Nuclear sequences reveal mid-range isolation of an imperilled deep-water coral population

RON I. EYTAN,* MARSHALL HAYES,† PATRICIA ARBOUR-REILY,* MARGARET MILLER‡ and MICHAEL E. HELLBERG*

Department of Biological Sciences, Louisiana State University, Baton Rouge, LA 70803, USA, †Department of Plant Pathology and Plant-Microbe Biology, Cornell University, Ithaca, NY 14853, USA, ‡NOAA Fisheries, Southeast Science Center, 75 Virginia Beach Dr, Miami, FL 33149 USA

Abstract

The mitochondrial DNA of corals and their anthozoan kin evolves slowly, with substitution rates about two orders of magnitude lower than in typical bilateral animals. This has impeded the delineation of closely related species and isolated populations in corals, compounding problems caused by high morphological plasticity. Here we characterize rates of divergence and levels of variation for three nuclear gene regions, then use these nuclear sequences as markers to test for population structure in *Oculina*, a taxonomically confused genus of corals. Rates of sequence divergence (obtained by comparison to *Solenastrea hyades*) were at least five (and sometimes over 10) times faster for the three nuclear markers than for a mitochondrial reference sequence. Nuclear sequence variation was also high within populations, although it tended to decline north of Cape Canaveral. Significant subdivision was evident among samples from 10 locations from between North Carolina and the Florida Panhandle, but neither nominal species designation nor population depth explained much of this variation. Instead, a single population from the unique deep (> 70 m) water reefs at the Oculina Banks off central Florida was a strong genetic outlier: all pairwise measures of subdivision involving this population were greater than those involving all other populations, and multilocus clustering recognized the Oculina Banks as distinct from other populations, despite its close proximity (≤ 36 km) to populations from shallower waters nearby and its location at the centre of the sampled range. Genetic isolation of the Oculina Banks population suggests that focused efforts will be needed to conserve the foundation species of these monotypic reefs and that depth may play a role in isolating marine populations and perhaps facilitating initial steps towards speciation.

Keywords: multilocus genotyping, Oculina Banks, *Oculina*, population isolation

Received 19 December 2008; accepted 22 March 2009

Introduction

Appearances can be a misleading guide to distinguishing closely related, but genetically isolated species. The species delineation problem is especially difficult for taxa with simplified morphologies (e.g. cave salamanders, Niemiller *et al.* 2008; earthworms, King *et al.* 2008; parasitic acanthocephalans, Steinauer *et al.* 2007; sponges Klautau *et al.* 2003), especially when combined with high levels of

Correspondence: Michael E. Hellberg, Fax: (225)578 2597; E-mail: mhellbe@lsu.edu

phenotypic plasticity (e.g. sponges, Erwin & Thacker 2007; octocorals, Prada *et al.* 2008; freshwater mussels, Baker *et al.* 2003). Despite these difficulties, species delineation is a necessary component of understanding the speciation process, of characterizing ecological variation among habitats and geographical locales, and of targeting imperilled species and populations for conservation efforts.

DNA sequences offer a large potential pool of characters for inferring species boundaries and relationships. Until recently, genetic studies aimed at understanding differentiation near the species interface have used primarily mitochondrial DNA (mtDNA) sequences (at least for bilateral

animals). One region of this genome (*coxI*) has even been nominated for use as a universal genetic barcode (Hebert *et al*. 2003). Mitochondrial DNA has several advantages relative to nuclear DNA in these contexts: high rates of nucleotide substitution produce abundant identifiable variants and these variants sort to reciprocal monophyly quickly due to low effective population size. However, the lack of recombination means that all mtDNA sequences from an individual share a single history that may not necessarily reflect species relationships, as demonstrated by recent examples of phylogenies based on multiple nuclear loci conflicting with mtDNA-based trees (Leache & McGuire 2006; Carling & Brumfield 2008). Moreover, and more germane for the issue of species delineation, rapid rates of mtDNA change cannot be assumed for a substantial proportion of eukaryotes.

