Species boundaries in the absence of morphological, ecological or geographical
differentiation in the Red Sea octocoral genus *Ovabunda* (Alcyonacea: Xeniidae)

Catherine S. McFadden¹, Roxanne Haverkort-Yeh², Alexandra M. Reynolds¹, Anna
Halász³, Andrea M. Quattrini¹, Zac H. Forsman², Yehuda Benayahu³, Robert J. Toonen²

¹Department of Biology, Harvey Mudd College, 1250 N. Dartmouth Ave., Claremont,
CA 91711, USA

²Hawai'i Institute of Marine Biology, School of Ocean & Earth Sciences & Technology,
University of Hawai'i at Mānoa, Kāne'ohe, HI 96744, USA

³School of Zoology, George S. Wise Faculty of Life Sciences, Tel Aviv University,
Ramat Aviv, Tel Aviv 69978, Israel

corresponding author: mcfadden@g.hmc.edu

Abstract

The development of coalescent-based and other multilocus methods for species delimitation has facilitated the identification of cryptic species complexes across the tree of life. A recent taxonomic revision of the ecologically important soft coral genus *Ovabunda* validated 11 morphospecies, all with type localities and overlapping geographic ranges in the Red Sea. A subsequent molecular phylogenetic analysis using mitochondrial and 28S nrDNA genes divided the genus into just two clades, with no apparent genetic distinctions among morphospecies. To further explore species boundaries among morphospecies of *Ovabunda* we sequenced three additional nuclear genes (*ITS*, *ATPS α* , *ATPS β*), and obtained data for 1,332 unlinked SNPs from restriction-site associated DNA sequencing. Both coalescent-based and allele-sharing species delimitation analyses supported four species of *Ovabunda*, each of which included multiple morphotypes encompassing the full range of morphological variation observed within the genus. All four species occurred over the same depth range of 5–41 m, and were sympatric at sites separated by 1,100 km in the Red Sea. The only characters that have been found to distinguish three of the four species are diagnostic substitutions in the nuclear genome; the fourth differs by exhibiting polyp pulsation, a behavioral trait that can be assessed only in live colonies. The lack of any obvious morphological, life history, ecological or geographical differences among these four species begs the question of what drove the evolution and maintenance of reproductive isolating mechanisms in this cryptic species complex.

Keywords: coral reef, cryptic species complex, molecular phylogenetics, reproductive isolation, species delimitation

1. Introduction

Use of molecular tools to assess population structure and biodiversity has led to the discovery of numerous cryptic species across the tree of life (Knowlton 2000; Bickford et al. 2007). The existence of cryptic species—species that are superficially indistinguishable—is often suspected *a priori* when slight morphological or ecological differences are observed among geographically separated populations. Hypothesized species boundaries remain challenging to test in such situations, however, because of the difficulty of assessing reproductive isolation in allopatry (Fujita et al. 2012). In contrast, when morphological or ecological variants co-occur sympatrically, species concepts based on reproductive isolation can be tested directly by observations of interbreeding, or indirectly by assessing genetic exchange among individuals (e.g., McFadden et al. 1997; McFadden 1999). Disjunct allele frequencies and linkage disequilibrium among segregating genetic loci provide powerful evidence of reproductively isolated groups (e.g., cryptic species) within sympatric populations (e.g., Shaffer and Thomson 2007; Flot et al. 2010; Piggott et al. 2011; Ladner and Palumbi 2012).

Species boundaries among anthozoan cnidarians (corals and sea anemones) have been particularly difficult to assess as a result of their often environmentally plastic morphologies (Kim et al. 2004; Prada et al. 2008; Forsman et al. 2009; Marti-Puig et al. 2013; Paz-Garcia et al. 2015) coupled with a general lack of taxon-specific molecular tools necessary to evaluate gene flow at the population level. The mitochondrial genomes of anthozoans evolve notoriously slowly (Shearer and Coffroth 2008; Huang et al. 2008), and few single-copy nuclear loci have been identified that vary appropriately or can be sequenced consistently across taxa (Neigel et al. 2007; Concepcion et al. 2008;

McFadden et al. 2010). Consequently, it has frequently been a challenge to determine if distinct, co-occurring ecomorphs of common taxa represent different species or merely environmentally-induced phenotypic variants (e.g., Gutiérrez-Rodríguez et al. 2009; Prada et al. 2008; Bilewitch et al. 2010; Forsman et al. 2010; Paz-García et al. 2015). Coalescent-based methods of species delimitation that have been applied to other taxa with considerable success (Fujita et al. 2012) require data from a greater number of independently segregating loci than have typically been sequenced in studies of species boundaries in anthozoans (reviewed by Stat et al. 2012), however the number of loci available are gradually increasing, and multi-locus methods are becoming more common (e.g., Reitzel et al. 2013; Boulay et al. 2014; Prada et al. 2014; Pante et al. 2015; Herrera & Shank 2016; Hellberg et al. 2016).

Ovabunda Alderslade, 2001 is a common and ecologically important genus of xeniid octocorals that occupy shallow reef habitats in the Red Sea (Halász et al. 2013). A recent taxonomic revision based on morphological examination of the original type material identified 11 nominal species, all with type localities in the Red Sea; eight of these species are known exclusively from the Red Sea, while three have also been recorded in the western Indian Ocean (Madagascar and the Seychelles) (Halász et al. 2013). A twelfth species was recently described from Thailand (Janes et al. 2014). *Ovabunda* species are opportunists that recruit to open substrate and grow quickly, often monopolizing large areas of primary reef substrate by excluding other corals, including reef-building scleractinians (Benayahu and Loya 1981, 1985).

Halász et al.'s (2013) revision identified a set of morphological characters that have been used previously in the classical taxonomic literature and are considered to be

diagnostic for distinguishing species of *Ovabunda*. These include the number of rows and numbers per row of pinnules (lateral projections) on the polyp tentacles, as well as the dimensions and spacing between individual pinnules. These soft-tissue characters are, however, sometimes difficult to assess; both numbers of rows and numbers of pinnules per row are variable within species and may also vary among polyps within an individual (Halász et al. 2013). In contrast to most other soft corals, the sclerites (calcitic skeletal elements) of the different *Ovabunda* species are uniform in shape and size-range, and thus cannot be used for species diagnosis (Halász et al. 2013).

Subsequent molecular phylogenetic studies have failed to support the distinctions among *Ovabunda* morphospecies. Haverkort-Yeh et al. (2013) examined octocoral diversity at 14 locations within the Red Sea using a suite of two mitochondrial (*mtMutS*, *COI*) and two nuclear (*ITS*, *ATPS α*) markers. They were unable to resolve genetic differences among 5 nominal morphospecies of *Ovabunda*, and also found *Ovabunda* to be paraphyletic with the morphologically similar genus, *Xenia*. Halász et al. (2014) further examined the phylogenetic relationships among 10 nominal morphospecies of *Ovabunda* using DNA sequences from three mitochondrial genes (*mtMutS*, *ND2*, *COI*) and a fragment of nuclear ribosomal 28S rDNA. Their analysis separated the genus into just two clades within each of which there was little to no sequence variation among individuals. Each clade included representatives of 6–10 morphospecies spanning the full range of variation in pinnule characters. Based on the lack of support for any genetic distinctions among the traditionally defined morphospecies, Halász et al. (2014) suggested that the genus *Ovabunda* might consist of just two species that can be

distinguished by the presence or absence of polyp pulsation, a behavioral trait obviously seen only in live colonies.

In the present study we expanded the work of Halász et al. (2014) to include an additional four nuclear loci as well as SNPs derived from RAD-tag sequencing (Toonen et al. 2013; Knapp et al. 2016), and utilized a combination of coalescent-based and allele-sharing species delimitation analyses to estimate the number of species of *Ovabunda* that are present in the central and northern Red Sea. Application of a general lineage species concept (de Queiroz 1998) suggests that *Ovabunda* comprises four sympatric species whose boundaries do not align with any observed morphological, ecological or geographical differences among taxa.

