

Supplementary Material and Figures:

JAMP pipeline rationale

Using a preliminary data set, we compared JAMP (Elbrecht 2020), Leray's CROP pipeline (Leray and Knowlton 2017), OBITools using Swarm clustering (Mahe et al. 2015; Boyer et al. 2016; Wangensteen and Turon 2017), and an in-house custom pipeline integrating various programs but using Vsearch (Rognes et al. 2016) for clustering followed by a LULU post-cluster step (Froslev et al. 2017). Overall output and general trends were similar enough among pipelines that it did not seem to matter which was used. Only the rare things differed among pipelines, which is known to occur even when using the same pipeline on replicated data (Leray and Knowlton 2017). Thus, we chose JAMP for the following reasons: 1) It's an R modular package making it pretty user-friendly; 2) The package creator, Vasco Elbrecht, has thoroughly tested the pipeline for a variety of metabarcoding applications; and 3) we really like the metadata integration at each step to support the examination of your fasta files enabling you to better understand what is happening, how much you are losing (in terms of sequences), and the resulting output. Ultimately, the choice of pipeline is up to the user and in some cases can alter results (e.g., Shafer et al. 2017; Pauvert et al. 2019) but a formal comparison of pipelines is beyond the scope of this work. In our case, we selected JAMP as the easiest and fastest pipeline to use based on our initial comparison, and it performed well in this comparison.

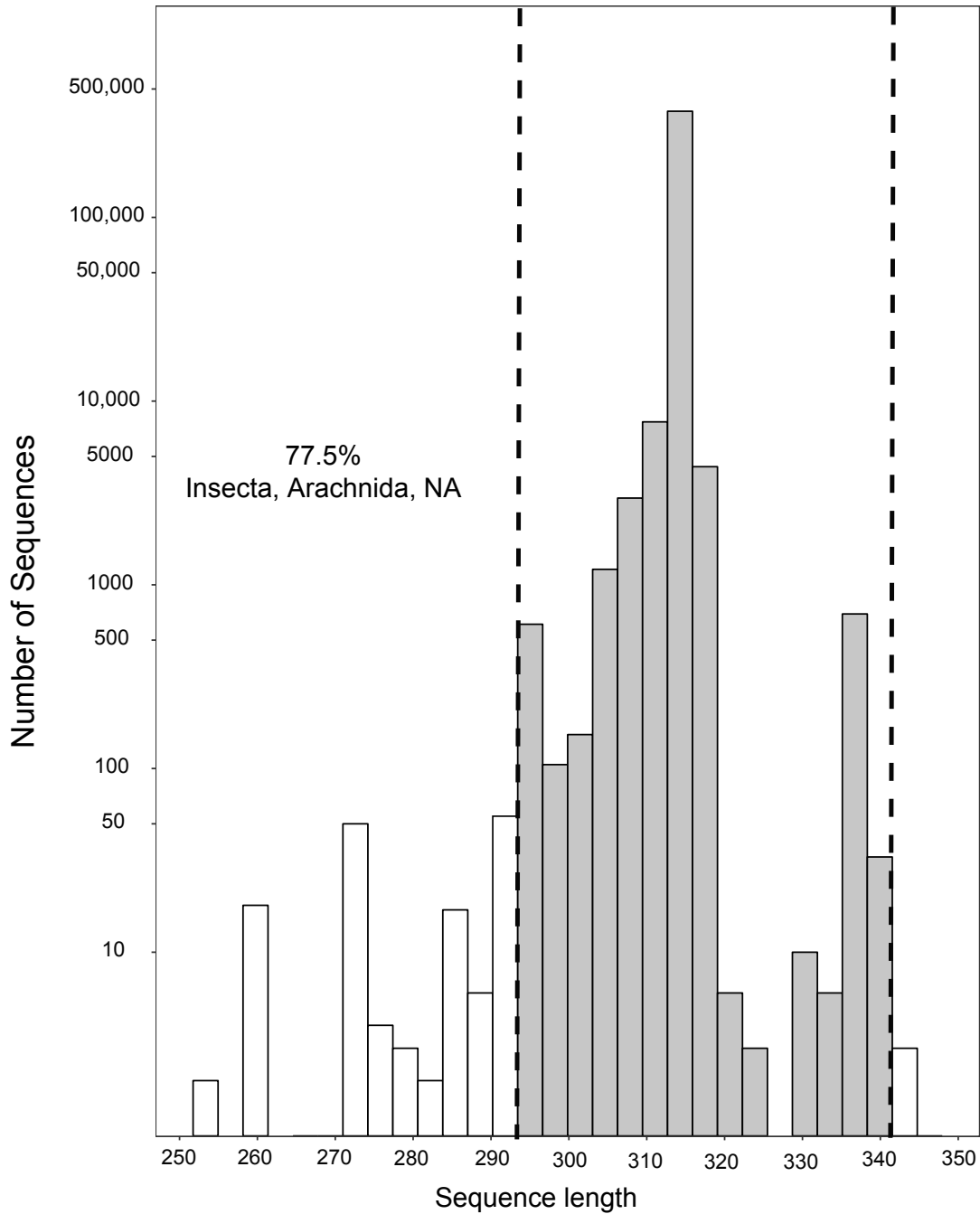
References:

- Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). OBITOOLS: a UNIX-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, 16(1), 176–182. <https://doi.org/10.1111/1755-0998.12428>
- Frøslev, T.G., Kjøller, R., Bruun, H.H. *et al.* Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. *Nat Commun* 8, 1188 (2017). <https://doi.org/10.1038/s41467-017-01312-x>
- Leray, M., & Knowlton, N. (2017). Random sampling causes the low reproducibility of rare eukaryotic OTUs in Illumina COI metabarcoding. *PeerJ*, 5, e3006. <https://doi.org/10.7717/peerj.3006>
- Mahé, F., Rognes, T., Quince, C., de Vargas, C., & Dunthorn, M. (2015). Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ*, 3, e1420. <https://doi.org/10.7717/peerj.1420>
- Pauvert, C., Buée, M., Laval, V., Edel-Hermann, V., Fauchery, L., Gautier, A., Lesur, I., Vallance, J. and Vacher, C., 2019. Bioinformatics matters: the accuracy of plant and soil fungal community data is highly dependent on the metabarcoding pipeline. *Fungal Ecology*, 41, 23-33. <https://doi.org/10.1016/j.funeco.2019.03.005>
- Rognes, T., Flouri, T., Nichols, B., Quince, C., & Mahé, F. (2016). VSEARCH: A versatile open source tool for metagenomics. *PeerJ*, 4, e2584. <https://doi.org/10.7717/peerj.2584>
- Shafer, A.B., Peart, C.R., Tusso, S., Maayan, I., Brelsford, A., Wheat, C.W. and Wolf, J.B., 2017. Bioinformatic processing of RAD-seq data dramatically impacts downstream population genetic inference. *Methods in Ecology and Evolution*, 8, 907-917. <https://doi.org/10.1111/2041-210X.12700>
- Wangensteen O.S., & Turon X. (2017) Metabarcoding Techniques for Assessing Biodiversity of Marine Animal Forests. In: Rossi S., Bramanti L., Gori A., Orejas C. (eds) Marine Animal Forests. Springer, Cham. https://doi.org/10.1007/978-3-319-21012-4_53
- Vasco Elbrecht (2020). JAMP: Just Another Metabarcoding Pipeline. R package version 0.77.

Figure S1: Tank set-up at the Hawai'i Institute of Marine Biology in Kāne'ōhe Bay on the island of O'ahu and an image of the 2-tiered modified ARMS units that were placed within the tanks.



Figure S2 – cutadapt length filtration rationale: The frequency and length distribution of sequences within the primer binding site using in silico PCR on the Midori unique database. Nearly 400K sequences were obtained and scored against a profile HMM to remove any spurious virtual amplicons (Wilkinson 2018). While most sequences were around 313 bp, there were a number of sequences at 337 bp assigned to the phylum Platyhelminthes and at 295 bp assigned to bivalves and trematodes. As a result, we used a min/max 295/340 bp cutoff represented by the gray bars within the dashed lines. Sequences outside of the dashes were removed as 78% were classified as Insecta, Arachnida, or NA - non-marine eukaryotes.



