Next-generation Sequencing of Whole Mitochondrial Genomes in Soft Corals

HHMI

Kevin Heath, Catherine S. McFadden
Department of Biology, Harvey Mudd College

HARVEY MUDD COLLEGE

Introduction

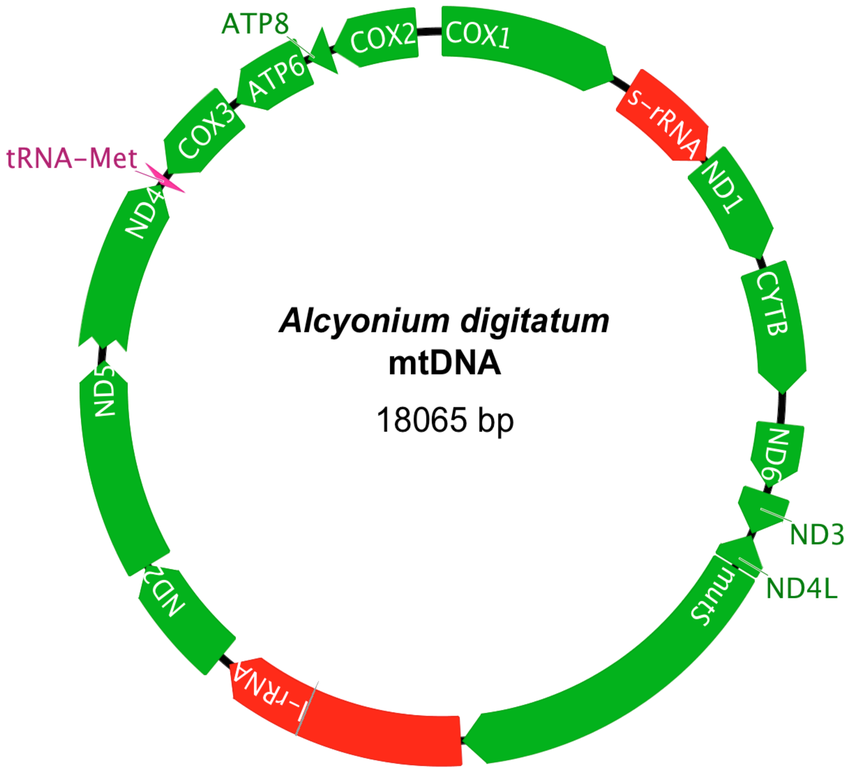
There is evidence that *A. digitatum* and *A.* sp. A, an undescribed species, interbreed to produce hybrid offspring in the field. It is difficult to confirm hybridization genetically because few loci differ between the parents. Additionally, past hybridization and speciation events confound single-locus genetic markers and make it hard to distinguish between individuals of different species. However, next-generation sequencing techniques (RADseq) make it possible to sequence numerous loci for a large number of individuals. By screening thousands of loci, we can identify SNPs that distinguish the parent species. Since hybrids should share alleles from both parents, we can used these fixed differences to genetically confirm the existence of hybrids.

Benefits of RADseq

RADseq is an effective method of sequencing large amounts of genomic data at a reasonable cost. Unlike other sequencing methods, RADseq targets DNA adjacent to restriction sites, producing a collection of short reads from many places in a genome. It is therefore more likely to detect rare or novel variations. This property of RADseq makes it useful for genotyping species. By using RADseq, we hope to identify enough unique SNPs to distinguish between the two coral species and their potential hybrid offspring.

Constructing Mitochondrial Genomes

Mitochondrial genomes are useful for phylogenetic studies of coral. Both variations in protein coding regions and gene order can be used to differentiate species and determine maternal lineages. RADSeq allows us to recover large portions of a coral's mitochondrial genome. By aligning the short reads to the *Dendronephthyia gigantea* mitochondrial genome, we were able to reconstruct most the mitochondrial genomes for *A. digitatum* and *A.* sp. A. We then used the genome to design PCR primers to target the incomplete regions for sequencing.



Gene	*A. digitatum* Size	*A.* sp. A Size	SNPs
COI	1597	1598	6
s-rRNA	919	915	9
ND1	969	969	2
CYTB	1128	1092	4
ND6	558	557	3
ND3	368	366	2
ND4L	294	294	0
MutS	2945	2941	19
l-rRNA	2193	2194	10
ND2	1027	1028	2
ND5	1818	1819	6
ND4	1482	1504	4
tRNA-Met	71	71	0
COIII	775	800	2
ATP6	708	708	21
ATP8	216	215	1
COII	762	761	2
IGR Totals	786	803	14

Comparison of *A. digitatum* and *A.* sp. A

The most variation in the coding regions of these genomes occurs in MutS and ATP6, which have 19 and 21 SNPs, respectively. Additionally, the COIII coding region in *A. digitatum* is also 25 base pairs shorter than that in *A.* sp A. The intergenic regions of the corals are very similar, with only 14 fixed differences. However, the intergenic regions of *A.* sp. A is 17 base pairs longer than in *A. digitatum*. In comparison to soft corals in the *Dendronephthya* genus, *Alcyonium* mitochondrial genomes are approximately 800 base pairs shorter. This difference is primarily the result of different-sized intergenic regions.

Acknowledgements

We thank the Howard Hughes Medical Institude for funding this research. We also thank the Toonen Lab at the Hawaiian Institute of Marine Biology for RADseq library preparation and sequencing. Finally, thanks to Jeho Park for helping us access XSEDE supercomputing resources.

References

Davey, J. W., and M. L. Blaxter. "RADSeq: Next-generation Population Genetics." Briefings in Functional Genomics 9.5-6 (2011): 416-23. Web.
Toonen R, Puritz J, Forsman Z, Whitney J, Fernandez-Silva I, Andrews K, Belcaid M, Bird C (In prep.) ezRAD: a simplified method for genomic genotyping and SNP discovery pipeline for non-model organisms.
Park, Eunji, et. al. "Estimation of divergence times in cnidarian evolution based on mitochondrial protein-coding genes and the fossil record." Molecular Phylogenetics and Evolution 62.1 (2012): 329-345.