

# Mitochondrial cytochrome b sequence data are not an improvement for species identification in scleractinian corals

John P. Wares

Department of Genetics, University of Georgia, Athens, GA, USA

## ABSTRACT

There are well-known difficulties in using the cytochrome oxidase I (COI) mitochondrial gene region for population genetics and DNA barcoding in corals. A recent study of species divergence in the endemic Caribbean genus *Agaricia* reinforced such knowledge. However, the growing availability of whole mitochondrial genomes may help indicate more promising gene regions for species delineation. I assembled the whole mitochondrial genome for *Agaricia fragilis* from Illumina single-end 250 bp reads and compared this sequence to that of the congener *A. humilis*. Although these data suggest that the cytochrome b (CYB) gene region is more promising, comparison of available CYB sequence data from scleractinian and other reef-building corals indicates that multilocus approaches are still probably necessary for phylogenetic and population genetic analysis of recently-diverged coral taxa.

**Subjects** Biodiversity, Evolutionary Studies, Genetics, Marine Biology, Zoology

**Keywords** Coral, Barcoding, Mitochondrial DNA, Species

## INTRODUCTION

Coral reefs are widely recognized as being important representatives and biogenic harbors of biodiversity (*Plaisance et al., 2011*). At the same time as coral reefs are in crisis due to disease and habitat change, there is still new diversity being explored (*Breedy, Williams & Guzman, 2013; Breedy & Guzman, 2014*). With the incredible phenotypic diversity that may be found in an individual taxon (*Bruno & Edmunds, 1997; Veron, 2011*), as well as the potential for hybridization among taxa (*Vollmer & Palumbi, 2002*), biologists often turn to molecular techniques for separating taxa.

In many animal species, this approach has been relatively straightforward and has often relied on a single gene that is both highly variable at silent nucleotide positions as well as highly conserved for amino acid sequence (*Folmer et al., 1994; Hebert et al., 2003*). This combination allows the sequencing of this gene with universal primers, yet the discovery of tremendous amounts of nucleotide variation that may be used to distinguish taxa. However, this gene region has proven nearly useless in corals (*Shearer & Coffroth, 2008*). Researchers have also tried using other protein-coding loci such as *atp6* (*Bongaerts et al., 2013*) to explore phylogenetic diversity in Agariciid corals, but still tended to recover nonmonophyletic taxa. Similarly, *Meyers (2013)* showed that using intron regions within the mitochondrial ND5 locus (*Concepcion, Medina & Toonen, 2006*) could not resolve many species in the genus *Agaricia*.

Submitted 11 July 2014  
Accepted 16 August 2014  
Published 9 September 2014

Corresponding author  
John P. Wares, [jpwares@uga.edu](mailto:jpwares@uga.edu)

Academic editor  
Mónica Medina

Additional Information and  
Declarations can be found on  
page 5

DOI [10.7717/peerj.564](https://doi.org/10.7717/peerj.564)

© Copyright  
2014 Wares

Distributed under  
Creative Commons CC-BY 4.0

**OPEN ACCESS**

The focus for such work has often been mitochondrial regions because the DNA is abundant in animal tissues, often variable within and among populations, and the lower effective size of the mitochondrial genome—a haploid genome that is typically maternally inherited—tends to result in diagnostic nucleotide characters for a population in less time than for a nuclear locus (Avisé, 2000). For both historical and empirical reasons, some groups of systematists and population geneticists have widely used other mitochondrial regions with success. Population genetics in fishes, for example, frequently explore cytochrome b or ND4, and some have used non-coding regions (e.g., ribosomal or the D-loop origin of replication) (Muss et al., 2001; Taylor & Hellberg, 2003; Hyde & Vetter, 2007).