Wilkinson, S., Davy, S., Bunce, M., & Stat, M. (2018). Taxonomic identification of environmental DNA with informatic sequence classification trees. *PeerJ*. <https://doi.org/10.7287/peerj.preprints.26812v1>

Figure S3: Comparison of COI and 28S rRNA neighbor-joining phylogenetic trees. The 28S rRNA tree is missing *Ancorinidae* sp. 2, *Hymeniacidon* sp. 5, *Poecilosclerida* sp. 4, *Oscarella* sp. 6, *Leucosolenida* sp. 18, and *Heteroscleromorpha* sp. 1 due to short sequence length. Coloration represents the three sponge classes: yellow = Homoscleromorpha, light blue = Demospongiae, and light purple = Calcarea. Dashed lines separate the taxa within the orders from each tree.

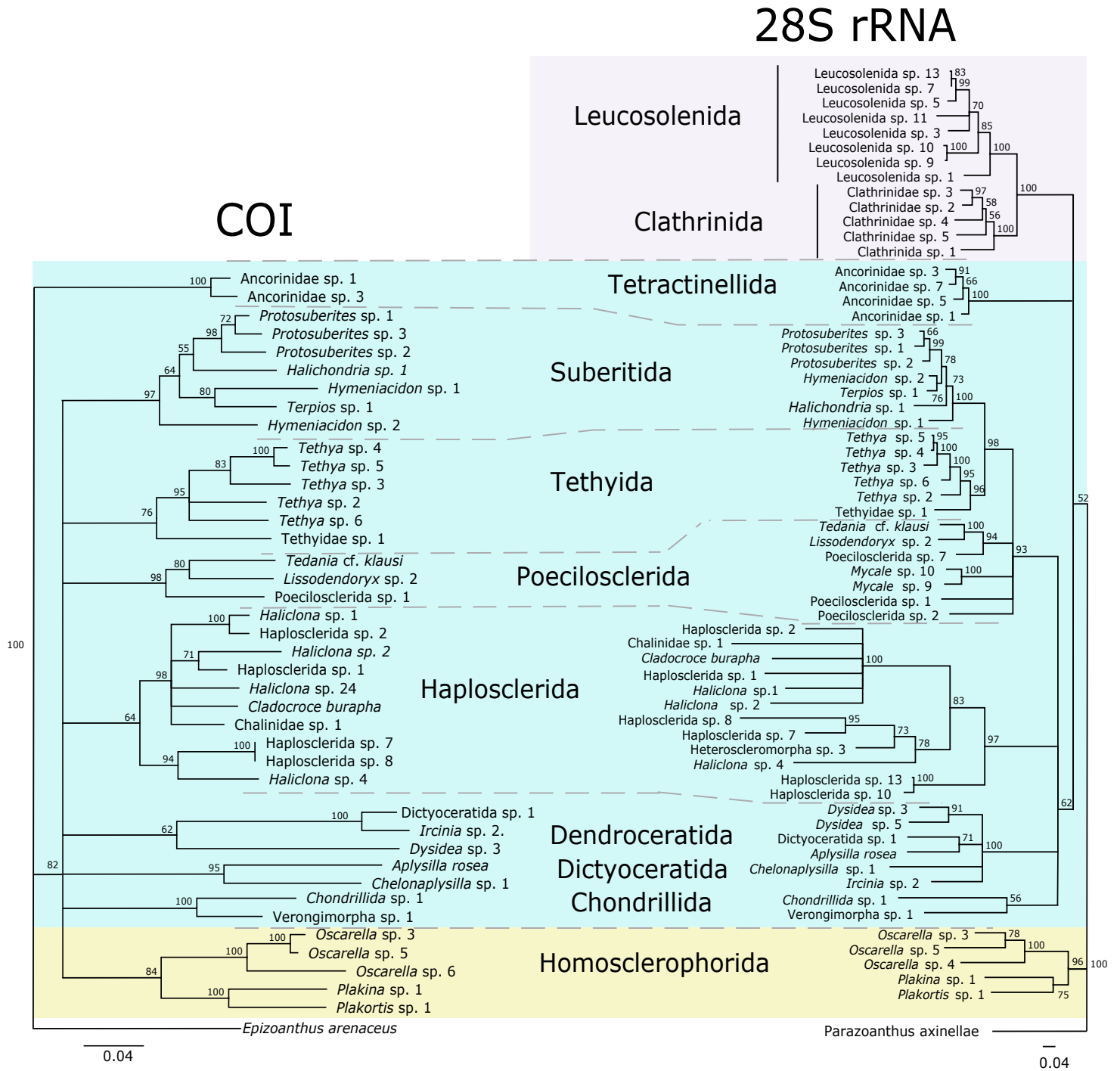


Figure S4: (a) All suberitids found in this study express the color yellow naturally or under stress making it challenging to identify. (b) Additional sponges that express yellow within Poecilosclerida and the sponge *Biemna fistulosa* found in the metabarcoding data but not from the barcoded morphologies which could have been overlooked by the taxonomist.

