Appendix S1

Metabolic differentiation facilitates coexistence in two coral reef fish species

Simon J. Brandl1,2,3,4, Calvin N. Quigley5, Jordan M. Casey2,3,Alexandre Mercière3, Nina M.D. Schiettekatte2,3, Tommy Norin6, Valeriano Parravicini2,3, Isabelle M. Côté1

*1 Department of Biological Sciences, Simon Fraser University, Burnaby, BC, Canada*

*2 PSL Université Paris: EPHE-UPVD-CNRS, USR 3278 CRIOBE, Université de Perpignan, 66860 Perpignan, France*

*3 Laboratoire d’Excellence “CORAIL,” Perpignan, France*

*4 CESAB, Centre for the Synthesis and Analysis of Biodiversity, Institut Bouisson Bertrand, 34000 Montpellier*

*5 School of Biological Sciences, Victoria University of Wellington, Wellington, New Zealand*

*6 DTU Aqua: National Institute of Aquatic Resources, Technical University of Denmark, 2800 Kgs. Lyngby, Denmark*

Email: Simon J. Brandl ([simonjbrandl@gmail.com](mailto:simonjbrandl@gmail.com))

*Gut content DNA metabarcoding*

Library preparation and sequencing were performed by Jonah Ventures (Boulder, Colorado, USA). We targeted the mitochondrial cytochrome *c* oxidase subunit I (COI) gene region, which performs well across metazoans, and the chloroplast 23S rRNA gene region, which performs well across autotrophs. For the COI markers, we used the m1COIintF forward primer (Leray et al. 2013) and the jgHCO2198 reverse primer (Geller et al. 2013). For the 23S markers, we used the p23SrV\_f1 and Diam23Sr1 23S primers (Sherwood and Presting 2007, Hamsher et al. 2011, Cannon et al. 2016). We ran two-step PCR reactions at 25 μL according to the Promega PCR Master Mix guidelines (Promega, Madison, Wisconsin, USA): 12.5 μL Master Mix, 0.5 μM of each primer, 1 μL gDNA, and 10.5 μL DNase/Rnase-free water. We followed the following conditions for PCR amplification for the COI marker: denaturation at 94°C for 2 min, followed by 45 cycles of 15 s at 94°C, 30 s at 50°C, and 1 min at 72°C, then a final elongation at 72°C for 10 min. We used the following conditions for PCR amplification for the 23S marker: denaturation at 94°C for 3 min, followed by 40 cycles of 30 s at 94°C, 45 s at 55°C, and 1 min at 72°C, then a final elongation at 72°C for 10 min. Following the PCR amplifications, we visually inspected each reaction with a 2% agarose gel.

The following library preparation and sequencing protocols apply to both the COI and 23S gene. We performed clean-ups by incubating amplicons with Exo1/SAP for 30 min at 37°C, followed by inactivation at 95°C for 5 min. We then stored the products at -20°C. Afterwards, we ran a second indexing PCR to bind a unique 12-nucleotide index sequence, which included Promega Master mix, 0.5 μM of each primer, and 2 μL of template DNA under the following conditions: denaturation at 95°C for 3 min, followed by 8 cycles at 95°C for 30 s, 55°C for 30 s, and 72°C for 30 s. Each reaction was visually inspected with a 2% agarose gel to ensure successful amplification.

We cleaned and normalized 25 μL of each indexed amplicon with the SequalPrep Normalization Kit (Life Technologies, Carlsbad, California, USA), pooled 5 μL of each sample, and sent the library pools to the Genohub service provider (Austin, Texas, USA). Bead cleaning with Agencourt AMPure XP beads (Beckman Coulter, Brea, California, USA) was performed to remove < 200 bp amplicons, samples were quantified with a Qubit Fluorometer (Invitrogen, Carlsbad, California, USA), and amplicon average size was analyzed with an Agilent TapeStation 4200 (Agilent, Santa Clara, California, USA). Finally, sequencing was performed on an Illumina HiSeq using the HiSeq Rapid SBS Kit v2, 500-cycles (Illumina, San Diego, California, USA).

