New approaches to species delimitation and population structure of anthozoans: two case studies of octocorals using ultraconserved elements and exons

Running Title: New genomic approaches for coral delimitation

Katie L. Erickson[1#], Alicia Pentico[1#], Andrea M. Quattrini[1,2*], Catherine S. McFadden[1]

[1] Department of Biology, Harvey Mudd College, Claremont CA 91711 USA
[2] Department of Invertebrate Zoology, National Museum of Natural History, Smithsonian Institution, Washington DC 20560 USA

[#]equal contributions
*corresponding author quattrinia@si.edu

Abstract

As coral populations decline worldwide in the face of ongoing environmental change, documenting their distribution, diversity and conservation status is now more imperative than ever. Accurate delimitation and identification of species is a critical first step. This task, however, is not trivial as morphological variation and slowly evolving molecular markers confound species identification. New approaches to species delimitation in corals are needed to overcome these challenges. Here, we test whether target enrichment of ultraconserved elements (UCEs) and exons can be used for delimiting species boundaries and population structure within species of corals by focusing on two octocoral genera, *Alcyonium* and *Sinularia*, as exemplary case studies. We designed an updated bait set (29,181 baits) to target-capture 3,023 UCE and exon loci, recovering a mean of $1,910 \pm 168$ SD per sample with a mean length of $1,055 \pm 208$ bp. Similar numbers of loci were recovered from *Sinularia* ($1,946 \pm 227$ SD) and *Alcyonium* ($1,863 \pm 177$ SD). Species-level phylogenies were highly supported for both genera. Clustering methods based on filtered SNPs delimited species and populations that are congruent with previous allozyme, DNA barcoding, reproductive and ecological data for *Alcyonium,* and offered further evidence of hybridization among species. For *Sinularia,* results were congruent with those obtained from a previous study using Restriction Site Associated DNA Sequencing. Both

case studies demonstrate the utility of target-enrichment of UCEs and exons to address a wide range of evolutionary and taxonomic questions across deep to shallow time scales in corals.

Keywords: cnidaria, population genomics, target enrichment, UCE, Octocorallia

Introduction

Corals are arguably one of the most ecologically important groups of metazoans on earth. By their ability to produce large colonies by asexual propagation and CaCO$_3$ precipitation, they create massive biogenic structures, thereby engineering entire reef-based ecosystems in both shallow and deep waters (Riegl & Dodge, 2019; Roberts, Wheeler, & Freiwald, 2006). These ecological communities are among those most imperiled by climate change, with dire predictions suggesting that they may not survive the next century (Carpenter et al., 2008; Pandolfi, Connolly, Marshall, & Cohen, 2011; Hughes, Bellwood, Connolly, Cornell, & Karlson, 2014). As coral populations and species are in decline, it is important to establish baselines of abundance, distribution, and diversity, all of which rely on accurately delimiting and identifying species. Unfortunately, our current lack of understanding of basic taxonomy and species boundaries in many groups of corals prohibits us from better understanding their roles in ecosystems, and ultimately, from effectively managing their habitats.

Traditionally, morphological traits such as colony growth form and skeletal structure have been used to distinguish corals at all taxonomic levels (Fabricius & Alderslade, 2001; Veron, 2000). At the species level, however, morphological differences can be subtle, and both phenotypic variation and plasticity of growth form often confound assessment of species boundaries (e.g., Forsman, Barshis, Hunter, & Toonen, 2009; Marti-Puig et al., 2014; McFadden et al., 2017; Paz-García, Hellberg, García-de-León, & Balart, 2015). The advent of integrated taxonomic approaches that utilize molecular characters to inform taxonomy has greatly improved our understanding of species boundaries in numerous coral taxa, and led to the recognition of homoplasy in many morphological characters previously considered to be diagnostic of species and higher taxa (Arrigoni et al., 2014; Benayahu, Ofwegen, & McFadden, 2018; Budd & Stolarski, 2009; McFadden et al., 2017). The relative invariance of mitochondrial genes in corals and most other anthozoans (e.g., sea anemones) has, however, hindered the application of integrated approaches to species delimitation in this ecologically important group (Huang, Meier, Todd, & Chou, 2008; McFadden et al. 2011).

As a result of their slow rates of mitochondrial gene evolution coupled with a lack of alternative marker development, both DNA barcoding efforts and studies of phylogeography in most corals lag far behind many other groups. The universal COI barcode does not reliably distinguish species within anthozoans (Huang et al., 2008; McFadden et al., 2011; Shearer & Coffroth, 2008). No alternative, universally agreed upon single-locus barcode has been proposed or widely used for either the scleractinian corals or related sea anemones, and a multilocus barcode that has been used for the octocorals (*mtMutS* + COI + 28S rDNA) has recognized

limitations (McFadden et al., 2011; McFadden, Brown, Brayton, Hunt, & Ofwegen, 2014; Quattrini et al., 2019). Relatively few studies have developed alternative markers such as nuclear gene regions (Ladner & Palumbi, 2012; Prada et al. 2014) or microsatellites (e.g., Baums, Boulay, Polato, & Hellberg, 2012; Gutiérrez-Rodríguez & Lasker, 2004; LeGoff, Pybus, & Rogers, 2004; Quattrini, Baums, Shank, Morrison, & Cordes, 2015) for use in species delimitation or phylogeographic studies. As a result, understanding of population structure, phylogeographic patterns, and the processes driving evolutionary diversification in corals and other anthozoans lags far behind the progress that has been made in other ecologically important marine phyla (Cowman, Parravicini, Kulbicki, & Floeters, 2017; Crandall et al., 2019; Hodge & Bellwood, 2016).

The recent development of genomic sequencing methods holds considerable promise for resolution of species boundaries and phylogeography in corals and other anthozoans. Several studies have now demonstrated that Restriction Site Associated DNA sequencing (RAD-Seq) effectively delimits species when conventionally-used barcodes have failed (Forsman et al., 2017; Herrera & Shank, 2016; Johnston et al., 2017; Pante et al., 2015; Quattrini et al., 2019). However, few studies have yet documented the effectiveness of RAD-Seq for population-level work in anthozoans (Leydet, Grupstra, Coma, Ribes, & Hellberg, 2018; Bracco, Liu, Galaska, Quattrini, & Herrera, 2019). There are also a few inherent disadvantages to RAD-Seq relative to sequence enrichment methods. These include a stark drop in orthologous loci that can be extracted from distantly related taxa (Cairou, Duret, & Charlat, 2013; Viricel, Pante, Dabin, & Simon-Bouhet, 2013), limiting the ability of RAD-Seq to resolve deep phylogenetic relationships (Harvey, Smith, Glenn, Faircloth, & Brumfield, 2016; but see Eaton, Spriggs, Park, & Donahue, 2017), and rendering this method impractical for combining datasets of taxa that first diverged hundreds of millions of years ago (Stolarski et al. 2011; Quattrini et al., accepted).

Target capture enrichment of Ultraconserved Elements (UCEs) and exons could be a welcome addition to the methods used for species delimitation and studies of phylogeography and population structure in corals and other anthozoans. UCEs consist of highly conserved regions shared across species, which makes them ideal loci for target-capture enrichment because a single RNA probe set can be applied to a large sample of highly divergent species (Faircloth et al., 2012; McCormack et al., 2012; Quattrini et al., 2018). Recently, Quattrini et al. (2018; accepted) demonstrated that UCEs and exons can resolve deeper-level phylogenetic relationships within class Anthozoa, and Quek et al. (2020) applied a similar transcriptome-based target-enrichment approach to resolve genera of scleractinian corals. However, since highly divergent and phylogenetically informative regions flank both UCE and exon loci (Faircloth et al., 2012; Wolf et al., 2018), they have the potential to resolve phylogenetic relationships at shallow evolutionary scales as well. Previous studies have successfully demonstrated that UCE and/or exon data can be used for species delimitation (e.g., Blair et al., 2019; Hedin et al., 2018; Zarza et al., 2018) and population genomics (Oswald et al., 2019; Winker, Glenn, & Faircloth, 2018; Zarza et al., 2016) in a wide range of taxa (e.g., arthropods, frogs, snakes, birds) across a range of evolutionary time scales. Moreover, the target-capture method has proven successful

with highly degraded DNA, such as is often recovered from museum specimens or specimens not preserved specifically for genomic work (Derkarabetian, Benavides, & Giribet, 2019; Hawkins et al., 2016; McCormack et al., 2016; Ruane & Austin, 2017). Here, we test the ability of target-capture of UCEs and exons to delimit species and geographically discrete populations of corals, comparing the results directly to previous studies that have applied RAD-Seq, single-locus barcoding and other molecular methods to the same specimens. If this method is in fact successful at delimiting species and populations, then a single comprehensive dataset could be used to simultaneously estimate population structure within a species, discriminate closely related species, and identify phylogenetic relationships across a variety of timescales to help resolve the anthozoan tree of life.