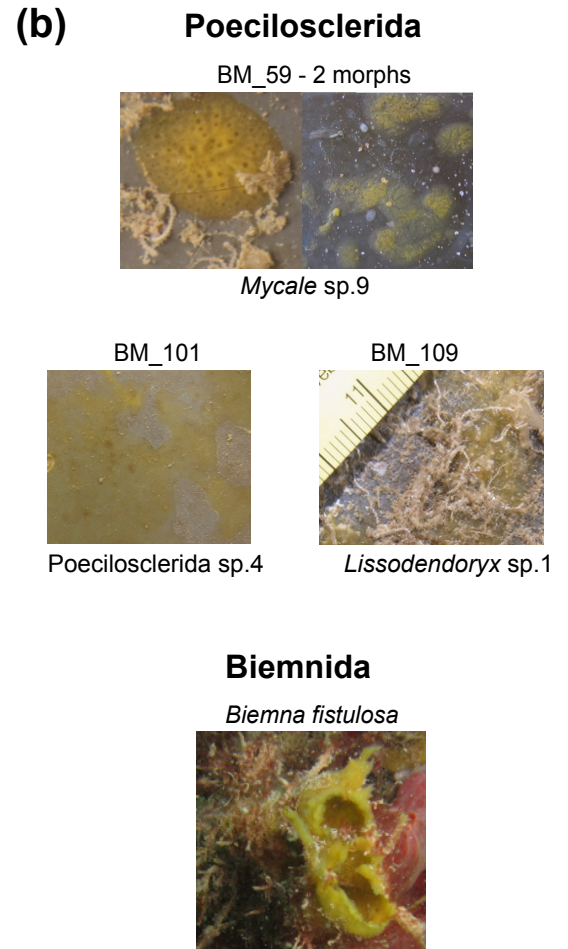
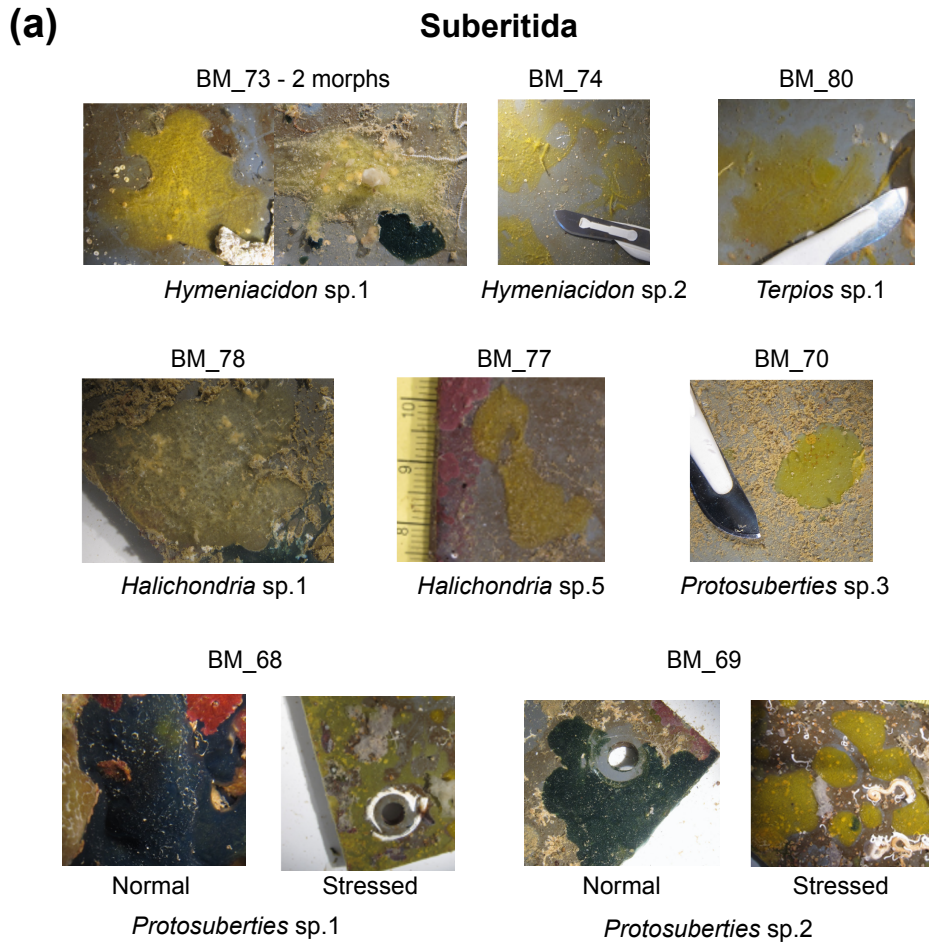