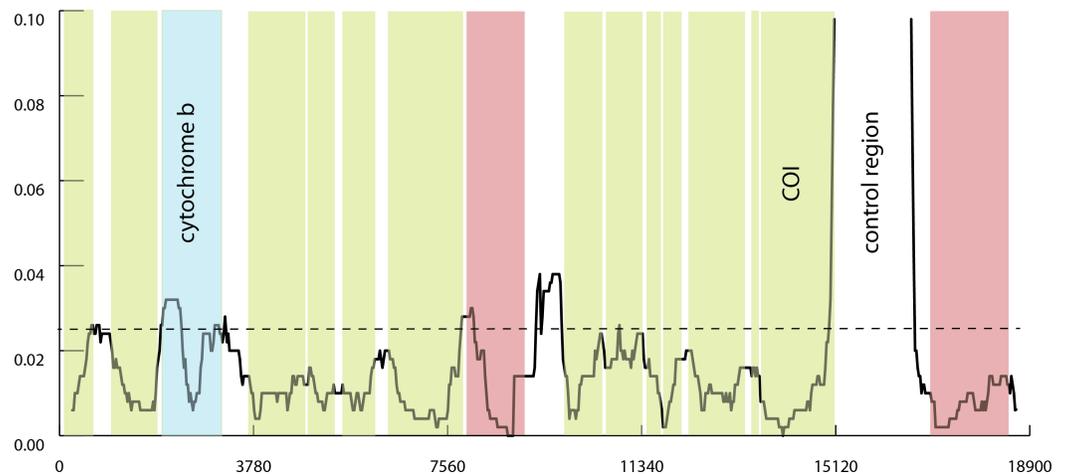
The brief goal of this study is to attempt to identify another useful mitochondrial region for population genetics and systematics studies in scleractinian corals. Here I focus on the Agariciidae; no complete phylogeny yet sufficiently resolves the endemic Caribbean genus *Agaricia* (Meyers, 2013; Bongaerts et al., 2013), and overall the family is an important one for reef development but needs further exploration of its biogeographic and phylogenetic history (Luck et al., 2013). Thus, this study first compares whole mitochondrial sequences between two divergent taxa of *Agaricia* (*A. fragilis* and *A. humilis*), and identifies the most divergent protein-coding region (using coding regions for increased likelihood of conserved primer development). I then analyze divergence in this region (cytochrome b, CYB) across available scleractinian data to show that this region alone is unlikely to improve our ability to separate taxa using DNA sequence-based methods.

## METHODS

To identify potential regions on the easily-sequenced mitochondrial locus, a single individual of *A. fragilis* (AS1943, collected in the Upper Florida Keys and detailed in Meyers, Porter & Wares (2013)) was shotgun sequenced with a single Illumina MiSeq library preparation as in Wares (2013). Resultant single-end 250nt reads were trimmed and mapped to the *A. humilis* mitochondrial genome (GenBank [DQ643831](#)) using Geneious 7.1.4 (Biomatters). The alignment process included up to 5 iterations, with maximum gapped sites per read of 10%, maximum mismatches per read of 20% and a minimum overlap of 20 nucleotides. Annotation of this assembled genome was initiated using MITOS (Bernt et al., 2013) and corrected via re-alignment with the *A. humilis* sequence.

Aligned coding sequences were evaluated for K2P divergence between the two genomes using PAUP\*4.0b10 (Swofford, 2000) as in Shearer & Coffroth (2008); a sliding-window measure of divergence was calculated for 500-bp regions in 25-bp increments along the whole mitochondrial genome.

Subsequently, sequence data for the CYB locus from coral studies that included ‘Scleractinia’ (these studies often include other taxa from the Anthozoan subclass Hexacorallia) were downloaded from GenBank using Geneious 7.1.4 and aligned in CodonCode Aligner v4.2.2. Again, K2P distances among all sequences were obtained using PAUP\*, and all pairwise distances were coded as conspecific, congeneric (excluding



**Figure 1** Divergence of mitochondrial genomes in *Agaricia*. Sliding window divergence between mitochondrial genomes of *A. fragilis* and *A. humilis*. Window size was 500 bp, measured every 25 bp. Coding regions are shaded in green; ribosomal regions in red. Other non-coding regions (tRNA and the intron region for ND5) not indicated. Cytochrome b is shaded blue and harbors highest mean divergence of 0.024 (dotted line).

conspecific), or “other”. The distances observed for these 3 classes of comparison were density plotted using ggplot2 in the R computational environment.