To investigate the utility of UCEs and exons in delimiting coral species and populations, we present two case studies, both of which have been the focus of previous species delimitation efforts using other molecular methods. The first case study is the genus *Alcyonium*, a common soft coral found circumglobally in temperate and cold waters. *Alcyonium* provides an ideal study system because species boundaries among a group of 10 North Atlantic and Mediterranean species have been confirmed previously using a combination of morphological, reproductive and genetic data (McFadden, 1999; McFadden, Donahue, Hadland, & Weston, 2001; McFadden & Hutchinson, 2004). These 10 species (two currently undescribed) belong to three well-defined clades (the *digitatum*, *coralloides* and *acaule* groups). Pairs of species in each clade nonetheless share identical or nearly identical haplotypes or genotypes at *mtMutS, COI, 28S rDNA* and *ITS*, the markers that have been used most widely for species delimitation and DNA barcoding in octocorals (McFadden et al., 2001, 2011, 2014; McFadden, Sánchez, & France, 2010). Differences in their reproductive biology (McFadden et al., 2001) as well as fixed allelic differences at allozyme loci (McFadden, 1999) suggest that genetically very similar species are nonetheless reproductively isolated. In addition, the population structure of *A. coralloides* spanning the species' range in the Mediterranean and eastern North Atlantic has been documented previously using allozymes (McFadden, 1999). Therefore, we can evaluate the utility of using UCEs and exons for both population genetic and species delimitation studies of corals using *Alcyonium* as an exemplar.

Our second case study is the genus *Sinularia*, one of the most speciose, widespread and ecologically important soft corals found on shallow-water reefs throughout the Indo-Pacific. As many as 38 morphospecies of *Sinularia* have been found to co-occur on a single reef (Manuputty & Ofwegen, 2007; Ofwegen, 2002), where they contribute to reef formation and biodiversity maintenance. Despite the important roles that they play in coral reef ecosystems, delimitation of closely related *Sinularia* species based on morphological characters and DNA barcodes remains challenging. Two clades in particular, denoted as clades 4 and 5C, each comprise many nominal species whose morphological distinctions are subtle and often confusing (McFadden, Ofwegen, Beckman, Benayahu, & Alderslade, 2009), and DNA barcodes do not always reliably differentiate distinct morphospecies (McFadden et al., 2014; Benayahu et al., 2018). Recently, Quattrini et al. (2019) combined RAD-Seq data with coalescent and clustering methods to

delimit *Sinularia* species, also using the results to evaluate the reliability of character-based DNA barcodes in the group. This work has given us the opportunity to use the same individuals as in Quattrini et al. (2019) to determine if species delimitations based on RAD-Seq vs. UCE/exon loci are congruent.

By targeting the same individuals used in these previous studies, we address whether species and population boundaries delimited with target-capture of UCEs and exons are congruent with findings from other methods. Specifically, we re-designed our original anthozoa-v1 bait set (Quattrini et al., 2018) to capture more loci with a higher capture rate in both subclasses of anthozoans, including octocorals (herein) and hexacorals (Cowman et al., 2020). In this study, we then target-captured these loci within two exemplar genera of alcyonacean soft corals and examined phylogenetic and phylogeographic relationships. In addition, we used SNPs from UCE and exon loci in Bayesian and multivariate analyses to delimit species within *Alcyonium* and *Sinularia* and to infer population structure within *A. coralloides*.

Methods

Case Study 1: Alcyonium

Alcyonium specimens were collected between 1989-1996 from locations in the North Atlantic and Mediterranean using SCUBA, with the exception of *A. palmatum* which was obtained from deep water by dredge (McFadden, 1999; McFadden et al., 2001). All samples were preserved in liquid nitrogen immediately after collection and stored in a -80°C freezer. Sixty-eight *Alcyonium* samples that had been included in previous studies were sequenced herein. Samples included at least three specimens from each of the 10 putative species in three distinct clades (*digitatum, coralloides, acaule*) as denoted by *ITS, mtMutS* and/or *COI* barcodes (McFadden et al., 2001, 2011). In addition, a total of 32 specimens of *A. coralloides* that were analyzed with allozymes (McFadden, 1999) were included, including five individuals from each of four locations: two in the Mediterranean, from Marseille, France (*A. coralloides*-M1) and Banyuls-sur-Mer, France (*A. coralloides*-M2); and two in the eastern North Atlantic, from Iles Chausey, France (*A. coralloides*-CH) and Sagres, Portugal (*A. coralloides*-SG). *Alcyonium haddoni,* collected from Argentina, was also included as it had been sequenced in Quattrini et al. (2018).

Case Study 2: Sinularia

Sinularia from clades 4 and 5C were collected using SCUBA at 13 sites in the shallow fore-reef zone (3-21m) at Dongsha Atoll Marine National Park (Taiwan) in 2011, 2013, and 2015 (Benayahu et al., 2018; Quattrini et al., 2019). All specimens were preserved in 70% EtOH, with small tissue subsamples preserved in 95% EtOH for DNA (Quattrini et al., 2019). Benayahu et al. (2018) identified these individuals to species using morphological characters and DNA barcodes. Subsequently, Quattrini et al. (2019) used RAD-Seq from 95 individuals to confirm species boundaries among morphospecies. A subset of 13 individuals of four morphospecies in

clade 4 and 31 individuals of nine morphospecies in clade 5C were selected for sequencing using target-capture methods.

RNA Bait Design

To have higher specificity for and target more loci in specific taxonomic groups, we re-designed a bait set (anthozoa-v1, Quattrini et al., 2018) that was produced to target-capture 720 UCE and 1,071 exon loci from across the subclass Anthozoa. Target-capture results with the original bait set from 234 anthozoans (Quattrini et al., accepted) were screened to keep baits that either performed well across all anthozoans or within particular groups more specifically. Thus, we retained 12,320 baits targeting 1,580 loci (667 UCE and 913 exon loci) from the anthozoa-v1 bait set. Using the program Phyluce (Faircloth, 2016) and following methods in Quattrini et al., (2018), we added more octocoral-specific baits to 1) improve target-capture performance of the anthozoa-v1 loci in octocorals and 2) target additional loci, which were not included in the anthozoa-v1 bait set. A similar procedure was followed for hexacorals, as reported elsewhere (Cowman et al., 2020). Methods are outlined in Supplementary File 1, but more details can also be found in both Quattrini et al. (2018) and the PHYLUCE documentation (https://phyluce.readthedocs.io/en/latest/tutorial-four.html, Faircloth, 2016).

Following bait design, the exon and UCE bait sets were concatenated and then screened to remove duplicate baits ($\geq$50% identical over >50% of their length), resulting in a final non-duplicated octocoral-v2 bait set of 29,181 baits targeting 3,023 loci (1,337 UCE and 1,686 exon loci). Baits designed against transcriptomes have "design: octotrans" in the .fasta headers. It is possible that both "UCE" loci and "exon" loci contain both conserved coding and non-coding regions, but we retain the distinction here based on how the baits were designed (e.g., exons designed from transcriptomes, UCEs from genomes) and following naming conventions in prior studies (Branstetter, Longino, Ward, & Faircloth, 2017; Starrett et al., 2017; Quattrini et al., 2018). Baits were synthesized by Arbor BioSciences (Ann Arbor, MI).

DNA extractions, library preps, and sequencing

DNA was extracted using the Qiagen DNeasy Blood & Tissue kit or a modified CTAB extraction protocol (McFadden et al., 2001). Approximately 250 ng of DNA from each sample was then sheared to an average size of 400-800 bp (Q800R QSonica Inc. Sonicator). Libraries were prepared for sequencing using the KAPA HyperPrepKit (Kapa Biosystems) with dual-indexed iTru adapters following Quattrini et al. (2018). For 66 *Alcyonium* samples and 38 *Sinularia* samples, the MyBaits protocol v IV (Arbor BioSciences, Ann Arbor, MI) was used to target and enrich UCEs and exons in pools of eight individuals using the newly designed bait set. For three *Alcyonium* and six *Sinularia*, loci were captured with the anthozoa-v1 bait set as described in Quattrini et al. (2108). Library size was assessed using a Bioanalyzer (UC Riverside) and then sent to Oklahoma Medical Research Foundation (Oklahoma City, OK) for sequencing. Samples target-enriched with the anthozoa-v1 bait set were sequenced with other specimens on an Illumina HiSeq 3000 (120 samples multiplexed) and samples target-enriched

with the newly designed octocoral-v2 bait set were sequenced on a NovaSeq (214 samples multiplexed). More methodological details can be found in Quattrini et al. (2018).

Bioinformatic processing, assembly, and alignment of UCEs.