2. Materials and Methods

2.1. Sample collection

Specimens of *Ovabunda* and the closely related genus *Xenia* were collected from 2008–2011 using SCUBA at eight sites in Eilat, Israel, northern Gulf of Aqaba, and from four locations on the west coast of Saudi Arabia, central Red Sea (Fig. 1). Details of these collections have been published previously (Haverkort-Yeh et al. 2013; Halász et al. 2014). Vouchers of all specimens have been deposited at the Zoological Museum of Tel Aviv University (Israeli material) and the Smithsonian Institution (Saudi Arabian material). All specimens were identified to morphospecies using the species-diagnostic polyp characters presented in Halász et al. (2013, 2014), namely the number of rows of pinnules on the tentacles and the number and spacing of pinnules in the outermost row, as well as the occurrence of polyp pulsation in live colonies (Table 1).

2.2. Sanger sequencing of individual gene loci

Extraction of DNA, amplification using polymerase chain reaction (PCR) and sequencing followed published protocols for three mitochondrial genes (*mtMutS*, *COI*, *ND2*), two regions of the nuclear ribosomal gene complex (*28S rDNA*, *ITS*), and a nuclear intron, *ATPS α* (Haverkort-Yeh et al. 2013; Halász et al. 2014). Two additional nuclear introns, *SRP54* and *ATPS β* , were amplified and sequenced using the primers and PCR protocols of Concepcion et al. (2008) and Jarman et al. (2002), respectively. Phase (Stephens et al. 2001) implemented in DNAsp5 (Librado and Rozas 2009) was used to resolve alleles in heterozygous nuclear genotypes in cases where alignment of forward and reverse sequence chromatograms was possible. When alleles could not be phased reliably from direct sequencing, PCR products were cloned and re-sequenced as described in Haverkort-Yeh et al. (2013).

2.3. Restriction-site associated DNA sequencing

Based on preliminary results of multi-locus species delimitation analyses, DNA from individuals belonging to each of three clades (A, B, C) representing putative species was pooled, and a single library was prepared for each clade (Toonen et al. 2013). DNA was quantified using a NanoDrop spectrophotometer, and 100 ng per individual was pooled for clades B (n = 23) and C (n = 27). Because fewer individuals were available for clade A (n = 15), 150 ng per individual was pooled to yield equivalent totals of approximately 2.5 μ g DNA per clade. Insufficient DNA was available for clade D individuals to include them in RADseq analyses. Libraries were prepared using the non-PCR ezRAD protocol

(Toonen et al. 2013; Knapp et al. 2016). DNA was digested simultaneously with the restriction enzymes MboI and Sau3AI (NEB), cleaned using AmpureXP beads, and inserted into the Illumina TruSeq Nano DNA Library Prep Kit. Libraries were validated using an Agilent 2100 BioAnalyzer, quantified using qPCR, and pooled for sequencing as paired-end 2x300 bp runs on an Illumina GAIIX (all steps performed by the Hawai'i Institute of Marine Biology EPSCoR Core sequencing facility).

2.4. Phylogenetic analyses of single-locus sequence data

Sequences were aligned using the L-INS-i method in MAFFT (Katoh et al. 2005), and the best-fit model of evolution for each gene region was subsequently determined using jModelTest v. 2.1.6 (Darriba et al. 2012; Guindon & Gascuel 2003). Maximum likelihood analyses initially were run separately for each gene using Garli 2.0 (Zwickl 2006). In cases where individuals had heterozygous genotypes, both alleles were included in the phylogenetic analysis. All gene regions (with the exception of *SRP54*, see Results 3.2) were then concatenated, and maximum likelihood analyses were run on the combined dataset. For the polymorphic nuclear genes (*ITS*, *ATPS α* and *ATPS β*) only one allele for each individual was included in the concatenated dataset, and all analyses were run twice, once using one randomly chosen allele from each heterozygous individual, and again with the alternative allele included instead. Partitioned analyses were run with different models of evolution applied to each of four separate data partitions: (1) the three mitochondrial genes ("mt") (HKY); (2) *28S rDNA* (K80); (3) *ITS* (HKY + G); and (4) the two ATP synthase introns, *ATPS α* + *ATPS β* (HKY + G). A Bayesian analysis was run on the concatenated dataset using MrBayes v. 3.2 (Ronquist et al. 2012) with the same four

data partitions and models of evolution specified as for the ML analysis. MrBayes was run for 50×10^6 generations (until standard deviation of split partitions < 0.015) with a burn-in of 25% and default Metropolis coupling parameters.

2.5. Species delimitation using a discovery approach

Among several computational methods currently in use for delimiting species without *a priori* assignment of individuals to putative species, Rittmeyer & Austin (2012) recently identified the program Structurama (Huelsenbeck & Andolfatto 2007; Huelsenbeck et al. 2011) as a particularly powerful approach. Structurama successfully delimits recently diverged species without requiring an estimated gene genealogy to be used as a guide tree. Structurama uses the same Bayesian clustering algorithm implemented in Structure (Pritchard et al. 2000) but can treat the number of clusters (k) as a random variable estimated by a Dirichlet process prior.

Sequences were converted to numbered alleles using SNAP Map (Aylor et al. 2006). The three mitochondrial genes were concatenated and treated as a single, haploid locus, with individuals for which we lacked sequence data for any one of the three gene regions coded as missing data. All alleles were coded for *ITS*, *ATPS α* and *ATPS β* , with indels recoded and included in the allele assignment. All four nuclear loci were treated as diploid with individuals assigned phased homozygous or heterozygous genotypes as appropriate. Although the *ITS* and *28S* loci are linked, we were unable to phase alleles between them because they had been sequenced independently. As a result, we ran separate analyses on datasets consisting of five loci (mt, *28S*, *ITS*, *ATPS α* , *ATPS β*) or four loci only with either *28S* or *ITS* removed.

We ran Structurama 2.0 (Huelsenbeck et al. 2011) with model parameters set to a Dirichlet process prior (rv) to allow the number of populations to vary (Pella and Masuda 2006) both with and without admixture. Following the approach taken by Carstens & Satler (2013), alpha was allowed to vary under a gamma probability distribution with shape and scale parameters of (1, 1), (1, 5), (1, 10), (0.1, 1), (0.1, 5), and (0.1, 10). Analyses were run for 10^7 generations and sampled every 1,000 generations with a burn-in of 1,000 (10%), after which mean number of populations, mean partition, and togetherness were run to examine population structure in the dataset.

2.6. Bayesian species delimitation

The Bayesian species delimitation program BPP v2.2 (Rannala & Yang 2003; Yang & Rannala 2010) was used to validate the results of the species discovery approach taken with Structurama. BPP required (1) individuals to be assigned to provisional species *a priori*, and (2) a user-supplied, fully resolved species tree to be used as a guide tree. BPP uses reversible jump Markov Chain Monte Carlo (rjMCMC) methods to evaluate whether or not each ancestral node in the guide tree should be retained or collapsed. *BEAST v. 1.8.1 (Heled & Drummond 2010) was used to estimate a species tree to be used as a guide tree. Analyses were run using only the 4 nuclear loci (*28S rDNA*, *ITS*, *ATPS α* , *ATPS β*) and the set of 38 individuals for which sequence data were complete. One allele was retained per individual; a second analysis that was run using the alternative allele yielded identical results, therefore further resampling of allelic combinations was determined to be unnecessary. Analyses were run for 50 million generations and sampled every 5,000 generations with 20% of 10,000 trees discarded as burnin. Tracer v. 1.6

(Rambaut & Drummond 2007) was used to assess convergence of likelihood plots over time, and the maximum clade credibility tree was used as the guide tree for BPP.