*Sequence bioinformatics*

We processed all raw COI and 23S sequences with the JAMP pipeline (<https://github.com/VascoElbrecht/JAMP>). After demultiplexing, forward and reverse reads were pair-end merged with USEARCH v11.0.667 (Edgar 2010), and primers were trimmed from both ends using Cutadapt v1.18 (Martin 2011). Quality filtering was performed with expected error filtering through USEARCH (Edgar and Flyvbjerg 2015), and all sequences affected by sequencing and PCR error were removed using the UNOISE algorithm (Edgar 2016). Exact sequence variants (ESVs) were then compiled into ESV tables, and taxonomy was assigned to each ESV using the basic local alignment search tool (BLAST) hit when query coverage and percent identity exceeded 95% and 80%, respectively, or by mapping them against a 23S database from Silva (Yilmaz et al. 2013), specifying zero deviations to ensure mapping accuracy. Consensus taxonomy was generated from the hit tables, first considering 100% matches, then decreasing by 1% until hits were available for each ESV. Taxonomy that was present in at least 90% of the hits was reported; otherwise, an “NA” was assigned when several different taxa matched the ESV. For error reduction due to misidentified taxa, the bracket was increased to 2% when matches of 97% and higher were present, but no family-level or lower taxonomy was assigned. We then removed all self-hits from the COI dataset, which resulted from sequencing the host tissue, and omitted all ESVs that only had a single sequence read across the entire dataset. For the network analysis, we also removed all ESVs that were only used by a single individual to ensure convergence of the modularity metric. Further detail on the treatment of the assembled sequencing datasets is provided in the appended *R* script.

Literature Cited

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

Cannon, M., J. Hester, A. Shalkhauser, E. R. Chan, K. Logue, S. T. Small, and D. Serre. 2016. In silico assessment of primers for eDNA studies using PrimerTree and application to characterize the biodiversity surrounding the Cuyahoga River. Scientific reports 6:22908.

Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Pena, J. K. Goodrich, and J. I. Gordon. 2010. QIIME allows analysis of high-throughput community sequencing data. Nature methods 7:335.

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

Edgar, R. C. 2013. UPARSE: highly accurate OTU sequences from microbial amplicon reads. Nature methods 10:996.

Edgar, R. C. 2016. UNOISE2: improved error-correction for Illumina 16S and ITS amplicon sequencing. BioRxiv:081257.

Edgar, R. C., and H. Flyvbjerg. 2015. Error filtering, pair assembly and error correction for next-generation sequencing reads. Bioinformatics 31:3476–3482.

Geller, J., C. Meyer, M. Parker, and H. Hawk. 2013. Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all‐taxa biotic surveys. Molecular ecology resources 13:851–861.

Hamsher, S. E., K. M. Evans, D. G. Mann, A. Poulíčková, and G. W. Saunders. 2011. Barcoding diatoms: exploring alternatives to COI-5P. Protist 162:405–422.

Leray, M., J. Y. Yang, C. P. Meyer, S. C. Mills, N. Agudelo, V. Ranwez, J. T. Boehm, and R. J. Machida. 2013. A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: application for characterizing coral reef fish gut contents. Frontiers in zoology 10:34.

Martin, M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet. journal 17:10–12.

Sherwood, A. R., and G. G. Presting. 2007. Universal primers amplify a 23S rDNA plastid marker in eukaryotic algae and cyanobacteria. Journal of phycology 43:605–608.

Yilmaz, P., L. W. Parfrey, P. Yarza, J. Gerken, E. Pruesse, C. Quast, T. Schweer, J. Peplies, W. Ludwig, and F. O. Glöckner. 2013. The SILVA and “all-species living tree project (LTP)” taxonomic frameworks. Nucleic acids research 42:D643–D648.