The PHYLUCE pipeline (Faircloth, 2016) was used to perform quality control on the demultiplexed reads, assemble loci, and align loci for phylogenetic analyses. Reads were cleaned using illumiprocessor (Faircloth, 2013) and Trimmomatic v 0.35 (Bolger, Lohse, & Usadel, 2014). Contigs were assembled using the program SPAdes v 3.1 (Bankevich et al., 2012; with the --careful and --cov-cutoff 2 parameters). The *phyluce_assembly_match_contigs_to_probes* command was used to match probes to contigs to identify loci with a minimum coverage of 70% and a minimum identity of 70%. Loci were then extracted using *phyluce_assembly_get_fastas_from_match_counts* and aligned with MAFFT v7.130b (Katoh et al., 2002)*. Edges of the aligned loci were then trimmed using *phyluce_align_seqcap_align*. Coverage of each locus was estimated using *phyluce_assembly_get_trinity_coverage* and *phyluce_assembly_get_trinity_coverage_for_uce_loci*. Finally, a data matrix including all loci concatenated with at least 75% of taxa present was generated using the program *phyluce_align_get_only_loci_with_min_taxa*.

Phylogenetics

Maximum likelihood inference was run on concatenated alignments of the 75% complete UCE and exon data matrix to obtain a phylogenetic tree. Phylogenies were constructed using the GTRGAMMA model by searching for the maximum-likelihood (ML) tree. The program RAxML was used to perform 100 bootstrap searches and 20 ML searches using rapid bootstrapping (Stamatakis, 2014). A phylogeny was constructed for the *Alcyonium* samples along with three *Sinularia* spp. as outgroups. RAD-Seq data from *Sinularia* as described in Quattrini et al. (2019) were subset to include only the individuals used in target-capture analysis. ML was performed on both RAD-Seq and UCE/exon datasets using the above methods on clades 4 and 5C to determine the congruence between constructed phylogenies. Clade 4 was rooted to *S. humilis* while a midpoint root was used for Clade 5C as in Quattrini et al. (2019).

SNP calling

SNPs were extracted from both genera following methods outlined in previous studies (Zarza et al. 2016, 2018; Derkarabetian et al. 2019) using modified scripts from Zarza et al. 2016. In each taxon set, the individual with the largest number of UCE/exon contigs as identified by *phyluce_assembly_get_match_counts* was chosen as a reference individual for SNP calling. The programs *phyluce_assembly_get_match_counts* and *phyluce_assembly_get_fastas_from_match_count* were re-run on the reference individual to create a fasta of UCE and exon contigs found only in the reference. Reference loci were indexed using bwa-version 0.7.7 (Li & Derbin, 2009). BAM files for each individual in the analysis were created by mapping their cleaned, trimmed reads to the reference using BWA-MEM (Li, 2013),

sorting the reads using SAMtools and removing duplicates using Picard v 1.106-0 (http://picard.sourceforge.net). GATK v3.4 (McKenna et al., 2010) was then used to realign BAMs around indels, call variants, and filter variants based on vcf tools (Danecek et al., 2011). For each clade, SNPs were called at loci for which at least 75% of taxa were represented and for which each locus had at least 10X coverage. Only individuals target-captured with the new octocoral-v2 bait set were included in SNP calling analyses to maximize number of SNPs called. SNPs were filtered to one SNP per locus for loci $\leq$ 1,000 bp, or to one per 1,000 bp if loci were $\geq$ 2,000 bp. All code for SNP calling can be found on figshare.

Species delimitation

Two methods were used to delimit species: Discriminant Analysis of Principal Components (DAPC) and STRUCTURE. Both of these programs used the filtered SNPs to identify clusters of genetically-related individuals for each clade. DAPC (Jombart et al., 2010) clusters species based on their genetic similarity and can provide estimates for how many populations (K) to assume. DAPC was conducted using the adegenet package (Jombart, 2008) in R v 3.6 (R Core Team, 2019). First the program *find.clusters* found the optimal number of clusters needed to minimize the Bayesian Information Criterion (BIC), and the function *optim.a.score* was used to determine how many Principal Component (PC) axes and Discriminant Functions (DFs) needed to be retained. The DAPC results were displayed as scatterplots.

STRUCTURE (Pritchard, Stephens, & Donnelly, 2000) uses a Bayesian clustering approach to probabilistically infer population structure given K. STRUCTURE was run in parallel using StrAutoParallel (Chhatre & Emerson, 2017, generations=1M, burnin=250K). The maximum value for K ($K = 4$–10 depending on clade analyzed) was chosen to be higher than the optimal number of clusters designated by DAPC. Five runs were completed at each value of K. The five runs at the DAPC-selected K value were combined and results were visualized using *starmie* (https://github.com/sa-lee/starmie) in R (R Core Team, 2019).

Results

Sequencing and Locus Summary

The total number of reads obtained from the HiSeq ranged from 2,680,770 to 9,150,274 and the NovaSeq from 696 to 26,147,550 total reads per sample. We removed two *Alcyonium* samples that failed sequencing, and then quality and adapter trimmed raw reads for the remaining samples, resulting in a mean of 8,021,057 ± 83,566 SD PE trimmed reads per sample (Supplementary Table 1). Trimmed reads were assembled into a mean of 83,566 ± 57,749 SD contigs per sample (range: 22,813 to 285,695) with a mean length of 400 ± 32 bp (Supplementary Table 1). Read coverage per contig ranged from 0.4 to 5.5X with the original bait set and HiSeq sequencing and 0.1 to 9.6X with the new bait set and NovaSeq sequencing.

With the new bait set and NovaSeq sequencing, a total of 2,977 loci (out of 3,023 targeted loci) were recovered from the assembled contigs (Supplementary Table 1). The mean

number of loci recovered with this new bait set was $1,910 \pm 168$ SD per sample (range: 1,019 to 2,173) with a mean length of $1,055 \pm 208$ bp (range:115 to 1,803 bp). This method produced an average of 2X more loci that were 200 bp longer as compared to target capture with the anthozoa-v1 bait set (994 ± 194 SD loci recovered at 886 ± 191 bp) and sequencing with the HiSeq. We recovered similar numbers of loci from *Sinularia* ($1,991 \pm 114$ SD) as compared with *Alcyonium* ($1,863 \pm 177$ SD). Read coverage per UCE contig ranged from 1 to 12.8X with the original bait set and HiSeq and 1.2 to 135X with the new bait set and NovaSeq.

Case Study 1: Alcyonium

The 75% taxon occupancy matrix included a concatenated alignment of 1,264 UCE and exon loci with an alignment length of 1,280,803 bp. ML inference on this dataset produced a highly supported phylogeny (most nodes >95% b.s.), except for the sister relationship between the *A. digitatum* clade and *A. haddoni* (15% b.s. support) (Fig. 1). In the *digitatum* clade, each putative species (*A.* sp. A, *A. digitatum, A. siderium)* was monophyletic. In the *coralloides* clade, most of the putative species were monophyletic, except that one *A.* sp. M2 was sister to the remaining species in the *coralloides* clade. In the *acaule* clade, *A. acaule* was monophyletic but *A. glomeratum and A. palmatum* were nested together.

Both DAPC and STRUCTURE analyses based on filtered SNPs supported the phylogenetic results for each group (Fig. 2). For the *digitatum* clade, DAPC (1,974 filtered SNPs) indicated K=3 clusters matching the three putative species. STRUCTURE (K=3) further indicated that two *A.* sp. A individuals were admixed with *A. digitatum* (Fig. 2A-B). For the *coralloides* clade, DAPC (1,980 filtered SNPs) suggested K=4 clusters, with Mediterranean *A. coralloides,* north Atlantic *A. coralloides, A. bocagei,* and *A. hibernicum/*sp. M2 forming distinct clusters (Fig. 2C-D). STRUCTURE at K=4 also revealed that *A. hibernicum* was admixed between north Atlantic *A. coralloides* and *A.* sp. M2. For the *acaule* clade, DAPC (2,114 filtered SNPs) indicated K=3 clusters corresponding to the three putative species, and STRUCTURE at K=3 indicated that all *A. glomeratum* individuals were admixed with *A. palmatum* (Fig. 2E-F).

We also examined whether this method could be used to discern populations of *A. coralloides* in the Mediterranean Sea from those in the eastern North Atlantic (Fig. 3A). From an initial 4,775 unfiltered SNPs recovered from individuals used in analyses (Table 1), 152 filtered SNPs remained post-filtering. Nevertheless, the two populations in the Atlantic (*CH, SG*) were distinct from one another and were clearly separated from those in the Mediterranean (*M1, M2*). In addition, the Mediterranean populations were differentiated from one another, although a few individuals were admixed (Fig. 3B, C).

Case Study 2: Sinularia

For *Sinularia* clade 4, the 75% taxon occupancy matrices included concatenated alignments of 2,231,641 bp (1,939 loci) and 549,862 bp (6,343 loci) for target-capture and RAD-Seq, respectively. The clade 4 tree topologies for the target-capture and RAD-Seq datasets were identical, with highly supported (> 95% b.s.) monophyletic clades for five putative species (Fig.