Bayesian species delimitation using BPP was conducted on the same set of 38 individuals used to estimate the guide tree; numbers of individuals per provisional species ranged from two (*Ovabunda* clade D) to 11 (*Ovabunda* clade B). The three mitochondrial loci (*mtMutS*+*COI*+*ND2*) were concatenated and added to the dataset as a fifth locus, with missing data for 5 of 38 individuals. rjMCMC was run for 200,000 generations with a sampling interval of 1 and a burnin of 20,000, using algorithm 0, $\epsilon = 2.0$ and automatic fine-tuning of parameters. To accommodate the different effective population sizes expected for autosomal vs. mitochondrial loci, θ was allowed to vary by applying a heredity multiplier of 0.25 to the mitochondrial loci. Because posterior probabilities are affected by the prior distributions on the ancestral population size (θ) and root age (τ_0) (Yang & Rannala 2010), we followed the approach of Leaché and Fujita (2010) and compared the results of three different combinations of priors: (1) $\theta \sim G(2, 2000)$, $\tau_0 \sim G(2, 2000)$, representing small ancestral population sizes and shallow divergence among species; (2) $\theta \sim G(1, 10)$, $\tau_0 \sim G(1, 10)$, representing large ancestral population sizes and deep divergence among species; and (3) $\theta \sim G(1, 10)$, $\tau_0 \sim G(2, 2000)$, representing large ancestral population sizes and shallow divergence among species. The latter is the most conservative combination and should favor models that support fewer species (Leaché & Fujita 2010). Analyses were repeated with different initial seeds, starting species models, and both with and without gaps and ambiguities removed from the sequences.

2.7. RADseq data processing and species delimitation

pyRAD v3.0 (Eaton 2014) was used to trim and cluster sequences across the three clades (A, B and C) for identification of homologous loci (following Pante et al. 2015; Herrera and Shank 2016). Paired-end reads were merged using PEAR (Zhang et al. 2014), and input into the pyRAD pipeline. The clustering threshold, which is the similarity threshold used to cluster sequences within and across individuals, was set to c 0.85 and the minimum number of individuals per locus (m) to either 2 or 3. In all analyses, the maximum number of low quality (phred quality score <20), ambiguous bases (Ns) retained in each read was set at 10 while the number of SNPs allowed in a final locus was set at 50. Reads with more than 10 N's and short reads (<50 bp) were discarded. We also evaluated the minimum depth of coverage per cluster to make a base call (d 4, 6) and the maximum number of shared polymorphic sites at a locus (p 0.10, 0.25), but no differences were found in the number of loci or the number of SNPs obtained between these parameter settings. The standalone version of DeconSeq v 0.4.3 (Schmieder and Edwards 2011) was used to identify potential contamination by symbiotic zooxanthellae in the pyRAD datasets by screening the *Ovabunda* loci against *Symbiodinium kawagutii* (Lin et al. 2015; http://web.malab.cn/symka_new/).

One unlinked SNP was randomly sampled from each locus using pyRAD to create unlinked SNP datasets. The set of unlinked SNPs generated from the pyRAD datasets run with (c 0.85, m 2) and (c 0.85, m 3) were analyzed using Structurama v2.0 (Huelsenbeck et al. 2011) with the same parameter settings described above.

StrAutoParallel v 1.0 (Chhatre and Emerson 2016) was run with K set from 1 to 4, and data from the $K=3$ run were summarized and plotted.

2.8. Diagnostic characters and DNA barcoding

Subsequent to species delimitation analyses, pairwise genetic distance values (Kimura 2-parameter) at each genetic locus were calculated within and among species using MEGA v. 5 (Tamura et al. 2011). Sequence alignments for each gene were scrutinized to identify SNPs representing pure characteristic attributes (PuCAs, sensu DeSalle et al. 2005) that could be used as diagnostic characters to identify individuals to species.

3. Results

3.1. Morphological identifications

All phylogenetic and species delimitation analyses were conducted with a set of 99 individuals (88 *Ovabunda*, 11 *Xenia*) for which at least six of eight loci had been sequenced (Appendix Table A.1); an additional 40 specimens for which only 3–5 loci were obtained were not included in any subsequent analyses. Morphological analyses (Halász et al. 2013, Haverkort-Yeh et al. 2013) identified the 88 *Ovabunda* specimens to 17 different morphospecies (*O. ainex* (Reinicke, 1997), *O. arabica* (Reinicke, 1995), *O. biseriata* (Verseveldt & Cohen, 1971), *O. crenata* (Reinicke, 1997), *O. faraunensis* (Verseveldt & Cohen, 1971), *O. gohari* (Reinicke, 1997), *O. hamsina* (Reinicke, 1997), *O. impulsatilla* (Verseveldt & Cohen, 1971), *O. macrospiculata* (Gohar, 1940), *O. verseveldti* (Benayahu, 1990), plus seven unidentified morphospecies that could not be matched to a nominal species). Five morphospecies (*O. ainex*, *O. biseriata*, *O. faraunensis*, *O. impulsatilla*, *O. macrospiculata*) were particularly common ($n > 9$ specimens), while none of the other morphospecies were represented by more than four

specimens (Table 1). All 11 *Xenia* specimens were identified as either *X. umbellata* or an unidentified morphospecies.

3.2. Phylogenetic analyses

Maximum likelihood analyses of individual loci recovered poorly resolved, but largely congruent, gene tree topologies for each of the three mitochondrial genes and the nuclear genes *28S rDNA*, *ITS rDNA*, and *ATPS β* (Appendix Figs. A.1-A.4). All of these gene trees revealed moderate (>50%) to strong (>70%) bootstrap support for a clade of *Xenia* and for the pulsating clade of *Ovabunda* (clade A) recognized in previous analyses (Halász et al. 2014). Relationships among the remaining, non-pulsating *Ovabunda* species were poorly resolved in all gene trees. The gene tree for *ATPS α* supported the separation of the *Xenia* clade from *Ovabunda*, but did not resolve a separate clade within *Ovabunda* for the pulsating species; this gene tree was not, however, in direct conflict with the others, merely less well resolved (Appendix Fig A.5). The gene tree for *SRP54* was, however, in conflict with the other seven gene trees; most notably, *SRP54* divided the *Xenia* species among two divergent clades, one of which was internal to a poorly resolved *Ovabunda* clade (Appendix Fig A.6). In addition, we identified up to four *SRP54* alleles in some individuals, suggesting that this locus may not be single-copy in this group as previous studies have assumed for other taxa (Jarman et al. 2002; Concepcion et al. 2008). Consequently, we did not include results from the *SRP54* locus in further analyses.

Maximum likelihood and Bayesian analyses of the concatenated alignment of the seven loci with congruent gene trees (total length = 4,447 bp) recovered identical species

tree topologies (Fig. 2), with support for five major clades. All 11 individuals of *Xenia* belonged to a well-supported clade distinct from *Ovabunda*. Within *Ovabunda*, there was strong to moderate support for four clades. Clade A consisted of 20 pulsating specimens identified as either *O. macrospiculata* or one of 5 unidentified morphospecies discussed in previous work (Halász et al. 2014). Clade D included two specimens, identified as *O. biseriata* and *O. ainex*, that shared distinct, unique genotypes at all nuclear loci. The remaining 67 *Ovabunda* specimens belonged to a well-supported clade within which there was moderate support for two distinct sub-clades, B and C.

3.3. Allele-sharing and linkage disequilibrium

Reproductive isolation among the four clades of *Ovabunda* (A, B, C, D) was supported by the distribution of alleles at polymorphic nuclear loci. Seventeen of the 18 alleles detected at *28S rDNA* were private (i.e. restricted to a single clade), with just a single individual of clade C (A124) sharing an allele that was otherwise found only in clade B individuals. At *ATPS β* , all but one of 39 alleles was private, with just one clade C individual (A017) sharing an allele with clade B. At *ATPS α* , 47 of 49 alleles were private, with two alleles shared among individuals of clades B and C. None of the 72 alleles found at *ITS* were shared across clades. Haploweb diagrams that connect the two alleles found in individuals with heterozygous nuclear genotypes further illustrate the lack of alleles shared among clades (Appendix Figs. A.3-A.5).

