligo39
Unclassified Thiomicrospiraceae Oligo4
Unclassified Thiomicrospiraceae Oligo42
Unclassified Thiomicrospiraceae Oligo43
Unclassified Thiomicrospiraceae Oligo49
Unclassified Thiomicrospiraceae Oligo51
Unclassified Thiomicrospiraceae Oligo55
Unclassified Thiomicrospiraceae Oligo60
Unclassified Thiomicrospiraceae Oligo69
Unclassified Thiomicrospiraceae Oligo7
Unclassified Thiomicrospiraceae Oligo72
Unclassified Thiomicrospiraceae Oligo76

451 **Figure 3**

452

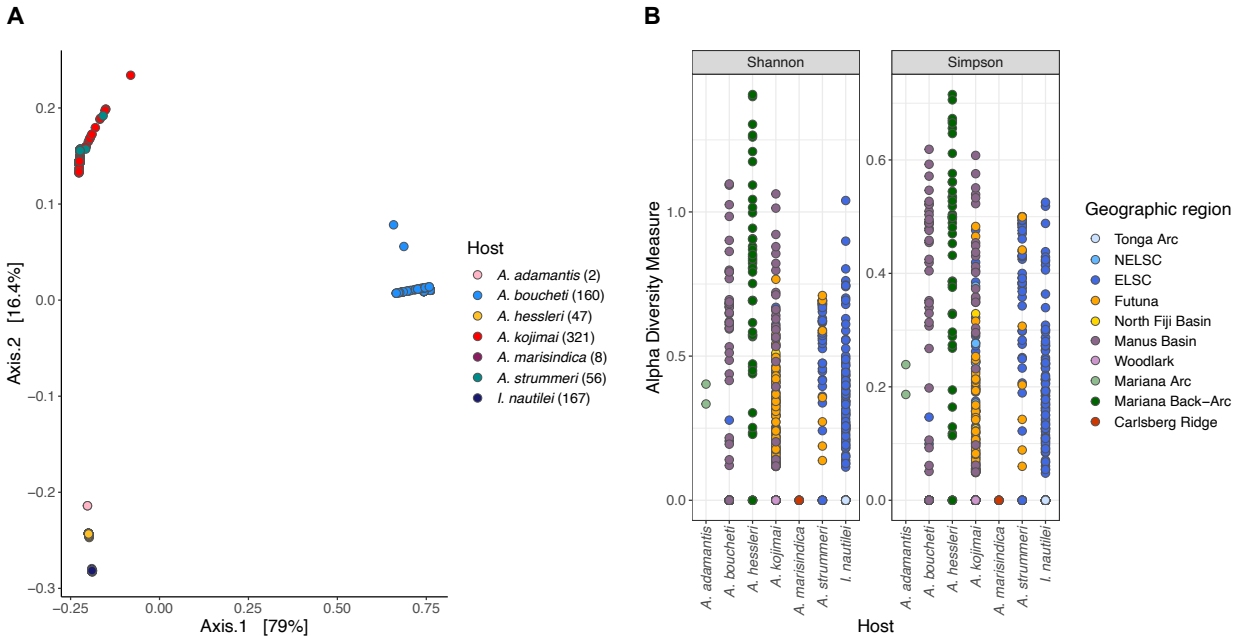


453

454

455 **Figure 4**

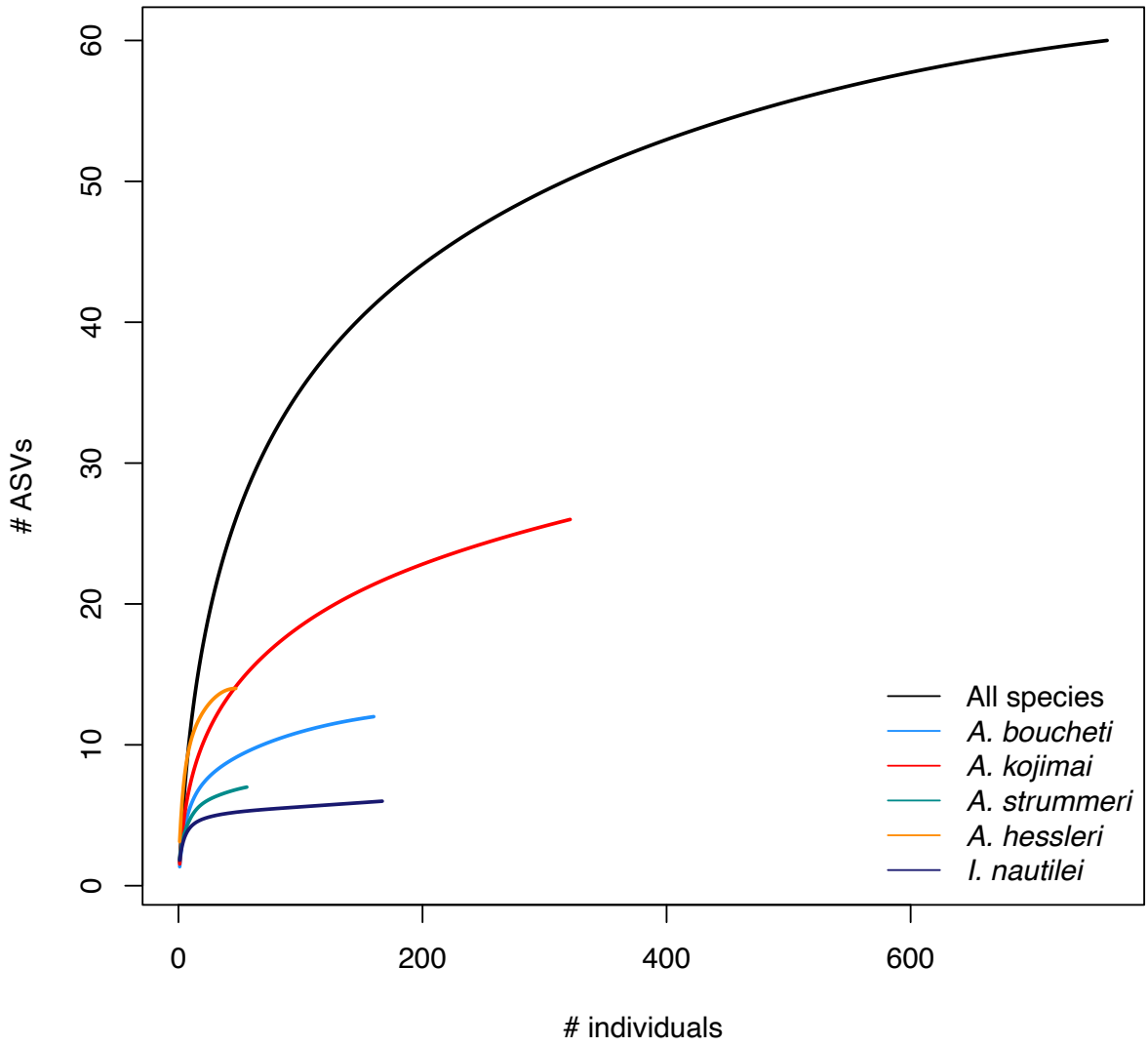
456



457

458

459 **Figure 5**



460

461