4). The concatenated alignments for *Sinularia* clade 5C were 1,549,273 bp (1,240 loci) and 701,433 bp (8,060 loci) for target-capture and RAD-Seq datasets, respectively. Clade 5C tree topologies for both datasets were also highly supported with monophyletic clades for seven putative species; however, *S. penghuensis* and *S. slieringsi* formed a clade and *S. acuta* and *S. abrupta* were nested together (Fig. 5). There was also a notable difference between topologies: the clade containing *S. acuta* and *S. abrupta* was sister to *S. slieringsi/penghuensis* in the RAD-Seq phylogeny, but sister to the other five species in the phylogeny produced with target-capture data. In addition, the relationships among individuals within the *S. lochmodes* and *S. slieringsi/penghuensis* clades differed between datasets, and were less well-supported in target-capture phylogenies.

We also extracted SNPs from the UCE and exon loci for both *Sinularia* clades. For clade 4, 2,030 filtered SNPs were used in the DAPC and STRUCTURE analyses. DAPC indicated that *Sinularia* clade 4 contains K=5 clusters: *S. ceramensis, S. humilis, S. pavida, S. tumulosa A,* and *S. tumulosa B* (Fig. 6A). However, one *S. tumulosa* A individual grouped with *S. tumulosa* B and one *S. ceramensis* individual grouped with *S. humilis*. Plotting K=5 clusters in the STRUCTURE analysis revealed four species that were well separated: *S. humilis, S. pavida, S. tumulosa* A, and *S. tumulosa* B, except for the one *S. tumulosa* A that was admixed with *S. tumulosa* B (Fig. 6B). Although the DAPC plot separated most *S. ceramensis* from the remaining species, the STRUCTURE plot suggested all *S. ceramensis* individuals were admixed with *S. humilis.*

The DAPC analysis indicated K=6 clusters for *Sinularia* Clade 5C based on 2,742 SNPs (Fig. 6C). Five of the six clusters corresponded to the prior group designations from RAD-Seq analyses (Quattrini et al., 2019): *S. slieringsi/penghuensis, S. penghuensis, S. densa, S. abrupta/acuta,* and *S. lochmodes*. The sixth group consisted of three putative species: *S. exilis, S. maxima,* and *S. wanannensis.* The STRUCTURE plot (K = 6) also indicated clear separation of the six groups, with some admixture observed in *S. slieringsi/penghuensis* individuals with *S. penghuensis* (Fig. 6D). We also plotted clusters at K=8 in both DAPC and STRUCTURE analyses based on results of RAD-Seq data (Quattrini et al., 2019). The DAPC plot separated each putative species, but the STRUCTURE analysis indicated that *S. wanannensis* and *exilis* clustered together, but were distinct from *S. maxima* (Fig. 6E,F).

Discussion

In the present study, we designed a bait set to improve target-capture performance of the anthozoa-v1 bait set (Quattrini et al., 2019) and target additional UCE and exon loci in octocorals (octocoral-v2, 29,181 baits targeting 3,023 loci). Testing this bait set in two well-studied octocoral genera demonstrated the utility of target-capture genomics to delimit both species and populations of octocorals. To date, only three studies have used genomic approaches to delimit among closely-related octocorals, and all three used RAD-Seq data (Herrera & Shank, 2016; Pante et al., 2015; Quattrini et al., 2019), which precludes combining or including other taxonomic groups in those datasets. Here, we demonstrated that phylogenies constructed with RAD-Seq or target-capture data were congruent and contained similar levels of resolution. This

target-capture approach was also able to confirm species boundaries, known from reproductive and allozyme data, that were only weakly supported by the widely used DNA barcodes for octocorals (e.g., McFadden et al., 2010). Our enhanced octocoral bait set combined with the companion hexacoral bait set (Cowman et al., 2020) are now available to help resolve the anthozoan tree of life across deep to shallow evolutionary time scales.

Case Study 1: Alcyonium

        Alcyonium was one of the first coral genera for which molecular evidence was used to delimit species (McFadden, 1999; McFadden et al., 2001), and its northern hemisphere representatives have continued to serve as test cases for the development of markers for DNA barcoding (McFadden et al., 2011, 2014). Previous studies have used allozymes and various single-locus barcoding markers to: (1) identify two new, cryptic species within the genus (McFadden, 1999; McFadden et al. 2001); (2) support species boundaries among eight of the 10 putative species belonging to three distinct clades (McFadden, 1999; McFadden et al. 2001, 2011, 2014); and (3) suggest that hybridization may have occurred in the past or be ongoing within two of those clades (McFadden & Hutchinson, 2004; McFadden, Rettig, & Beckman, 2005). The results of species delimitation using evidence from sequence capture of UCEs and exons provide strong support for each of these hypotheses. Moreover, they strengthen and corroborate formerly weak evidence suggesting: (4) a species boundary between north Atlantic and Mediterranean populations of *A. coralloides* (McFadden, 1999), and (5) no genetic differentiation between *A. palmatum* and *A. glomeratum* (McFadden et al. 2001, 2011, 2014).

        The four species comprising the *A. coralloides* clade formerly were synonymized under that name (Groot & Weinberg, 1982), but fixed allelic differences at four to six allozyme loci between sympatrically occurring morphotypes support the distinction of *A. hibernicum*, *A. bocagei* and *A.* sp. M2 (an undescribed species from the Mediterranean) from *A. coralloides sensu stricto* (McFadden, 1999). Molecular evidence for a species boundary between distinctly different north Atlantic and Mediterranean morphotypes of *A. coralloides*, has, however, been equivocal (McFadden, 1999; McFadden et al., 2011). Our phylogenetic results indicate monophyly for each of these species, with the exception of one individual of *A.* sp. M2 that was recovered sister to the rest of the *coralloides* clade. STRUCTURE and DAPC analyses (Figs. 2C, D) also support species boundaries between *A.* sp. M2, *A. bocagei,* and the north Atlantic and Mediterranean morphotypes of *A. coralloides.* However, DAPC analysis grouped *A. hibernicum* with *A.* sp. M2, and STRUCTURE results indicated that all (n=4) *A. hibernicum* individuals were admixed between *A.* sp. M2 and *A. coralloides*. Based on observations of shared *ITS* polymorphisms, McFadden & Hutchinson (2004) proposed that *A. hibernicum*, a putatively asexual species (Hartnoll, 1977), arose as the product of hybridization between *A. coralloides* and *A.* sp. M2. They could not, however, conclusively rule out the possibility that these three species might share *ITS* sequences due to recent ancestry and incomplete lineage sorting. STRUCTURE analysis (Fig. 2D) demonstrated that *A. hibernicum* was equally admixed between *A.* sp. M2 and the Atlantic morphotype of *A. coralloides*, supporting the hypothesis that *A.*

hibernicum is a hybrid species that arose from a cross between those two species; shared mitochondrial gene sequences suggest that *A.* sp. M2 was likely the maternal parent (McFadden et al. 2011). Because the present-day distributions of the two parental species are allopatric, this event likely happened at some time in the past, perhaps prior to *A.* sp. M2 (re)colonizing the Mediterranean subsequent to the Messinian crisis (Duggen et al., 2003).

Whereas the hybridization event that gave rise to the hybrid species, *A. hibernicum*, likely occurred in the past, it has also been proposed that hybridization may be ongoing within the *digitatum* clade, between *A. digitatum* and *A.* sp. A, an undescribed species from the British Isles (McFadden et al., 2001; McFadden et al., 2005). Where these two sympatric species co-occur in close proximity, individuals that are intermediate in morphology are occasionally found. Genetic evidence from a single allozyme locus and from shared polymorphisms in *ITS* are consistent with the conclusion that these intermediate individuals are F1 hybrids (McFadden et al., 2005), but the extreme genetic similarity of the two parent species at all other loci that have been examined to date made it difficult to establish that fact with certainty. Although only two putative hybrid individuals with intermediate morphologies were sequenced in this study, STRUCTURE analyses (Fig. 2B) indicated that both of those individuals appear to be admixed equally with *A. digitatum* and *A.* sp. A. Sequencing of additional hybrid and parental individuals will be necessary to confirm the status of the hybrids as F1s, and to support the hypothesis that they propagate asexually (McFadden et al., 2005).

Within the *acaule* clade, *A. palmatum* and *A. glomeratum* are distinct from one another morphologically, geographically and ecologically, yet share identical haplotypes or genotypes at *ITS, 28S rDNA, mtMutS* and *COI* (McFadden et al., 2001, 2011, 2014). *A. palmatum* is restricted to deep-water soft sediment habitats in the Mediterranean and north Atlantic, while *A. glomeratum* occurs on rocky substrates in shallow water in the north Atlantic only (Manuel, 1981). In addition to these ecological differences, *A. palmatum* and *A. glomeratum* also differ in colony growth form, with *A. glomeratum* more closely resembling *A. acaule,* an ecologically similar species found on hard substrate in the Mediterranean (Weinberg, 1977). As in previous studies (McFadden et al. 2001, 2011), our phylogenetic and species delimitation analyses indicated that *A. acaule* was distinct from *A. glomeratum* and *A. palmatum,* but the evidence that *A. glomeratum* and *A. palmatum* are distinct species is less clear. Although they formed slightly separated clusters in DAPC analysis (Fig. 2E), neither species was monophyletic, and STRUCTURE analysis (Fig. 2F) indicated admixture. Recent sequence data from an Atlantic population of *A. palmatum* support the hypothesis of ongoing gene flow between these two species (unpubl. data), but more extensive sampling will be required to confirm that possibility.