There was strong linkage disequilibrium between mitochondrial haplotypes and nuclear genotypes as well as between alleles at unlinked nuclear loci (i.e., between rDNA and the ATPS introns). In only one case (A080) did an individual have a mitochondrial

haplotype characteristic of one clade (B) combined with nuclear genotypes typical of a different clade (C). Likewise, with only two exceptions (A017, A124, both discussed above), alleles at the two rDNA loci were always linked to *ATPS* β alleles private to the same clade.

3.4. *Species delimitation analyses*

Structurama consistently assigned individuals to five clusters (Table 2) whose membership matched almost exactly the five clades detected in the phylogenetic analyses. Cluster 1 represented the *Xenia* clade, and clusters 2–5 corresponded to *Ovabunda* clades A–D. When either *28S* or *ITS* was removed from the dataset and only four unlinked loci were included, a higher probability was assigned to the existence of only four clusters rather than five (Table 2). In particular, when *28S* was omitted the two individuals belonging to clade D were assigned to other clusters. In addition, the cluster assignment of the three individuals of *Ovabunda* that shared alleles across clades (A017, A080, A124) varied somewhat depending on the model parameters used, and was not always in agreement with their position in the phylogenetic analysis.

Bayesian species delimitation analyses using BPP found strong support for all five of the putative species identified by Structurama (Fig. 3). This outcome was not affected by the prior distributions for θ and τ_0 , or by different starting parameters (seed, species model), rjMCMC algorithm, or exclusion of ambiguities and gaps.

3.5. *Species delimitation based on unlinked SNPs*

RADseq yielded a total of 7,427,277 reads. The mean read size per sample ranged from 278 ± 44 bp to 284 ± 37 bp. Following the removal of low quality and short reads, 6,781,094 reads were retained. An approximately equal (31–35%) proportion of reads was recovered for each of clades A, B and C. A total of 1,619 loci (1,332 unlinked SNPs) were present in the pyRAD dataset generated with parameters $c\ 0.85$, $m2$; this number was reduced to 233 loci (186 unlinked SNPS) in the ($c\ 0.85$, $m3$) dataset. Fewer than 0.7% of loci were identified as belonging to *Symbiodinium*. Structurama identified three populations for all shape and scale parameters investigated for both SNP datasets ($c0.85$, $m2$ and $c0.85$, $m3$) run with and without admixture (Appendix Fig. A.7).

3.6. Diagnostic characters and DNA barcoding

Within each of the five distinct clades supported by phylogenetic and species delimitation analyses, mean genetic distances (Kimura 2-parameter) among individuals within clades ranged from 0.02% at the nearly invariant *ND2* gene to 2.41% at *ATPS α* (Table 3). Mean genetic distances among clades were typically 3–5X higher, ranging from 0.07% (*ND2*) to 3.46% (*ATPS α*). The ranges of genetic distance values observed between individuals in different clades were, however, very similar to those between individuals in the same clade; interclade distances of 0 were recorded at all loci, and in one case (*ATPS α*) the maximum genetic distance observed within a clade exceeded the maximum value observed among clades (Table 3). As a consequence of this broad overlap between intra- and interclade genetic distances, specimens cannot be reliably identified to clade based solely on genetic distance from a reference taxon (i.e. there is no barcoding gap among clades).

The number of pure characteristic attributes (PuCAs) ranged from just one in *ND2* (distinguishing *Ovabunda* clade A from the other four clades) to 30 in the *ITS* region, and totaled 54 across all seven genes (Fig. 4). *Ovabunda* clade B had only three PuCAs, whereas the *Xenia* clade was distinguished by 25, a majority of them (16) located in the *ITS* region. In addition to these species-specific PuCAs, six substitutions distinguished clades B and C from the other three clades but not from one another (Fig. 5). When those sites are included in a compound CA (DeSalle et al. 2005), all five clades can be distinguished using either the *28S* (7-nt compound CA) or *ITS* (36-nt compound CA) regions alone. None of the other five gene regions sequenced had pure or compound CAs able to discriminate all five clades.

3.7. Geographical and ecological differentiation of clades

The five genetic clades (*Xenia* plus four clades of *Ovabunda*) recognized in the phylogenetic and species delimitation analyses could not be distinguished geographically. At least two of five clades were found at every site at which more than a single specimen was collected, and three or more clades were present at 7 of 10 sites (Fig. 1). All five clades were collected from the relatively small reef area of only a few hundred square meters adjacent to the Interuniversity Institute for Marine Sciences (IUI) in Eilat, northern Gulf of Aqaba. All clades except D were also found along the coast of Saudi Arabia in the central Red Sea (Fig. 1), suggesting that they are sympatric over the 1,100 km distance between the two geographic regions sampled in this study.

Collection data also indicate that clades do not appear to be spatially or ecologically separated by depth (Fig. 5). There was extensive overlap among all five

clades in the range of depths at which they were collected. Although *Ovabunda* clade A and the *Xenia* clade tended to be more frequently collected at depths of 0–10 m while *Ovabunda* clades B and C were more frequently collected at depths greater than 10 m, all clades were found over the entire sampled depth range. The two individuals of clade D, both collected at Eilat, came from depths of 5 m and 41 m.

4. Discussion

Multilocus phylogenetic and species delimitation analyses support the presence of four reproductively isolated *Ovabunda* species found throughout the northern Gulf of Aqaba and central Red Sea. This number is more than the two taxa suggested by previous molecular phylogenetic analyses of primarily mitochondrial loci (Halász et al. 2014), but fewer than the 10 morphospecies described using traditional taxonomic characters (Halász et al. 2013, 2014). Despite the lack of genetic differentiation at the mitochondrial loci typically used for molecular barcoding (McFadden et al. 2011, 2014), the distributions of alleles at nuclear loci suggest that these four species have been reproductively isolated for long enough for lineage sorting to be nearly complete. Only three of 99 individuals had multilocus genotypes suggestive of either incomplete lineage sorting or hybridization between groups.

Reproductive isolation does not appear to have been accompanied by any obvious morphological, geographical, or ecological differentiation among species apart from the behavioral feature of polyp pulsation in species A (Halász et al. 2014). Species A, B and C each encompass the full range of morphological variation in pinnule characters observed within the genus (from 1–3 rows of pinnules and 7–26 pinnules per row; Table

1), and inhabit overlapping depths across a broad range of sympatry. At this time, no characters other than multilocus nuclear genotypes (Fig. 4) have been identified that reliably discriminate the three non-pulsating species (B, C, D) from one another or distinguish them from non-living material of the pulsating species, A. Species A has been proposed to correspond to the common morphospecies *O. macrospiculata* (Halász et al. 2013, 2014). In the absence of diagnostic morphological characters, appropriate binomials have yet to be determined for the other species.

The present inability to distinguish clades B, C and D in the field limits the inferences that can be made about the mechanisms that might enforce reproductive isolation among these syntopic species. All xeniid octocorals, including *Ovabunda*, brood their offspring (Benayahu 1991; Kahng et al. 2011), typically releasing planula larvae over a reproductive period that lasts several months (Benayahu and Loya 1984a, b). Previous studies of *O. macrospiculata* have suggested that sperm and egg maturation peak in June–July, with release of planulae occurring bi-weekly from late May through September (Benayahu and Loya 1984a, b). Details of the timing and mechanism of fertilization remain unknown, however, for this species or any other brooding octocoral (Benayahu 1991; Kahng et al. 2011). In addition, given the lack of diagnostic morphological traits, it is likely that previous studies of reproduction in *O. macrospiculata* have inadvertently included a mix of individuals from different species. Halász et al. (2014), however, collected planulae from *Ovabunda* colonies that were subsequently genotyped to confirm their identity. Their sample sizes were very small (4 colonies of species A and one each of species B and C) but nonetheless indicated some overlap in the timing of planulation among species. Any significant differences in the

timing of fertilization (as opposed to release of brooded larvae) that might act as a reproductive barrier among species remain unknown, and would be difficult to ascertain in these brooding species that have internal fertilization.