Figure S5: Images represent the top and bottom sides of the three plates from ARMS Unit 13. The sponge *Plakortis* sp. 1, outlined in green, dominated Unit 13 both in terms of plate coverage and metabarcoding reads (65735 reads – Table S3).

Plate 1-Top

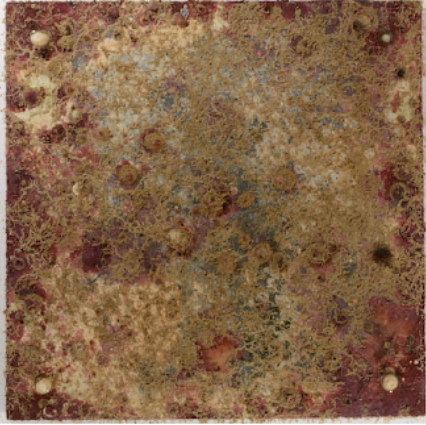


Plate 1-Bottom

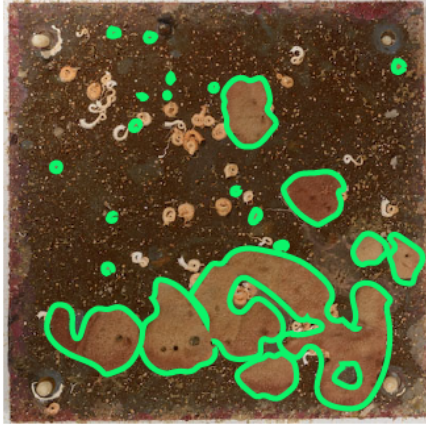


Plate 2-Top



Plate 2-Bottom

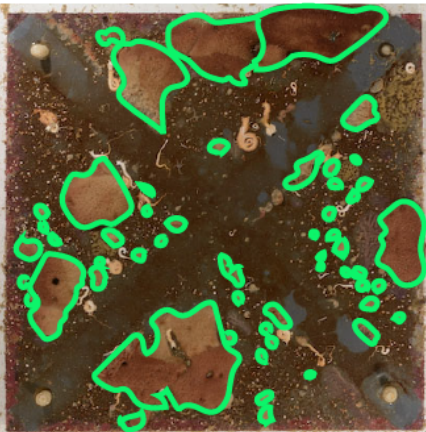


Plate 3-Top

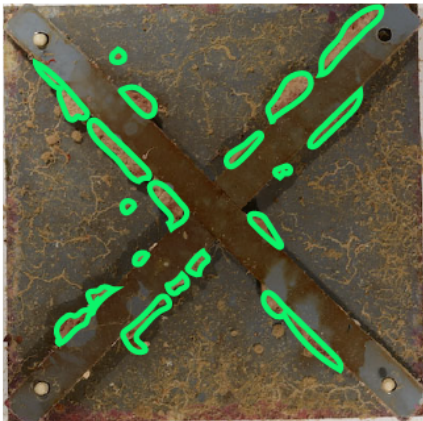


Plate 3-Bottom



Figure S6: Plate imagery of ARMS Unit 5 with all sponges found on the bottom of Plate 2 outlined in respective colors to represent the abundance and variety of sponges on one plate surface. The *Veronigimorpha* sp. 1, BM_89, was only found in this ARMS unit in small quantities on three of the plate surfaces (colored in red) and has an extremely low biomass, as demonstrated in the below right image showing its sheer layer over vermetid mollusks and serpulid worms. Regardless, this sponge was also detected by metabarcoding to be present only in this ARMS unit.