The target-capture approach also enabled us to differentiate among populations of *A. coralloides,* even with relatively small numbers of SNPs (n = 152)*. The STRUCTURE and DAPC analyses (Fig. 3) revealed differentiation between the two north Atlantic populations (~1500 km apart) and between the two Mediterranean populations (~200 km apart); only a few individuals in the latter populations were admixed. Furthermore, analyses clearly differentiated the Mediterranean and Atlantic populations of *A. coralloides* and indicated no admixture,

suggesting that they may indeed be different species. Results correspond to previous analyses using allozymes (McFadden, 1999), and thus show potential in helping to discern populations in other species of corals.

Case Study 2: Sinularia

Phylogenies of *Sinularia* clade 4 and 5C constructed with target-capture UCE/exon data were congruent with those based on RAD-Seq data (Fig. 4-5 in Quattrini et al., 2019). Both analyses used 75% taxon-occupancy matrices, and recovered highly resolved trees, with the majority of nodes having 95-100% bootstrap support. Only a few nodes in each analysis had <95% b.s. support, although there was a trend for more highly supported nodes in the phylogeny constructed with UCE/exon data. Overall, the target-capture data consisted of fewer loci (~2K loci) compared to RAD-Seq data (~6-8k loci), but average locus size (1K bp) and alignments were longer (~1-2M bp) in UCE/exon data as compared to RAD-Seq data (~87 bp locus size, 500-700K bp alignment size). One observed difference between the clade 5C phylogenies was the placement of the *S. acuta/abrupta* clade, with this clade sister to *S. penghuensis/slieringsi* in the RAD-Seq phylogeny and sister to the remaining species in the UCE/exon phylogeny. Quattrini et al. (2019) also found incongruence in the placement of this clade depending on method used (e.g., SNAPP using SNPs or ML using concatenated data) and also indicated that some members of this clade were hybrids. Regardless of the unresolved placement of *S. acuta/abrupta*, topologies based on UCEs/exons were highly congruent with those produced from RAD-Seq data, further supporting that the method (RAD-Seq vs. UCE target capture) does not influence phylogenetic results (Harvey et al., 2016; Manthey et al., 2016).

Our results also suggest that the effectiveness of target enriching UCEs/exons for species delimitation is similar to the more frequently used RAD-Seq approach. DAPC and STRUCTURE results (Fig. 6) were generally congruent with those from RAD-Seq (Quattrini et al. (2019). In our study, however, we noted that one *S. tumulosa A* (D324) grouped with *S. tumulosa B* in DAPC and it was clearly admixed in STRUCTURE analysis (Fig. 6B). This specimen was placed in sp. A in prior RAD-Seq analyses but had the same *mtMutS* and *28S* barcodes as sp. B (Quattrini et al., 2019). Perhaps these two groups of *tumulosa* have recently diverged and alleles have not yet completely sorted, and so the incongruent placement depending on marker used is not that surprising. One other incongruence with prior RAD-Seq results (Quattrini et al., 2019) was that *S. ceramensis* appeared highly admixed with *S. humilis* in the current study. This was somewhat surprising as these species were clearly distinguished by DNA barcodes and RAD-Seq data in Quattrini et al. (2019). These species were pooled together in the target enrichment step, and because two *S. ceramensis* had the same i7 index as one *S. humilis* it is possible that index jumping occurred (i.e., mis-assignment of libraries to a different index than expected). However, it is also possible that this is due to hybridization. Hybridization between these two species was not tested in Quattrini et al. (2019), but it is not impossible as hybridization is an important evolutionary mechanism in the diversification of *Sinularia.*

Species delimitation results for clade 5C were also mostly congruent with prior RAD-Seq analyses (Quattrini et al. 2019), but there were also a few exceptions. Analyses in Quattrini et al. (2019) clearly indicated the presence of eight putative species: *S. acuta/abrupta*, *S. densa, S. exilis, S. lochmodes, S. maxima, S. penghuensis, S. slieringsi/penghuensis,* and *S. wanannensis*. Our results initially indicated the presence of only six putative species, grouping *S. exilis, S, maxima,* and *S. wanannensis* together, but upon further analyses at *K*=8, *S. maxima* separated from the other two species. In Quattrini et al. (2019), *S. exilis* grouped separately in DAPC analysis, but the STRUCTURE analysis suggested that it was equally admixed between *S. densa, S. maxima,* and *S. wanannensis.* Quattrini et al. (2019) discussed that a more complete picture would emerge by including more *S. exilis* individuals and including the missing morphospecies in clade 5C. The few differences noted between this study and Quattrini et al. (2019) are more likely a reflection of the recent divergence of *Sinularia* combined with rampant hybridization and inadequate taxon sampling of the clade, rather than the method used (i.e., RAD-Seq vs UCE target-capture) to delimit species.

Future Considerations

Our newly enhanced bait set designed to target UCEs and exons in octocorals works well to delimit among closely-related species and populations of octocorals. An additional test of this bait set on a wide diversity of octocorals recovered ~2,000 loci from each species (unpubl. data). Also, this newly designed bait set can be combined with the hexacoral-v2 bait set (Cowman et al. 2020) to address deep to shallow level evolutionary questions across the anthozoan tree of life. We retained the use of "UCE" and "exon" in this study because of how the targeted loci were designed (i.e., exon loci from transcriptomes and UCE loci from genomes). However, as demonstrated by others (Bossert & Danforth, 2018; Hedin et al., 2019), "UCE" loci likely contain coding regions as well, and similarly, it is possible that exon loci contain intronic regions. Therefore, we can perhaps refer to these regions as "conserved regions or elements" in future work, while further refining analytical methods (see Hedin et al. 2019).

A few additional methodological considerations would improve the use of this locus set in future studies. Although the mean alignment length was $1,139 \pm 639$ bp per locus, a few of the recovered alignments were very long (18 alignments >4,000 bp across both *Alcyonium* and *Sinularia* clades 4 and 5C). Many of these long, edge-trimmed alignments had large indels in the flanking regions, and it might be useful to remove these and/or clip alignments to approximately the same sizes to help reduce the possibility of aligning non-homologous regions. In addition, by following the phyluce tutorial for bait design, we ensured that few to no duplicate loci would be targeted, yet in analyses of specimens that were target-enriched with the new bait set and sequenced on the NovaSeq platform, phyluce randomly discarded 200-500 "duplicate" loci per specimen. These loci were removed because they either matched multiple contigs or contigs contained multiple UCE/exon loci, which perhaps is a reflection of large contig sizes obtained as compared to previous studies (e.g., Branstetter et al., 2017; Quattrini et al., 2018). It is also possible that some of these loci represent exons of the same protein as shown in Hedin et al.

(2019). Further investigation is needed to determine whether or not these "duplicates" were discarded unnecessarily. If they were not real duplicates, then keeping them could increase the number of loci retained per specimen and lead to more complete taxon occupancy matrices for analyses. In general, exploring additional bioinformatic tools to obtain and align loci might be beneficial for future studies.

The SNP calling pipeline used herein was successful at obtaining filtered SNPs for species delimitation and population differentiation of corals. To avoid the possibility of including linked SNPs in analysis, we filtered our SNPs to one SNP per locus or to one SNP per 1,000 bp for locus alignments $\geq$ 2,000 bp (~6% of all aligned loci were > 2,000 bp). This filter can easily be set at different numbers depending on user's goals. Another important consideration when using SNPs is having sufficient read coverage at each locus to avoid erroneous SNP calls. We called SNPs only at positions that had a read coverage of at least 10X. Although 10X might be considered low for some types of analyses, this parameter can also easily be changed by the user. Furthermore, it is possible to increase sequencing to maximize coverage, which could be important for population genomic studies that include specimens with highly degraded DNA (McCormack et al. 2016). Nevertheless, even with only 152 filtered SNPs, the data were sufficient to differentiate populations of *A. coralloides*, highlighting the promise of this method in population genomic studies.

Based on both prior studies (Quattrini et al. 2019) and ecological and reproductive data (McFadden 1999; McFadden et al. 2001), the clustering-based methods used in this study were successful at delimiting species. However, including additional species delimitation methods, such as multispecies coalescent models (e.g., BPP, Yang, 2015) and unsupervised machine learning (Derkarabetian et al. 2019) would provide more evidence of species boundaries. It has become clear that including phylogenomic data in integrative approaches to species delimitation in corals is necessary (Herrera & Shank 2016; Quattrini et al. 2019; Arrigoni et al., 2020).