## RESULTS

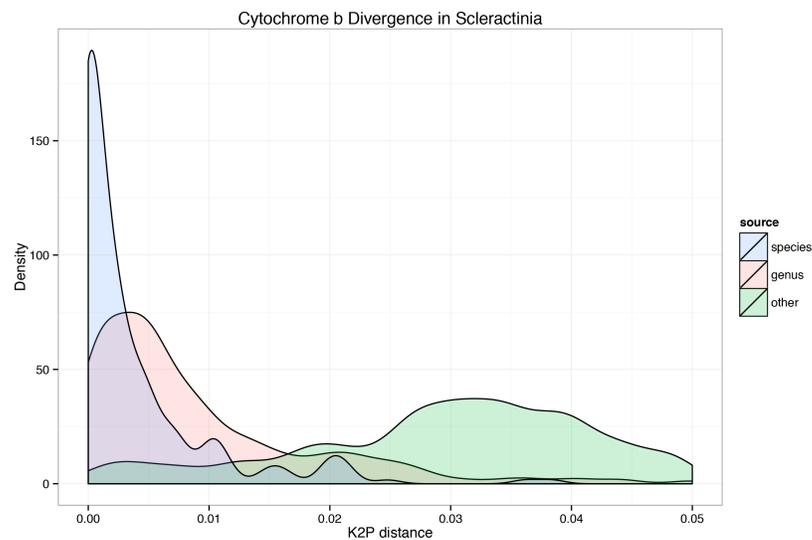
Illumina sequencing of the *A. fragilis* genomic DNA library resulted in a total of 31,957,468 reads. Mapping these reads to the *A. humilis* mitochondrial genome generated a single contig of 18,667 bp. The completed *A. fragilis* mitochondrial genome (Genbank [KM051016](#)) had no observed gene rearrangements and is consistent with the standard type SII for scleractinian corals ([Lin et al., 2014](#)).

Sliding window comparison of the two mitochondrial genomes is shown in [Fig. 1](#). The coding region with highest divergence between the two sequences is cytochrome b (CYB) with a mean divergence of 0.024 substitutions per site. All other coding regions exhibit lower divergence per nucleotide, with COI only about 1.6 percent divergent.

Evaluation of CYB sequences across the Scleractinia and other corals for their utility in separating congeneric and more distant species included 249 pairs of intraspecific contrasts, 2,098 intrageneric contrasts, and a total of 64,981 contrasts that included 203 species from 94 genera of corals. The sequence alignment of these data is provided as [File S1](#); a neighbor-joining tree generated from these data is provided as [File S2](#). These contrasts, shown in [Fig. 2](#), indicate that a divergence comparable to intraspecific diversity can be observed between members of the same genus or even more distantly related taxon pairs.

## DISCUSSION

The results of this study do less than hoped to advance molecular methods for species identification in corals. Mitochondrial DNA is often an optimal solution for metazoan species barcoding and a first attempt at species delineation ([Hebert et al., 2003](#)). Yet, in



**Figure 2** Cytochrome b divergence among scleractinian and related corals. Cytochrome b divergence, using K2P genetic distances, among taxon pairs. Plots are separated by intraspecific contrasts, intra-generic contrasts, and all other observed distances. Plot is truncated at 0.05 for clarity; all intraspecific and almost all intra-generic contrasts are shown.

corals—or anthozoans in general—the processes of mutation and DNA error correction in mitochondrial DNA (Hellberg, 2006), along with the propensity for hybridization among some taxa, has led to the need for more laborious locus development for such goals. Many studies are relying on microsatellite development (Concepcion *et al.*, 2010; Concepcion, Baums & Toonen, 2014), which enables additional variation and the benefits of a multi-locus study; however, the direct identification and analysis of shared, derived characters that distinguish populations is sometimes more complicated with such data.