In the absence of any obvious geographical, ecological or life history differences among the four clades of *Ovabunda*, we can only speculate about the mechanisms that may have driven their speciation. The Red Sea is a relatively geologically young (3–4 my) biogeographic region that has experienced high variability in temperature and salinity throughout the Pleistocene (reviewed in DiBattista et al. 2016). Recent paleoclimate models coupled with phylogenetic studies suggest that some species may have survived these abrupt periodic environmental changes in refugia located at the extreme northern and southern ends of the Red Sea or even immediately outside it in the Gulf of Aden (DiBattista et al. 2016). Phylogenetic evidence from endemic lineages of fish suggest that peripatric speciation may have driven diversification under these circumstances (DiBattista et al. 2016). Changes in the timing or chemistry of fertilization in *Ovabunda* populations during times when they were isolated in such refugia is one possible mechanism that could explain their present-day reproductive isolation. We currently lack an appropriate time-calibrated phylogeny for xeniid octocorals (a group with no fossil record) that could be used to test this hypothesis.

Successive waves of colonization by species that had evolved in allopatry in the Indian Ocean is another potential explanation for the co-occurrence of reproductively isolated cryptic species of *Ovabunda* in the Red Sea. Little is currently known, however, about the distribution and diversity of *Ovabunda* outside of the Red Sea, and the lack of congruence between morphological characters and clades hinders any attempt to confirm

the identity of the few species that have been reported from the Indian Ocean. The latter include *O. impulsatilla*, *O. hamsina* and *O. benayahui* (Reinicke, 1995) in the Seychelles (Janes 2008), and *O. hamsina* and a species dubiously identified as *O. macrospiculata* (Halász et al. 2013) in Madagascar (Verseveldt 1971; Janes 2008). No specimens corresponding morphologically to *O. benayahui* were collected in the present study, so we have no information about its clade affinities. Specimens of the other three morphospecies were found in clades A (*O. macrospiculata*), B (*O. hamsina*, *O. impulsatilla*) and C (*O. impulsatilla*, *O. macrospiculata*), allowing for the possibility that all three of those clades occur in the Indian Ocean as well as in the Red Sea. Alternatively, Indian Ocean populations could represent one or more different species not found in the Red Sea, as seems to be the case for *O. andamanensis*, a species described recently from Thailand (Janes et al. 2014). Indeed, multilocus barcode genotypes (*mtMutS+COI+28S*) from specimens collected recently in Madagascar (identified morphologically as *O. gohari*, *O. impulsatilla* and *O. verseveldti*) represent genetic clades that are distinct from those documented to occur in the Red Sea (McFadden and Benayahu, unpubl. data). Additional genetic sampling of *Ovabunda* from throughout the Indian Ocean reef system will be necessary before phylogenetic relationships among Indian Ocean and Red Sea clades can be used to test the hypothesis that speciation occurred in the Indian Ocean prior to *Ovabunda* species colonizing the Red Sea.

5. Conclusions

Subsequent to the confirmation of species boundaries using molecular tools, it is usually possible to distinguish cryptic species based on geographic range (if they are

allopatric) or because morphological or ecological characteristics that had previously been interpreted as intra- or sub-specific variation are now recognized to be species-diagnostic. The *Ovabunda* cryptic species complex reported here is unusual in that all of its currently recognized members are sympatric, and appear to overlap extensively in habitat use, ecological characteristics and morphological variation. With the exception of polyp pulsation in *O. macrospiculata* (species A), no diagnostic morphological differences distinguish them. These cryptic species are also not reliably distinguished using conventional single-locus mitochondrial barcode markers such as *COI* or the *mtMutS* gene that has been more widely used in octocorals (McFadden et al. 2010). For now, species B, C and D can be distinguished only by diagnostic pure and compound characteristic attributes found in the *28S* and *ITS rDNA* sequences.

Inability to distinguish cryptic species of *Ovabunda* in the field hinders any inferences that can be made about their population dynamics and biological or ecological interactions. Past studies that have documented the rapid reproduction and competitive dominance of *O. macrospiculata* on northern Red Sea reefs (Benayahu & Loya 1981, 1984a, b, 1985) may in fact have documented the combined life history attributes of four evolutionarily separate lineages. Further ecological and reproductive studies coupled with genotyping of colonies to identify clade (e.g. Halász et al. 2014) will be necessary to determine if there are any significant functional differences between them and to shed further light on the evolution and maintenance of possible reproductive isolating mechanisms in this cryptic species complex.

Acknowledgments

The authors wish to thank A. Lee-Fong, B. Bowen, G. Concepcion, I.S.S. Knapp, A. Eggers, and the participants of the 2013 Edwin W. Pauley Summer Program for their discussions and assistance with this project; A. Shlagman for curatorial skills; and the Interuniversity Institute for Marine Sciences in Eilat (IUI) for facilities. Collection of animals in Eilat complied with a permit issued by the Israel Nature and National Parks Protection Authority. Funding for this work came in part from the US-Israeli Binational Science Foundation (2008186), the National Science Foundation (OCE#12-60169 and OCE-0929031), and the Pauley Foundation. Part of this research (Applications DE-TAF-662, AT TAF 2064, and GB TAF 3027) received support from the SYNTHESYS Project (<http://www.synthesys.info/>) which is financed by European Community Research Infrastructure Action under the FP7 "Capacities" Program. It also was in part supported by a Temminck Fellowship, the Naturalis Biodiversity Center, and The Israel Cohen Chair in Environmental Zoology to YB. ZHF would also like to thank the Seaver Foundation for providing support. AMR was supported by Howard Hughes Medical Institute Undergraduate Science Education Award 52006301 to Harvey Mudd College. RHY received partial support from the Graduate Student Organization at the University of Hawai'i at Manoa. None of these funding agencies played any role in the study design; collection, interpretation or analysis of data; or writing of this report. The comments of two anonymous reviewers helped improve the manuscript.

Appendix A

Supplementary material includes a complete list of all specimens and associated sequences analyzed here; individual gene trees and haploweb diagrams for polymorphic nuclear loci; and a Structure plot for the results of species delimitation using RAD-tag sequencing.

References

- Aylor DL, Price EW, Carbone I (2006) SNAP: Combine and Map modules for multilocus population genetic analysis. *Bioinformatics* 22: 399-401.
- Benayahu Y (1991) Reproduction and developmental pathways of Red Sea Xenidiidae (Octocorallia, Alcyonacea). *Hydrobiologia* 216/217: 125-130.
- Benayahu Y, Loya Y (1981) Competition for space among coral-reef sessile organisms at Eilat, Red Sea. *Bulletin of Marine Science* 31: 514-522.
- Benayahu Y, Loya Y (1984a) Life history studies on the Red Sea soft coral *Xenia macrospiculata* Gohar, 1940. I. Annual dynamics of gonadal development. *Biological Bulletin* 166: 32-43.
- Benayahu Y, Loya Y (1984b) Life history studies on the Red Sea soft coral *Xenia macrospiculata* Gohar, 1940. II. Planulae shedding and post larval development. *Biological Bulletin* 166: 44-53.
- Benayahu Y, Loya Y (1985) Settlement and recruitment of a soft coral: why is *Xenia macrospiculata* a successful colonizer? *Bulletin of Marine Science* 36: 177-188.
- Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I (2007) Cryptic species as a window on diversity and conservation. *Trends in Ecology and Evolution* 22: 148-155.
- Bilewitch JP, Coates KA, Currie DC, Trapido-Rosenthal HG (2010) Molecular and morphological variation supports monotypy of the octocoral *Briareum* Blainville, 1830 (Octocorallia: Alcyonacea) in the Western Atlantic. *Proceedings of the Biological Society of Washington* 123: 93-112.