References
Andras, J., & Rypien, K. (2009). Isolation and characterization of microsatellite loci in the Caribbean sea fan coral, *Gorgonia ventalina*. *Molecular Ecology Resources*, *9*, 1036–1038.

Arrigoni, R., Kitano, Y. F., Stolarski, J., Hoeksema, B. W., Fukami, H., Stefani, F., ... Benzoni, F. (2014). A phylogeny reconstruction of the Dendrophylliidae (Cnidaria, Scleractinia) based on

molecular and micromorphological criteria, and its ecological implications. *Zoologica Scripta*, *43*, 661–688.

Arrigoni, R., Berumen, M. L., Mariappan, K. G., Beck, P. S., Hulver, A. M., Montano, S., ... & Benzoni, F. (2020). Towards a rigorous species delimitation framework for scleractinian corals based on RAD sequencing: the case study of *Leptastrea* from the Indo-Pacific. *Coral Reefs*, 1-25.

Bankevich, A., Nurk, S., Antipov, S., Gurevich, A., Dvorkin, M., Kulikov, A. S., ... Pevzner, P.A. (2012). SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology, 19*(5), 455–477. doi: 10.1089/cmb.2012.002

Baums, I.B., Boulay, J.N., Polato, N.R., & Hellberg, M. (2012). No gene flow across the eastern Pacific barrier in the reef-building coral *Porites lobata*. *Molecular Ecology*, *21*, 5418–5433.

Benayahu, Y., Ofwegen, L. P. van, Dai, C.-F., Jeng, M.-S., Soong, K., Shlagman, A., ... McFadden, C. S. (2018). The octocorals of Dongsha Atoll (South China Sea): An iterative approach to species identification using classical taxonomy and molecular barcodes. *Zoological Studies, 57*(50). doi: 10.6620/ZS.2018.57-50

Benayahu, Y., Ofwegen, L. P. van, & McFadden, C. S. (2018) Evaluating the genus *Cespitularia* Milne Edwards & Haime, 1850 with descriptions of new genera of the family Xeniidae (Octocorallia, Alcyonacea). *ZooKeys*, *754*, 63–101.

Blair, C., Bryson Jr., R. W., Linkem, C. W., Lazcano, D., Klicka, J., & McCormack, J. E. (2019). Cryptic diversity in the Mexican highlands: Thousands of UCE loci help illuminate phylogenetic relationships, species limits and divergence times of montane rattlesnakes (Viperidae: *Crotalus*). *Molecular Ecology Resources*, *19*(2), 349–365.

Bolger, A. M., Lohse, M., & Usadel, B. (2014). Trimmomatic: A flexible trimmer for Illumina Sequence Data. *Bioinformatics, 30*(15), 2114–2120. doi: 10.1093/bioinformatics/btu170

Bossert, S., & Danforth, B. N. (2018). On the universality of target-enrichment baits for phylogenomic research. *Methods in Ecology and Evolution*, *9*(6), 1453-1460.

Bracco, A., Liu, G., Galaska, M. P., Quattrini, A. M., & Herrera, S. (2019). Integrating physical circulation models and genetic approaches to investigate population connectivity in deep-sea corals. *Journal of Marine Systems*, *198*, 103189.

Branstetter, M. G., Longino, J. T., Ward, P. S., & Faircloth, B. C. (2017). Enriching the ant tree of life: Enhanced UCE bait set for genome-scale phylogenetics of ants and other Hymenoptera. *Methods in Ecology and Evolution*, *8*(6), 768–776.

Budd, A.F., & Stolarski, J. (2009). Searching for new morphological characters in the systematics of scleractinian reef corals: comparison of septal teeth and granules between Atlantic and Pacific Mussidae. *Acta Zoologica*, *90*, 142–165.

Cairou, M., Duret, L., & Charlat, S. (2013). Is RAD-seq suitable for phylogenetic inference? An in silico assessment and optimization. *Ecology and Evolution, 3*(4), 846–852. doi: 10.1002/ece3.512

Carpenter, K. E., Abrar, M., Aeby, G., Aronson, R. B., Banks, S., Bruckner, A., ...Wood, E. (2011). One-third of reef-building corals face elevated extinction risk from climate change and local impacts. *Science, 321*, 560–563.

Chhatre, V. E., & Emerson, K. J. (2017). StrAuto: automation and parallelization of STRUCTURE analysis. *BMC Bioinformatics, 18*(1), 192. doi: 10.1186/s12859-017-1593-0

Cowman, P. F., Parravicini, V., Kulbicki.M., & Floeter, S. R. (2017). The biogeography of tropical reef fishes: endemism and provinciality through time. *Biological Reviews*, *92*, 2112–2130.

Cowman, P. F., Quattrini, A. M., Bridge, T. C. L., Watkins-Colwell, G. J., Fadli, N., Grinblat, M., ... Baird, A. H. (2020). An enhanced target-enrichment bait set for Hexacorallia provides phylogenomic resolution of the staghorn corals (Acroporidae) and close relatives. bioRxiv, doi.org/10.1101/2020.02.25.965517

Crandall, E. D., Riginos, C., Bird, C. E., Liggins, L., Trem, E., Beger, M., ... Gaither, M. R. (2019). The molecular biogeography of the IndoPacific: Testing hypotheses with multispecies genetic patterns. *Global Ecology & Biogeography*, *2019*, 1–18.

Danecek, P., Auton, A., Abescasis, G., Albers, C., Banks, E., DePristo, M.A. … 100 Genomes Project Analysis Group. (2011). The variant call format and VCFTools. *Bioinformatics, 27*(15), 2156–2158. doi: 10.1093/bioinformatics/btr330

Derkarabetian, S., Benavides, L. R., Giribet, G. (2019). Sequence capture phylogenomics of historical ethanol-preserved museum specimens: unlocking the rest of the vault. *Molecular Ecology Resources*, *19*(6), 1531-1544.

Derkarabetian, S., Castillo, S., Koo, P. K., Ovchinnikov, S., & Hedin, M. (2019). A demonstration of unsupervised machine learning in species delimitation. *Molecular Phylogenetics and Evolution*, *139*, 106562.

Duggen, S., Hoernle, K., Van den Bogaard, P., Rüpke, L., & Morgan, J. P. (2003). Deep roots of the Messinian salinity crisis. *Nature*, *422*(6932), 602-606.

Eaton, D. A., Spriggs, E. L., Park, B., & Donoghue, M. J. (2017). Misconceptions on missing data in RAD-seq phylogenetics with a deep-scale example from flowering plants. *Systematic Biology, 66*(3), 399–412.

Fabricius, K., & Alderslade, P. (2001). *Soft corals and sea fans: a comprehensive guide to the tropical shallow-water genera of the central-West Pacific, the Indian Ocean and the Red Sea.* Australian Institute of Marine Science.

Faircloth, B. C. (2013). illumiprocessor: a trimmomatic wrapper for parallel adapter and quality trimming. doi: 10.6079/J9ILL.

Faircloth, B. C. (2016). PHYLUCE is a software package for the analysis of conserved genomic loci. *Bioinformatics, 32*(5), 786–788. doi: 10.1093/bioinformatics/btv646

Faircloth, B. C., McCormack, J. E., Crawford, N. G., Harvey, M. G., Brumfield, R. T., & Glenn, T.C. (2012). Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. *Systematic Biology, 61*(5), 717–726. doi: 10.1093/sysbio/sys004

Forsman, Z. H., Barshis, D. J., Hunter, C. L., & Toonen, R. J. (2009). Shape-shifting corals: Molecular markers show morphology is evolutionarily plastic in *Porites*. *BMC Evolutionary Biology*, *9*, 45.

Forsman, Z. H., Knapp, I.S.S., Tisthammer, K., Eaton, D.A.R., Belcaid, M., & Toonen, R.J. (2017). Coral hybridization or phenotypic variation? Genomic data reveal gene flow between *Porites lobata* and *P. compressa*. *Molecular Phylogenetics and Evolution*, *111*, 132–148.

Groot, S., & Weinberg, S. (1982). Biogeography, taxonomical status and ecology of *Alcyonium* (*Parerythropodium*) *coralloides* (Pallas, 1766). *Marine Ecology*, *3*, 293–312.

Gutiérrez-Rodríguez, C., & Lasker, H.R. (2004). Microsatellite variation reveals high levels of genetic variability and population structure in the gorgonian coral *Pseudopterogorgia elisabethae* across the Bahamas. *Molecular Ecology*, *13*, 2211–2221.

Hartnoll, R. G. (1977). Reproductive strategy in two British species of *Alcyonium*. In B. F. Keegan, P. O. Ceidigh, & P. J. S. Boaden (Eds.), *Biology of benthic organisms* (pp. 321–328). Pergamon Press.