Ultimately the goal of species delineation is identification of character states that are diagnostic. Finding gene regions that provide sufficient information, above and beyond the variation found within a population, is the challenge. Some nuclear gene regions have shown promise. For example, Concepcion *et al.* (2008) identified the SRP54 exon-primed intron-crossing locus as being a single-copy locus that is typically more variable than non-coding regions such as the ribosomal internal transcribed spacer (ITS) regions. Other authors are combining data from several loci to attain the same goal (McFadden, Reynolds & Janes, 2014). Certainly it is now common to approach phylogenetic and population genetic questions with multi-locus data where possible, and the same rigor appears to be necessary for species delimitation in coral taxa.

A somewhat circular problem of using available data for consideration of barcode locus efficiency is that Genbank itself is rife with poorly identified data (Meier *et al.*, 2006; Kwong, Srivathsan & Meier, 2012). When a nucleotide sequence is labeled/described only to genus, as in ‘*Discosoma* sp.’, it may or may not be conspecific with *Discosoma nummiforme*. In these analyses, such taxa would be compared at the genus, not the species level; however, in the data included in this study this would affect at most 5 taxon pairs. Additionally, there is the occasional problem of species that phenotypically appear to be similar to the focal

taxon, but genetically divergent (for example, in a phylogeography study), and are labeled with ‘*cf.*’ as indication of this uncertainty. These may or may not be the same actual species; in this study, they are treated the same as with underdescribed taxa noted above. Other taxonomic concerns among the data could similarly affect the inclusion of a contrast as within- or among-genera, but are beyond the scope of this study to resolve.

The premise of this study, that more divergent regions could be found by comparing mitochondrial genomes, is directly relevant only to the genus *Agaricia* from which these sequences derive. Using only a single genome from each species presents an incomplete picture of overall net nucleotide divergence (*Nei & Li, 1979*). However, given the typical problem of developing such markers in corals it may make sense as a general strategy to first explore available genomic data—whether mitochondrial or whole-genome—rather than blindly tackle the problem with available primer regions or use the same gene region that has proven useful in other Metazoans. Here, for example, we see that while the divergence between members of *Agaricia* is low at most protein-coding regions of the mitochondrion, the ‘control region’ exhibits high divergence; *Luck et al. (2013)* used another novel non-coding region to separate the Agariciid taxa *Leptoseris* and *Pavona*. It remains to be seen whether using next-generation approaches, as in this study, to generate whole mitochondrial genome sequences, may be more informative (but see *Fukami & Knowlton (2005)*) and nearly as cost-effective as attempting to capture several distinct gene regions via PCR.

## ACKNOWLEDGEMENTS

Sequence data were generated at the Georgia Genomics Facility ([dna.uga.edu](http://dna.uga.edu)). The specimen of *Agaricia fragilis* was collected by MK Meyers. Thanks to [WriteLatex.com](http://WriteLatex.com) for aiding my escape from the hegemony of Microsoft.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

This work was funded by NSF-EID-1015342 to John P. Wares and colleagues at the University of Georgia. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the author:  
National Science Foundation, Ecology of Infectious Diseases: #1015342.

### Competing Interests

The author declares there are no competing interests.

### Author Contributions

- John P. Wares conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, wrote the paper, prepared figures and/or tables, reviewed drafts of the paper.

## DNA Deposition

The following information was supplied regarding the deposition of DNA sequences:

Sequence data were generated at the Georgia Genomics Facility and deposited with GenBank, accession [KM051016](#).

## Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.564#supplemental-information>.