- Boulay JN, Hellberg ME, Cortés J, Baums IB (2014) Unrecognized coral species diversity masks differences in functional ecology. *Proceedings of the Royal Society B* 281:20131580.
- Carstens BC, Satler JD (2013) The carnivorous plant described as *Sarracenia alata* contains two cryptic species. *Biological Journal of the Linnean Society* 109: 737-746.
- Chhatre VE, Emerson KJ (2016) StrAuto: Automation and parallelization of STRUCTURE analysis. <http://strauto.popgen.org>
- 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: 323-326.
- Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods* 9(8): 772.
- DeSalle R, Egan MG, Siddall M (2005) The unholy trinity: taxonomy, species delimitation and DNA barcoding. *Philosophical Transactions of the Royal Society B* 360: 1905-1916.
- DiBattista JD, Choat JH, Gaither MR, Hobbs JA, Lozano-Cortés DF, Myers RF, Paulay G, Rocha LA, Toonen RJ, Westneat MW, Berumen ML (2016) On the origin of endemic species in the Red Sea. *Journal of Biogeography* 43: 13-30.
- Eaton DA (2014) PyRAD: assembly of de novo RADseq loci for phylogenetic analyses. *Bioinformatics* 30: 1844-1849. doi: 10.1093/bioinformatics/btu121.

- Flot J-F, Couloux A, Tillier S (2010) Haplowebs as a graphical tool for delimiting species: a revival of Doyle's "field for recombination" approach and its application to the coral genus *Pocillopora* in Clipperton. *BMC Evolutionary Biology* 10:372.
- Forsman ZH, Barshis DJ, Hunter CL, Toonen RJ (2009) Shape-shifting corals: molecular markers show morphology is evolutionarily plastic in *Porites*. *BMC Evolutionary Biology* 9:45.
- Forsman ZH, Concepcion GT, Haverkort RD, Shaw RW, Maragos JE, Toonen RJ (2010) Ecomorph or endangered coral? DNA and microstructure reveal Hawaiian species complexes: *Montipora dilata/flabellata/turgescens* & *M. patula/verrilli*. *Plos One* 5(12): e15021.
- Fujita MK, Leaché AD, Burbrink FT, McGuire JA, Moritz C (2012) Coalescent-based species delimitation in an integrated taxonomy. *Trends in Ecology and Evolution* 27: 480-488.
- Guindon S, Gascuel O (2003) A simple, fast and accurate method to estimate large phylogenies by maximum-likelihood. *Systematic Biology* 52: 696-704.
- Gutiérrez-Rodríguez C, Barbeitos MS, Sánchez JA, Lasker HR (2009) Phylogeography and morphological variation of the branching coral *Pseudopterogorgia elisabethae*. *Molecular Phylogenetics and Evolution* 50: 1-15.
- Halász A, McFadden CS, Aharonovich D, Toonen RJ, Benayahu Y (2013) A revision of the octocoral genus *Ovabunda* (Alderslade, 2001) (Anthozoa, Octocorallia, Xeniidae). *ZooKeys* 373: 1-41.
- Halász A, Reynolds AM, McFadden CS, Toonen RJ, Benayahu Y (2014) Could polyp

- pulsation be the key to species boundaries in the genus *Ovabunda* (Octocorallia: Alcyonacea: Xenidiidae)? *Hydrobiologia* doi:10.1007/s10750-014-2106-z
- Haverkort-Yeh R, McFadden CS, Halász A, Berumen M, Benayahu Y, Toonen RJ. (2013) A taxonomic survey of Saudi Arabian Red Sea octocorals (Cnidaria: Alcyonacea). *Marine Biodiversity* 43: 279-291.
- Heled J, Drummond AJ (2010) Bayesian inference of species trees from multilocus data. *Molecular Biology and Evolution* 27: 570-580.
- Hellberg ME, Prada C, Tan MH, Forsman ZH, Baums IB (2016) Getting a grip at the edge: recolonization and introgression in eastern Pacific *Porites* corals. *Journal of Biogeography* 43: 2147-2159.
- Herrera S, Shank TM (2016) RAD sequencing enables unprecedented phylogenetic resolution and objective species delimitation in recalcitrant divergent taxa. *Molecular Phylogenetics and Evolution* 100: 70-79.
- Huang D, Meier R, Todd PA, Chou LM (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.
- Huelsenbeck JP, Andolfatto P (2007) Inference of population structure under a Dirichlet process model. *Genetics* 175: 1787-1802.
- Huelsenbeck JP, Andolfatto P, Huelsenbeck ET (2011) Structurama: Bayesian inference of population structure. *Evolutionary Bioinformatics* 7: 55-59.
- Janes MP (2008) A study of the Xenidiidae (Octocorallia, Alcyonacea) collected on the "Tyro" expedition to the Seychelles with a description of a new genus and species. *Zoologische Mededelingen Leiden* 82 (49): 599-626.

- Janes MP, McFadden CS, Chanmethakul T (2014) A new species of *Ovabunda* (Octocorallia: Xeniidae) from the Andaman Sea, Thailand, with notes on the biogeography of this genus. *ZooKeys* 431: 1-17.
- Jarman SN, Ward RD, Elliott NG (2002) Oligonucleotide primers for PCR amplification of coelomate introns. *Marine Biotechnology* 4: 347-355.
- Kahng S, Benayahu Y, Lasker HR (2011) Sexual reproduction in octocorals. *Marine Ecology Progress Series* 443: 265-283.
- Katoh K, Kuma K, Toh, H, Miyata T (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research* 33, 511-513.
- Kim E, Lasker HR, Coffroth MA, Kim K (2004) Morphological and genetic variation across reef habitats in a broadcast-spawning octocoral. *Hydrobiologia* 530/531: 423-432.
- Knapp ISS, Puritz JB, Bird CE, Whitney JL, Sudek M, Forsman ZH, Toonen RJ (2016) ezRAD- an accessible next-generation RAD sequencing protocol suitable for non-model organisms v3.1. doi:10.17504/protocols.io.e9pbh5n. In press.
- Knowlton N (2000) Molecular genetic analyses of species boundaries in the sea. *Hydrobiologia* 420: 73-90.
- Ladner JT, Palumbi SR (2012) Extensive sympatry, cryptic diversity and introgression throughout the geographic distribution of two coral species complexes. *Molecular Ecology* 21: 2224-2238.
- Leaché AD, Fujita MK (2010) Bayesian species delimitation in West African forest geckoes (*Hemidactylus fasciatus*). *Proceedings of the Royal Society B* 277: 3071-3077.

- Librado P, Rozas J (2009) DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.
- Lin S, Cheng S, Song B, Zhong X, Lin X, Li W, Li L, Zhang Y, Zhang H, Ji Z, Cai, M (2015) The *Symbiodinium kawagutii* genome illuminates dinoflagellate gene expression and coral symbiosis. *Science* 350: 691-694.
- Marti-Puig P, Forsman ZH, Haverkort RD, Maragos J, Toonen R (2013) Extreme phenotypic polymorphism in *Pocillopora* (Anthozoa: Scleractinia); corallite morphology corresponds to mitochondrial groups, while taxonomy based on colony morphology does not. *Bulletin of Marine Science* 90:1–22.
- McFadden CS (1999) Genetic and taxonomic relationships among Northeastern Atlantic and Mediterranean populations of the soft coral *Alcyonium coralloides*. *Marine Biology* 133: 171-184.
- McFadden CS, Benayahu Y, Pante E, Thoma JN, Nevarez PA, France SC (2011) Limitations of mitochondrial gene barcoding in Octocorallia. *Molecular Ecology Resources* 11: 19-31.
- McFadden CS, Brown AS, Brayton C, Hunt CB, Ofwegen LP van (2014) Application of DNA barcoding to biodiversity studies of shallow-water octocorals: molecular proxies agree with morphological estimates of species richness in Palau. *Coral Reefs* 33: 275-286.
- McFadden CS, Grosberg RK, Cameron BB, Karlton DP, Secord D (1997) Genetic relationships within and between clonal and solitary forms of the sea anemone *Anthopleura elegantissima* revisited: evidence for the existence of two species. *Marine Biology* 128: 127-139.