Harvey, M. G., Smith, B. T., Glenn, T. C., Faircloth, B. C., & Brumfield, R. T. (2016). Sequence capture versus restriction site associated DNA sequencing for shallow systematics. *Systematic Biology, 65*(5), 910–924. doi: 10.1093/sysbio/syw036

Hawkins, M. T. R., Leonard, J. A., Helgen, K. M., McDonough, M. M., Rockwood, L. L., & Maldonado, J. E. (2016). Evolutionary history of endemic Sulawesi squirrels constructed from UCEs and mitogenomes sequenced from museum specimens. *BMC Evolutionary Biology, 16*(80), 1–16. doi: 10.1186/s12862-016-0650-z

Hedin, M., Derkarabetian, S., Blair, J., & Paquin, P. (2018). Sequence capture phylogenomics of eyeless Cicurina spiders from Texas caves, with emphasis on US federally-endangered species from Bexar County (Araneae, Hahniidae). *ZooKeys*, *769*, 49–76.

Hedin, M., Derkarabetian, S., Alfaro, A., Ramírez, M. J., & Bond, J. E. (2019) Phylogenomic analysis and revised classification of atypoid mygalomorph spiders (Araneae, Mygalomorphae), with notes on arachnid ultraconserved element loci. *PeerJ*, *7*, e6864.

Herrera, S., & Shank, T. M. (2016). RAD sequencing enables unprecedented phylogenetic resolution and objective species delimitation in recalcitrant divergent taxa. *Molecular Phylogenetics and Evolution, 100*, 70–79. doi: 10.1016/j.ympev.2016.03.010

Hodge, J. R., & Bellwood, D. R. (2016). The geography of speciation in coral reef fishes: the relative importance of biogeographical barriers in separating sister-species. *Journal of Biogeography*, *43*, 1324–1335.

Huang, D., Meier, R., Todd, P.A., & Chou, L.M. (2008). Slow mitochondrial COI sequence evolution at the base of the Metazoan tree and its implications for DNA barcoding. *Journal of Molecular Evolution*, *66*, 167–174.

Hughes, T. P., Bellwood, D. R., Connolly, S. R., Cornell, H. V., & Karlson, R. H. (2014). Double jeopardy and global extinction risk in corals and reef fishes. *Current Biology*, *24*, 2946–2951.

Johnston, E. C., Forsman, Z. H., Flot, J.-F., Schmidt-Roach, S., Pinzón, J. H., Knapp, I. S. S., & Toonen, R. J. (2017). A genomic glance through the fog of plasticity and diversification in *Pocillopora*. *Scientific Reports*, *7*(1), 5991.

Jombart, T. (2008). adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics, 24*(11), 1403–1405. doi: 10.1093/bioinformatics/btn129.

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics, 11*(94). doi: 10.1186/1471-2156-11-94

Katoh, K., Misawa, K., Kuma, K., & Miyata, T. (2002). MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Research, 30*(14), 3059–3066. doi: 10.1093/nar/gkf436

Ladner, J. T., & Palumbi, S. R. (2012). Extensive sympatry, cryptic diversity and introgression throughout the geographic distribution of two coral species complexes *Molecular Ecology*, *21*, 2224–2238.

LeGoff, M.C., Pybus, O.G., & Rogers, A.D. (2004). Genetic structure of the deep-sea coral *Lophelia pertusa* in the northeast Atlantic revealed by microsatellites and internal transcribed spacer sequences. *Molecular Ecology*, *13*, 537–549.

Leydet, K. P., Grupstra, C. G. B., Coma, R., Ribes, M., & Hellberg, M. E. (2018) Host-targeted RAD-Seq reveals genetic changes in the coral *Oculina patagonica* associated with range expansion along the Spanish Mediterranean coast. *Molecular Ecology*, *27*, 2529–2543.
Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv preprint arXiv:1303.3997*.

Li, H. & Durbin, R. (2009) Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics, 25, 1754-60.*

Manuel, R. L. (1981). *British Anthozoa (Synopses of the British fauna, no. 18)*. Academic Press.

Manthey, J. D., Campillo, L. C., Burns, K. J., & Moyle, R. G. (2016). Comparison of target-capture and restriction-site associated DNA sequencing for phylogenomics: a test in cardinalid tanagers (Aves, Genus: Piranga). *Systematic Biology*, *65*(4), 640–650.

Manuputty, A. E. W., & Ofwegen, L. P. van. (2007). The genus *Sinularia* (Octocorallia: Alcyonacea) from Ambon and Seram (Moluccas, Indonesia). *Zoologische Mededelingen Leiden, 81* (11), 187–216.

Marti-Puig, P., Forsman, Z. H., Haverkort-Yeh, R. D., Knapp, I. S., Maragos, J. E., & Toonen,

R. J. (2014). Extreme phenotypic polymorphism in the coral genus *Pocillopora*: micro-morphology corresponds to mitochondrial groups, while colony morphology does not. *Bulletin of Marine Science*, *90*, 211–231.

McCormack, J. E., Faircloth, B. C., Crawford, N. G., Gowaty, P. A., Brumfield, R. T., & Glenn, T. C. (2012). Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species tree analysis. *Genome Research, 22*(4), 746–754. doi: 10.1101/gr.125864.111

McCormack, J. E., Tsai, W. E., & Faircloth, B. C. (2016). Sequence capture of ultraconserved elements from bird museum specimens. *Molecular Ecology Resources, 16*(5): 1189–1203. doi: 10.1111/1755-0998.12466

McFadden, C.S. (1999). Genetic and taxonomic relationships among Northeastern Atlantic and Mediterranean populations of the soft coral *Alcyonium coralloides*. *Marine Biology, 133*, 171–184. doi: 10.1007/s002270050456

McFadden, C.S., & Hutchinson, M.B. (2004). Molecular evidence for the hybrid origin of species in the soft coral genus *Alcyonium* (Cnidaria: Anthozoa: Octocorallia). *Molecular Ecology, 13*(6): 1495-1505. doi: 10.1111/j.1365-294X.2004.02167.x

McFadden, C.S., Benayahu, Y., Pante, E., Thoma, J.N., Nevarez, P.A., & France, S.C. (2011). Limitations of mitochondrial gene barcoding in Octocorallia. *Molecular Ecology Resources, 11*(1): 19–31. doi: 10.1111/j.1755-0998.2010.02875.x

McFadden, C.S., Brown, A.S., Brayton., C., Hunt, C.B., & Ofwegen, L.P. van (2014). Application of DNA barcoding in biodiversity studies of shallow-water octocorals: molecular proxies agree with morphological estimates of species richness in Palau. *Coral Reefs, 33*, 275–286. doi: 10.1007/s00338-013-1123-0

McFadden, C. S., Donahue, R., Hadland, B. K., & Weston, R. (2001). A molecular phylogenetic analysis of reproductive trait evolution in the soft coral genus *Alcyonium*. *Evolution*, *55*(1), 54–67. doi: 10.1111/j.0014-3820.2001.tb01272.x

McFadden, C. S., Ofwegen, L. P. van, Beckman, E. J., Benayahu, Y., & Alderslade, P. (2009). Molecular systematics of the speciose Indo-Pacific soft coral genus, *Sinularia* (Anthozoa: Octocorallia). *Invertebrate Biology, 128*(4), 303–323. doi:10.1111/j.1744-7410.2009.00179.x

McFadden, C. S., Haverkort-Yeh, R., Reynolds, A. M., Halász, A., Quattrini, A. M., Forsman, Z. H., Benayahu, Y., & Toonen, R. J. (2017). Species boundaries in the absence of morphological,

ecological or geographical differentiation in the Red Sea octocoral genus *Ovabunda* (Alcyonacea: Xeniidae). *Molecular Phylogenetics and Evolution, 112*, 174–184. doi.org/10.1016/j.ympev.2017.04.025

McFadden, C.S., Rettig, P. M., & Beckman, E. J. (2005). Molecular evidence for hybridization between two alcyoniid soft coral species with contrasting life histories. *Integrative and Comparative Biology*, *45*, 1042.

McFadden, C. S., Sánchez, J. A., & France, S. C. (2010). Molecular phylogenetic insights into the evolution of Octocorallia: a review. *Integrative and Comparative Biology, 50*(3), 389–410. doi: 10.1093/icb/icq056

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., ... & DePristo, M. A. (2010). The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, *20*(9), 1297–1303.

Ofwegen, L. P. van (2002). Status of knowledge of the Indo-Pacific soft coral genus *Sinularia* May, 1898 (Anthozoa: Octocorallia). *Proceedings of the 9th International Coral Reef Symposium*, *1*, 167–171.

Oswald, J. A., Harvey, M. G., Remsen, R. C., Foxworth, D. U., Dittmann, D. L., Cardiff, S. W., & Brumfield, R. T. (2019). Evolutionary dynamics of hybridization and introgression following the recent colonization of Glossy Ibis (Aves: *Plegadis falcinellus*) into the New World. *Molecular Ecology*, *28*(7), 1675–1691.