## REFERENCES

- Avise JC. 2000.** *Phylogeography, the history and formation of species*. Cambridge, MA: Harvard Univ. Press.
- Bernt M, Donath A, Juhling F, Externbrink F, Florentz C, Fritsch G, Putz J, Middendorf M, Stadler PF. 2013.** Mitos: improved de novo metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution* **69**(2):313–319 DOI [10.1016/j.ympev.2012.08.023](#).
- Bongaerts P, Frade PR, Ogier JJ, Hay KB, van Bleijswijk J, Englebert N, Vermeij MJA, Bak RPM, Visser PM, Hoegh-Guldberg O. 2013.** Sharing the slope: depth partitioning of agariciid corals and associated *Symbiodinium* across shallow and mesophotic habitats (2–60 m) on a Caribbean reef. *BMC Evolutionary Biology* **13**:205 DOI [10.1186/1471-2148-13-205](#).
- Breedy O, Guzman HM. 2014.** A new species of alcyonacean octocoral from the Peruvian zoogeographic region. *Journal of the Marine Biological Association of the United Kingdom* **94**(3):493–498 DOI [10.1017/S0025315413001835](#).
- Breedy O, Williams GC, Guzman HM. 2013.** Two new species of gorgonian octocorals from the Tropical Eastern Pacific biogeographic region (Cnidaria, Anthozoa, Gorgoniidae). *Zookeys* **350**:75–90 DOI [10.3897/zookeys.350.6117](#).
- Bruno JF, Edmunds PJ. 1997.** Clonal variation for phenotypic plasticity in the coral *Madracis mirabilis*. *Ecology* **78**(7):2177–2190 DOI [10.1890/0012-9658\(1997\)078\[2177:CVFPPI\]2.0.CO;2](#).
- Concepcion GT, Baums IB, Toonen RJ. 2014.** Regional population structure of *Montipora capitata* across the Hawaiian archipelago. *Bulletin of Marine Science* **90**(1):257–275 DOI [10.5343/bms.2012.1109](#).
- Concepcion GT, Crepeau MW, Wagner D, Kahng SE, Toonen RJ. 2008.** An alternative to ITS, a hypervariable, single-copy nuclear intron in corals, and its use in detecting cryptic species within the octocoral genus *Carijoa*. *Coral Reefs* **27**(2):323–336 DOI [10.1007/s00338-007-0323-x](#).
- Concepcion GT, Medina M, Toonen RJ. 2006.** Noncoding mitochondrial loci for corals. *Molecular Ecology Notes* **6**:1208–1211 DOI [10.1111/j.1471-8286.2006.01493.x](#).
- Concepcion GT, Polato NR, Baums IB, Toonen RJ. 2010.** Development of microsatellite markers from four Hawaiian corals: *Acropora cytherea*, *Fungia scutaria*, *Montipora capitata* and *Porites lobata*. *Conservation Genetics Resources* **2**(1):11–15 DOI [10.1007/s12686-009-9118-4](#).
- Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. 1994.** DNA primers for amplification of mitochondrial cytochrome c oxidase subunit i from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* **3**:294–299.
- Fukami H, Knowlton N. 2005.** Analysis of complete mitochondrial DNA sequences of three members of the *Montastraea annularis* coral species complex (Cnidaria, Anthozoa, Scleractinia). *Coral Reefs* **24**(3):410–417 DOI [10.1007/s00338-005-0023-3](#).