- McFadden CS, Sánchez JA, France SC (2010) Molecular phylogenetic insights into the evolution of Octocorallia: a review. *Integrative and Comparative Biology* 50: 389-410.
- Neigel J, Domingo A, Stake J (2007) DNA barcoding as a tool for coral reef conservation. *Coral Reefs* 26:487–499.
- Pante E, Abdelkrim J, Viricel A, Gey D, France SC, Boisselier MC, Samadi S (2015). Use of RAD sequencing for delimiting species. *Heredity* 114: 450-459.
- Paz-García DA, Hellberg ME, García-de-León FJ, Balart EF (2015) Switch between morphospecies of *Pocillopora* corals. *American Naturalist* 186: 434-440. doi: 10.1086/682363
- Pella J, Masuda MM (2006) The Gibbs and split-merge sampler for population mixture analysis from genetic data with incomplete baselines. *Canadian Journal of Fisheries and Aquatic Science* 63: 576-596.
- Piggott MP, Chao NL, Beheregaray LB (2011) Three fishes in one: cryptic species in an Amazonian floodplain forest specialist. *Biological Journal of the Linnean Society* 102: 391-403.
- Prada C, Schizas NV, Yoshioka PM (2008) Phenotypic plasticity or speciation? A case from a clonal marine organism. *BMC Evolutionary Biology* 8: 47.
- Prada C, DeBiasse MB, Neigel JE, Yednock B, Stake JL, Forsman ZH, Baums IB, Hellberg ME (2014) Genetic species delineation among branching Caribbean *Porites* corals. *Coral Reefs* 33:1019–1030.

- Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155: 945-959.
- de Queiros K (1998) The general lineage concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations, in: Howard DJ, Berlocher SH (Eds.), *Endless Forms: Species and Speciation*, Oxford University Press, Oxford, England, pp. 57-75.
- Rambaut A, Drummond AJ (2007) Tracer v. 1.4. <http://tree.bio.ed.ac.uk/software/tracer/>
- Rannala B, Yang Z (2003) Bayes estimation of species divergence times and ancestral population sizes using DNA sequences from multiple loci. *Genetics* 164: 1645-1656.
- Reitzel AM, Herrera S, Layden MJ, Martindale MQ, Shank TM (2013) Going where traditional markers have not gone before: utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. *Molecular Ecology* 22: 2953–2970.
- Rittmeyer EN, Austin CC (2012) The effects of sampling on delimiting species from multi-locus sequence data. *Molecular Phylogenetics and Evolution* 65: 451-463.
- Ronquist F, Teslenko M, van der Mark P, Ayres D, Darling A, Höhna S, Larget B, Liu L, Suchard MA, Huelsenbeck JP (2012) MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology* 61: 539-542. doi:10.1093/sysbio/sys029
- Schmieder R, Edwards R (2011) Fast identification and removal of sequence contamination from genomic and metagenomic datasets. *Plos One* 6(3): e17288.
- Shaffer HB, Thomson RC (2007) Delimiting species in recent radiations. *Systematic*

- Biology 56:896–906
- Shearer TL, Coffroth MA (2008) Barcoding corals: limited by interspecific divergence not intraspecific variation. *Molecular Ecology Resources* 8: 247-255.
- Stat M, Baker AC, Bourne DG, Correa AMS, Forsman Z, Huggett MJ, Pochon X, Skillings D, Toonen RJ, van Oppen MJH, Gates RD (2012) Molecular delineation of species in the coral holobiont. *Advances in Marine Biology* 63:1–65.
- Stephens M, Smith NJ, Donnelly P (2001) A new statistical method for haplotype reconstruction from population data. *American Journal of Human Genetics* 68: 978-989.
- Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S (2011) MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution* 28: 2731–2739, doi: 10.1093/molbev/msr121
- Toonen RJ, Puritz JB, Forsman ZH, Whitney JL, Fernandez-Silva I, Andrews KR, Bird CE (2013) ezRAD: a simplified method for genomic genotyping in non-model organisms. *PeerJ*, doi 10:7717/peerj.203
- Verseveldt J (1971) Octocorallia from North-Western Madagascar (Part II). *Zoologische Verhandelingen* 117: 1-73.
- Yang Z, Rannala B (2010) Bayesian species delimitation using multilocus sequence data. *Proceedings of the National Academy of Science USA* 107: 9264-9269.
- Zhang J, Kobert K, Flouri T, Stamatakis A (2014) PEAR: A fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics* 30(5): 614-620.

Zwickl DJ (2006) Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion. Ph.D. dissertation, University of Texas: Austin.

Table 1: Nominal and unidentified morphospecies of *Ovabunda* as discriminated by three primary taxonomic characters: number of rows of pinnules on the tentacles, numbers of pinnules per row, and polyp pulsation (Halász et al. 2013, 2014). # Seq = number of individuals of each morphospecies included in the present analysis. Clade = phylogenetic cluster to which specimens of each morphospecies belonged (see Fig. 2). *O. benayahui* and *O. andamanensis* were not included in the present study.

Morphospecies	Pinnules				Clade
	# Rows	# per Row	Polyp Puls.	# Seq.	
<i>O. benayahui</i> (Reinicke, 1995)	1	6-7	?	--	--
<i>Ovabunda</i> sp. 1	1	7-15	N	3	C
<i>O. verseveldti</i> (Benayahu, 1990)	1	15-19	N	1	C
<i>O. gohari</i> (Reinicke, 1997)	1	18-21	N	2	B,C
<i>Ovabunda</i> sp. 2	1	22-24	N	1	C
<i>O. andamanensis</i> Janes et al. 2014	1-2	13-19	N	--	--
<i>Ovabunda</i> sp. 3	2	8-11	Y	1	A
<i>O. impulsatilla</i> (Verseveldt & Cohen, 1971)	2	8-13	N	9	B,C
<i>Ovabunda</i> sp. 4	2	10-16	Y	1	A
<i>O. biseriata</i> (Verseveldt & Cohen, 1971)	2	10-18	N	24	A,B,C,D
<i>O. faraunensis</i> (Verseveldt & Cohen, 1971)	2	15-24	N	9	B,C
<i>Ovabunda</i> sp. 6	2	19-23	Y	1	A
<i>O. arabica</i> (Reinicke, 1995)	2	22-26	N	2	C
<i>O. crenata</i> (Reinicke, 1997)	2-3	13-17	N	4	B
<i>Ovabunda</i> sp. 5	2-3	14-20	Y	3	A
<i>O. macrospiculata</i> (Gohar, 1940)	3	12-18	Y	11	A,C
<i>O. ainex</i> (Reinicke, 1997)	3	13-22	N	12	B,C,D
<i>Ovabunda</i> sp. 7	3	15-25	Y	2	A
<i>O. hamsina</i> (Reinicke, 1997)	3	16-22	N	2	B

Table 2: Results of species delimitation analyses using Structurama. Number of clusters

(k) was treated as a random variable with a Dirichlet process prior (α) estimated by a

gamma probability distribution for which shape and scale parameters were varied.