Pandolfi, J.M., Connolly, S.R., Marshall, D.J., & Cohen, A.L. (2011). Projecting coral reef futures under global warming and ocean acidification. *Science*, *333*, 418–422.

Pante, E., Abdelkrim, J., Viricel, A., Gey, D., France, S. C., Boisselier, M. C., & Samadi, S. (2015). Use of RAD sequencing for delimiting species. *Heredity*, *114*, 450–459. doi:10.1038/hdy.2014.105

Paz-García, D. A., Hellberg, M. E., García-de-León, F. J., & Balart, E. F. (2015). Switch between morphospecies of *Pocillopora* corals. *American Naturalist*, *186*, 434–440.

Prada, C., DeBiasse, M. B., Neigel, J. E., Yednock, B., Stake, J. L., Forsman, Z. H., Baums, I. B., & Hellberg, M. E. (2014). Genetic species delineation among branching Caribbean *Porites* corals. *Coral Reefs*, *33*, 1019–1030.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics, 155*(2), 945–959.

Quattrini, A. M., Baums, I. B., Shank, T. M., Morrison, C. L., & Cordes, E. E. (2015). Testing the depth-differentiation hypothesis in a deepwater octocoral. *Proceedings of the Royal Society B*, 282, 20150008.

Quattrini, A. M., Faircloth, B. C., Dueñas, L. F., Bridge, T. C. L., Brugler, M. R., Calixto-Botía, I. F., … McFadden, C. S. (2018). Universal target-enrichment baits for anthozoan (Cnidaria) phylogenomics: New approaches to long-standing problems. *Molecular Ecology Resources, 18*(2), 281–295. doi: 10.1111/1755-0998.12736

Quattrini, A. M., Wu, T., Soong, K., Jeng, M-S., Benayahu, Y., & McFadden, C. S. (2019). A next generation approach to species delimitation reveals the role of hybridization in a cryptic species complex of corals. *BMC Evolutionary Biology, 19*(116). doi: 10.1186/s12862-019-1427-y

Quattrini, A.M., Rodriguez, E. Faircloth, B.C., Cowman, P.F…..McFadden, C.S. (accepted). Paleoclimate ocean conditions shaped the evolution of corals and their skeletons through deep time. *Nature Ecology and Evolution.*

Quek, R. Z. B., Jain, S. S., Neo, M. L., Rouse G. W., Huang, D. (2020) Transcriptome-based target-enrichment baits for stony corals (Cnidaria: Anthozoa: Scleractinia). *Molecular Ecology Resources.* doi.org/10.1111/1755-0998.13150

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Riegl, B.M., & Dodge, R.E. (Eds.). (2019). *Coral reefs of the world*. Springer.

Roberts, J.M., Wheeler, A.J., & Freiwald, A. (2006.) Reefs of the deep: the biology and geology of cold-water coral ecosystems. *Science*, *312*, 543–547.

Ruane, S., & Austin, C. C. (2017). Phylogenomics using formalin-fixed and 100+ year-old intractable natural history specimens. *Molecular Ecology Resources*, *17*(5),1003–1008. doi: 10.1111/1755-0998.12655

Shearer, T. L. & Coffroth, M. A. (2008). Barcoding corals: limited by interspecific divergence, not intraspecific variation. *Molecular Ecology Resources*, *8*, 247–255. doi: 10.1111/j.1471-8286.2007.01996.x

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics, 30*(9), 1312–1313. doi: 10.1093/bioinformatics/btu033

Starrett, J., Derkarabetian, S., Hedin, M., Bryson Jr, R. W., McCormack, J. E., & Faircloth, B. C. (2017). High phylogenetic utility of an ultraconserved element probe set designed for Arachnida. *Molecular Ecology Resources*, *17*(4), 812–823.

Stolarski J., Kitahara, M.V., Miller, D.J., Cairns, S.D., Mazur, M., Meibom, A. (2011). The ancient evolutionary origins of Scleractinia revealed by azooxanthellate corals. *BMC Evolutionary Biology, 11*(1): 316.

Veron, J. E. N. (2000). *Corals of the world: Vols. 1-3*. Australian Institute of Marine Science and CRR.

Viricel, A., Pante, E., Dabin, W., & Simon-Bouhet, B. (2013). Applicability of RAD-tag genotyping for interfamilial comparisons: empirical data from two cetaceans. *Molecular Ecology Resources, 14*(3), 597–605. doi: 10.1111/1755-0998.12206

Weinberg, S. (1977). Revision of the common Octocorallia of the Mediterranean circalittoral II. Alcyonacea. *Beaufortia, 25*(326), 131–166.

Winker, K., Glenn, T. C., & Faircloth, B. C. (2018). Ultraconserved elements (UCEs) illuminate the population genomics of a recent, high-latitude avian speciation event. *PeerJ, 6*, e5735.

Wolf, P. G., Robison, T. A., Johnson, M. G., Sundue, M. A., Testo, W. L., & Rothfels, C. J. (2018). Target sequence capture of nuclear-encoded genes for phylogenetic analysis in ferns. *Applications in Plant Sciences*, *6*(5), e1148.

Yang, Z. (2015). The BPP program for species tree estimation and species delimitation. *Current Zoology*, *61*(5), 854-865.

Zarza, E., Faircloth, B. C., Tsai, W. L., Bryson Jr, R. W., Klicka, J., & McCormack, J. E. (2016). Hidden histories of gene flow in highland birds revealed with genomic markers. *Molecular Ecology*, *25*(20), 5144–5157.

Zarza, E., Connors, E. M., Maley, J. M., Tsai, W. L., Heimes, P., Kaplan, M., & McCormack, J. E. (2018). Combining ultraconserved elements and mtDNA data to uncover lineage diversity in a Mexican highland frog (Sarcohyla; Hylidae). *PeerJ*, *6*, e6045.

Author Contributions

CSM and AMQ conceived and designed the study. KLE and AP conducted phylogenetic, SNP calling and species delimitation analyses with guidance from AMQ and CSM. AMQ designed the baits. KLE and AP wrote preliminary drafts of the manuscript. CSM and AMQ wrote the final draft of the manuscript. All authors conducted lab work and approved the final manuscript version.

Table 1. Number of SNPs recovered for each clade.

	# Unfiltered SNPs	# Filtered SNPs
Alcyonium		
digitatum clade	30,384	1,974
c*oralloides* clade	39,302	1,980
acaule clade	37,569	2,114
coralloides populations	4,775	152
Sinularia		
Clade 4	65,261	2,030
Clade 5C	71,400	2,742

Figure Captions

Figure 1. Maximum likelihood phylogeny of *Alcyonium* based on a 75% complete matrix containing 1,264 UCE loci (1,280,803 bp). All unlabeled nodes have 100% bootstrap support (100 b.s. replicates). The tree was rooted with *Sinularia* outgroups.

Figure 2. Probability of membership graph from STRUCTURE analyses and Discriminant Analysis of Principal Components (DAPC) plot for (A-B) *Alcyonium digitatum* clade (1,974 SNPs, *K=3*), (C-D) *Alcyonium coralloides* clade (1,980 SNPs, *K=4），* and (E-F) *Alcyonium acaule* clade (2,114 SNPs, *K=3).*

Figure 3. (A) Map of collection locations for *Alcyonium coralloides,* (B) Probability of membership graph from STRUCTURE analyses and (C) Discriminant Analysis of Principal Components (DAPC) plot for *A. coralloides* (152 SNPs, *K*=4) populations (M1: Marseille, France, M2: Banyuls-sur-Mer, France, CH: Iles Chausey, France, SG: Sagres, Portugal).

Figure 4. Maximum likelihood phylogenies of *Sinularia* clade 4 constructed from (A) UCE loci (1,939 loci, 2,231,641 bp) and (B) RAD-Seq loci (6,343 loci, 549,862 bp) based on 75% incomplete matrices. All unlabeled nodes have 100% bootstrap support (100 b.s. replicates). The phylogeny was rooted with *S. humilis.*

Figure 5. Maximum likelihood phylogenies of *Sinularia* clade 5C constructed from (A) UCE loci (1,240 loci, 1,549,273 bp) and (B) RAD-Seq loci (8,060 loci, 701,433 bp) based on 75% incomplete matrices. All unlabeled nodes have 100% bootstrap support (100 b.s. replicates). A midpoint rooting was used.

Figure 6. Probability of membership graph from STRUCTURE analyses and Discriminant Analysis of Principal Components (DAPC) plot for (A-B) *Sinularia* clade 4 (2,030 SNPs, *K*=5) and (C-D) *Sinularia* clade 5C (2,742 SNPs, *K*=6), and (E-F) *Sinularia* clade 5C (2,742 SNPs, *K*=8).

Supplementary Data
Supplementary File 1. Bait redesign protocol.
Supplementary Table 1. Assembly, locus, and collection data for specimens.