- Hebert P, Cywinska A, Ball S, deWaard J. 2003.** Biological identification through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences* **270**:313–321 DOI [10.1098/rspb.2002.2218](https://doi.org/10.1098/rspb.2002.2218).
- Hellberg ME. 2006.** No variation and low synonymous substitution rates in coral mtDNA despite high nuclear variation. *BMC Evolutionary Biology* **6**:24 DOI [10.1186/1471-2148-6-24](https://doi.org/10.1186/1471-2148-6-24).
- Hyde JR, Vetter RD. 2007.** The origin, evolution, and diversification of rockfishes of the genus *Sebastes* (Cuvier). *Molecular Phylogenetics and Evolution* **44**:790–811 DOI [10.1016/j.ympev.2006.12.026](https://doi.org/10.1016/j.ympev.2006.12.026).
- Kwong S, Srivathsan A, Meier R. 2012.** An update on DNA barcoding: low species coverage and numerous unidentified sequences. *Cladistics* **28**(6):639–644 DOI [10.1111/j.1096-0031.2012.00408.x](https://doi.org/10.1111/j.1096-0031.2012.00408.x).
- Lin MF, Kitahara MV, Luo H, Tracey D, Geller J, Fukami H, Miller DJ, Chen CA. 2014.** Mitochondrial genome rearrangements in the scleractinia/corallimorpharia complex: implications for coral phylogeny. *Genome Biology and Evolution* **6**(5):1086–1095 DOI [10.1093/gbe/evu084](https://doi.org/10.1093/gbe/evu084).
- Luck DG, Forsman ZH, Toonen RJ, Leicht SJ, Kahng SE. 2013.** Polyphyly and hidden species among Hawai'i's dominant mesophotic coral genera, *Leptoseris* and *Pavona* (Scleractinia: Agariciidae). *PeerJ* **1**:e132 DOI [10.7717/peerj.132](https://doi.org/10.7717/peerj.132).
- McFadden CS, Reynolds A, Janes MP. 2014.** DNA barcoding of xeniid soft corals (Octocorallia: Alcyonacea: Xeniidae) from Indonesia: species richness and phylogenetic relationships. *Systematics and Biodiversity* **12**:247–257 DOI [10.1080/14772000.2014.902866](https://doi.org/10.1080/14772000.2014.902866).
- Meier R, Shiyang K, Vaidya G, Ng PKL. 2006.** DNA barcoding and taxonomy in diptera: a tale of high intraspecific variability and low identification success. *Systematic Biology* **55**(5):715–728 DOI [10.1080/10635150600969864](https://doi.org/10.1080/10635150600969864).
- Meyers MK. 2013.** From molecules to communities: an examination of multiple scales of diversity in reef building corals. PhD thesis.
- Meyers M, Porter JW, Wares JP. 2013.** Genetic diversity of fluorescent proteins in Caribbean agariciid corals. *Journal of Heredity* **104**(4):572–577 DOI [10.1093/jhered/est028](https://doi.org/10.1093/jhered/est028).
- Muss A, Robertson DR, Stepien CA, Wirtz P, Bowen BW. 2001.** Phylogeography of *Ophioblennius*: the role of ocean currents and geography in reef fish evolution. *Evolution* **55**:561–572 DOI [10.1554/0014-3820\(2001\)055\[0561:POOTRO\]2.0.CO;2](https://doi.org/10.1554/0014-3820(2001)055[0561:POOTRO]2.0.CO;2).
- Nei M, Li WH. 1979.** Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the United States of America* **76**:5269–5273 DOI [10.1073/pnas.76.10.5269](https://doi.org/10.1073/pnas.76.10.5269).
- Plaisance L, Caley MJ, Brainard RE, Knowlton N. 2011.** The diversity of coral reefs: what are we missing? *PLoS ONE* **6**(10):e25026 DOI [10.1371/journal.pone.0025026](https://doi.org/10.1371/journal.pone.0025026).
- Shearer TL, Coffroth MA. 2008.** Barcoding corals: limited by interspecific divergence, not intraspecific variation. *Molecular Ecology Notes* **8**:247–255 DOI [10.1111/j.1471-8286.2007.01996.x](https://doi.org/10.1111/j.1471-8286.2007.01996.x).
- Swofford DL. 2000.** *PAUP\*. phylogenetic analysis using parsimony (\*and other methods)*. Sunderland: Sinauer Associates.
- Taylor MS, Hellberg ME. 2003.** Genetic evidence for local retention of pelagic larvae in a Caribbean reef fish. *Science* **299**:107–109 DOI [10.1126/science.1079365](https://doi.org/10.1126/science.1079365).
- Veron JEN. 2011.** Coral taxonomy and evolution. In: Dubinsky Z, Stambler N, eds. *Coral reefs: an ecosystem in transition*. Springer, 37–45 DOI [10.1007/978-94-007-0114-4\\_4](https://doi.org/10.1007/978-94-007-0114-4_4).

**Vollmer SV, Palumbi SR. 2002.** Hybridization and the evolution of reef coral diversity. *Science* 296:2023–2025 DOI [10.1126/science.1069524](https://doi.org/10.1126/science.1069524).

**Wares JP. 2013.** Mitochondrial evolution across lineages of the vampire barnacle *Notochthamalus scabrosus*. *Mitochondrial DNA* 1–3 Epub ahead of print DOI [10.3109/19401736.2013.825791](https://doi.org/10.3109/19401736.2013.825791).