Boldface highlights k with the highest probability under each parameter set. Runs with

28S or ITS excluded used gamma probability distribution $\alpha(0.1,10)$.

gamma probability distribution									
	k	$\alpha(1,1)$	$\alpha(1,5)$	$\alpha(1,10)$	$\alpha(0.1,1)$	$\alpha(0.1,5)$	$\alpha(0.1,10)$	no 28S	no ITS
marginal likelihood		-1215.84	-1215.73	-1214.39	-1214.64	-1214.03	-1181.34	-593.98	-95.62
probability (# pops)	1	0	0	0	0	0	0	0	0
	2	0	0	0	0	0	0	0	0
	3	0	0	0	0	0	0	0	0
	4	0.03	0.07	0.13	0.05	0.1	0.16	0.53	0.49
	5	0.46	0.62	0.67	0.53	0.65	0.68	0.33	0.45
	6	0.36	0.26	0.18	0.31	0.22	0.15	0.1	0.06
	7	0.12	0.05	0.02	0.09	0.03	0.01	0.03	0
	8	0.03	0	0	0.02	0	0	0	0

Table 3: Genetic distances (Kimura 2-parameter, expressed as %) among individuals within and among clades of *Ovabunda* at the 7 loci sequenced in this study. Upper value in cell is mean $\pm$ s.d., lower value is maximum value observed. Within clade: mean $\pm$ s.d., range and median of all pairwise genetic distances between individuals belonging to the same clade (i.e. intraspecific comparisons); between clade: mean $\pm$ s.d., range and median of all pairwise genetic distances between individuals in different clades (i.e. interspecific comparisons).

Species	Locus						
	<i>ND2</i>	<i>mtMutS</i>	<i>COI</i>	<i>28S rDNA</i>	<i>ITS rDNA</i>	<i>ATPSα</i>	<i>ATPSβ</i>
Clade A	0.04 $\pm$ 0.07 0.28	0.13 $\pm$ 0.23 0.69	0.08 $\pm$ 0.13 0.39	0.07 $\pm$ 0.21 0.71	0.99 $\pm$ 0.73 2.28	4.01 $\pm$ 2.69 9.59	1.99 $\pm$ 1.37 6.05
Clade B	0.00 $\pm$ 0.00 0.00	0.09 $\pm$ 0.09 0.28	0.08 $\pm$ 0.08 0.27	0.01 $\pm$ 0.03 0.14	0.25 $\pm$ 0.13 0.64	2.22 $\pm$ 1.33 4.48	0.37 $\pm$ 0.32 1.13
Clade C	0.03 $\pm$ 0.06 0.28	0.05 $\pm$ 0.09 0.41	0.00 $\pm$ 0.00 0.00	0.13 $\pm$ 0.13 0.43	0.42 $\pm$ 0.22 1.08	2.09 $\pm$ 1.13 5.35	1.43 $\pm$ 1.55 4.49
Clade D	0.00 $\pm$ na 0.00	na	0.13 $\pm$ na 0.13	0.29 $\pm$ 0.12 0.58	0.32 $\pm$ na 0.32	0.00 $\pm$ na 0.00	0.00 $\pm$ na 0.00
<i>Xenia</i>	0.00 $\pm$ 0.00 0.00	0.13 $\pm$ 0.23 0.55	0.02 $\pm$ 0.05 0.13	0.14 $\pm$ 0.32 0.87	0.96 $\pm$ 1.18 3.06	1.17 $\pm$ 1.64 5.30	2.05 $\pm$ 1.83 4.66
Mean Within Clades	0.02 $\pm$ 0.05	0.09 $\pm$ 0.14	0.05 $\pm$ 0.09	0.07 $\pm$ 0.15	0.50 $\pm$ 0.47	2.41 $\pm$ 1.80	1.18 $\pm$ 1.35
Mean Between Clades	0.07 $\pm$ 0.08	0.40 $\pm$ 0.19	0.23 $\pm$ 0.18	0.49 $\pm$ 0.25	1.77 $\pm$ 0.91	3.46 $\pm$ 1.50	3.01 $\pm$ 1.39
Range, Median Within Clades Range,	0.00–0.28 0.00	0.00–0.69 0.00	0.00–0.39 0.00	0.00–0.87 0.00	0.00–3.06 0.32	0.00–9.59 2.31	0.00–6.05 0.56

Median	0.00–0.42	0.00–0.69	0.00–0.55	0.00–1.01	0.00–3.97	0.00–9.14	0.00–6.11
Between	0.00	0.14	0.13	0.57	1.51	3.41	3.43
Clades							

Figure Legends

Fig. 1: Map of the Red Sea showing collection sites and numbers of individuals of each phylogenetic clade collected at each site. Inset: Northern Gulf of Aqaba. Site abbreviations, Saudi Arabia: AL, Al Lith; JE, Jeddah; TH, Thuwal. Israel: IUI, Interuniversity Institute for Marine Sciences; UO, Underwater Observatory; NR, Nature Reserve; OJ, Oil Jetty; DD, Dry Dock; UR, Underwater Restaurant; PY, Pyramid.

Fig. 2: Maximum likelihood tree of *Ovabunda* and *Xenia* from Israel and Saudi Arabia based on partitioned analysis of three mitochondrial and four nuclear loci. Support values above nodes are bootstrap percentages from ML; values below nodes are Bayesian posterior probabilities. *: individuals with mixed-clade genotypes (see text).

Fig 3: Maximum clade credibility tree from *BEAST analysis used as a guide tree for Bayesian species delimitation with BPP. Numbers above nodes are posterior probabilities from *BEAST, separate analyses of alternative alleles for each individual. Numbers below nodes are posterior probabilities of ancestral nodes estimated using BPP.

Fig. 4: Characteristic attributes distinguishing clades of *Ovabunda* (A–D) and *Xenia* (X). Each column represents a nucleotide position in one of 7 genes sequenced: N: *ND2*, C: *COI*, M: *mtMutS*, R: *28S rDNA*, I: *ITS*, B: *ATPS β* , A: *ATPS α* . Vertical numbers indicate nucleotide position within a sequence (for example, N600 is the nucleotide at position 600 in the sequence alignment for *ND2*). d = deletion. Gray indicates positions at which clades B and C share a character that distinguishes them from other clades.

Fig. 5: Number of specimens of each clade of *Ovabunda* (A–D) and *Xenia* (X) collected at 10 m depth intervals. Data are pooled across 10 collection sites in Israel (northern Gulf of Aqaba) and 3 locations in Saudi Arabia.

Appendix A

Table A.1: Specimens of *Ovabunda* and *Xenia* used for phylogenetic and species delimitation analyses. Specimens that were also included in the pooled RADseq analysis indicated in boldface.

Fig. A.1: Maximum likelihood analysis of three concatenated mitochondrial loci (*mtMutS*, *COI*, *ND2*). Individuals are color-coded to reflect clade identity: clade A, red; clade B, green; clade C, light blue; clade D, dark blue; *Xenia* clade, yellow.

Fig. A.2: Maximum likelihood analysis of *28S rDNA*. Individuals are color-coded to reflect clade identity: clade A, red; clade B, green; clade C, light blue; clade D, dark blue; *Xenia* clade, yellow.

Fig. A.3: Maximum likelihood analysis of *ITS rDNA*. Haplowebs (dashed lines) connect alleles found in the same heterozygous individual. Individuals are color-coded to reflect clade identity: clade A, red; clade B, green; clade C, light blue; clade D, dark blue; *Xenia* clade, yellow.

Fig. A.4: Maximum likelihood analysis of *ATPS β* . Haplowebs (dashed lines) connect alleles found in the same heterozygous individual. Individuals are color-coded to

reflect clade identity: clade A, red; clade B, green; clade C, light blue; clade D, dark blue; *Xenia* clade, yellow.

Fig. A.5: Maximum likelihood analysis of *ATPS α* . Haplowebs (dashed lines) connect alleles found in the same heterozygous individual. Individuals are color-coded to reflect clade identity: clade A, red; clade B, green; clade C, light blue; clade D, dark blue; *Xenia* clade, yellow.

Fig. A.6: Maximum likelihood analysis of *SRP54*. Individuals are color-coded to reflect clade identity: clade A, red; clade B, green; clade C, light blue; clade D, dark blue; *Xenia* clade, yellow.

Fig. A.7: Structure plot from analysis of 1332 unlinked SNPs sequenced using RADseq. Individuals of *Ovabunda* clades A, B and C were pooled prior to library construction.

	N	C	C	C	C	M	M	M	M	M	M	R	R	R	R	R	R	I	I	 
--	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---	---

mt: mtMutS + COI + ND2

28S rDNA

ITS

ATPS α

SRP54