For plants, fungi, and many lower metazoans, rates of nucleotide substitution for mtDNA are about 100-times slower than those in bilateral animals (Wolfe *et al*. 1987; Hellberg 2006). Mitochondrial DNA thus will have far less power to reveal phylogeographical structure and recognize cryptic species in eukaryotes with slow mtDNA than in bilateral animals (Huang *et al*. 2008). Nuclear gene sequences may provide an alternative to mtDNA. Rates of nucleotide substitution for nuclear DNA (nDNA) in plants and fungi are not exceptionally slow: they are similar to those in bilateral animals, or about 10 times faster than those for mtDNA from the same taxa (Wolfe *et al*. 1987). Allozyme surveys in taxa with slow mtDNA have revealed ample variation and subdivision (e.g. Hellberg 1994; McFadden *et al*. 1997), suggesting nuclear variation is not constrained as mtDNA is. Thus, nDNA sequences may succeed in flagging isolated populations in lineages where slowly evolving mtDNA cannot.

Hard (scleractinian) corals are perhaps the animal group most sorely in need of an alternative to mtDNA for delineating species and recognizing isolated populations (Lopez *et al*. 1999; Ridgway & Gates 2006). Several aspects of the biology of reef corals make the identification of recently diverged populations especially difficult. First, scleractinian morphology is highly plastic, both at the level of entire colonies and single corallites (Foster 1979; Willis 1985; Bruno & Edmunds 1997; reviewed in Todd 2008), which has led to extreme confusion in their taxonomy and systematics (Fukami *et al*. 2004, 2008). Second, coral mtDNA evolves too slowly to distinguish some close relatives (Shearer *et al*. 2002; Hellberg 2006); even comparisons of entire mtDNA genomes reveal only a small numbers of changes (e.g. in *Montastrea*, Fukami & Knowlton 2005; *Pocillopora*, Flot & Tillier 2007). Third, the potential for interspecific hybridization appears great in some reef corals, sometimes producing morphological intermediates (van Oppen *et al*. 2001; Marquez *et al*. 2002; Combosch *et al*. 2008; reviewed in Willis *et al*. 2006). Many reef corals face threats to their existence (Hoegh-Guldberg *et al*. 2007), therefore despite the aforementioned challenges, the identification of genetically isolated corals is critical to rational conservation efforts aimed at repopulation and the maintenance of genetic diversity (Knowlton 2001; Baums 2008).

Corals of the genus *Oculina* exemplify how the ability to distinguish isolated populations matters. Several nominal species of *Oculina* occur in coastal North American waters, generally living in waters that are cooler and more turbid than other tropical stony corals can tolerate. Members of this genus are gonochoric broadcast spawners (Brooke & Young 2003), and their larvae recruit well on to artificial hard substrate, especially when algal competitors are absent (Miller & Hay 1996). The ecological adaptability of shallow water *Oculina* is further supported by reports from the Mediterranean, where *Oculina patagonica* appears to have invaded over the last four decades (Fine *et al*. 2001). Such hardiness, however, does not mean all populations of *Oculina* are immune to anthropogenic disturbance. Off the southeastern coast of the USA, *Oculina varicosa* occurs as small (< 30 cm) facultatively zooxanthellate colonies at depths shallower than 30 m. In addition, off the eastern coast of central Florida, nominal populations of *O. varicosa* have formed extensive bioherms of unconsolidated coral rubble and sediment, capped with large colonies (~1–2 m) of living coral in deep (70–100 m) water. These deep-water colonies have more slender branches than shallow colonies and are azooxanthellate. These bioherms, collectively termed the Oculina Banks, have been heavily damaged by illegal trawling and dredging (Reed *et al*. 2007), despite Federal Protection that was initiated in 1984 (Reed 2002). Recovery of the framework species of this unique habitat depends critically on whether the *Oculina* populations at the Oculina Banks are isolated from other *Oculina* populations: ample recruits could be transplanted from shallow populations if they are genetically homogeneous, whereas distinct Oculina Banks populations would minimally require locally targeted recovery strategies, and more broadly warrant greater efforts to conserve the unique habitat they create.

Here we use single-copy nuclear DNA sequences to distinguish genetically isolated populations among continental North American populations of the coral genus *Oculina*. We first compare levels of divergence for three nDNA markers to that for a commonly used region of mtDNA, cytochrome oxidase I. Next, we assess levels of variation in these nDNA markers among four North American *Oculina* nominal species (*O. arbuscula, O. diffusa, O. robusta,* and *O. varicosa*). Finally, we use the nDNA markers to assess differentiation and genetic isolation among named morphospecies of *Oculina*, among geographically distant sites, and among populations found at different depths.
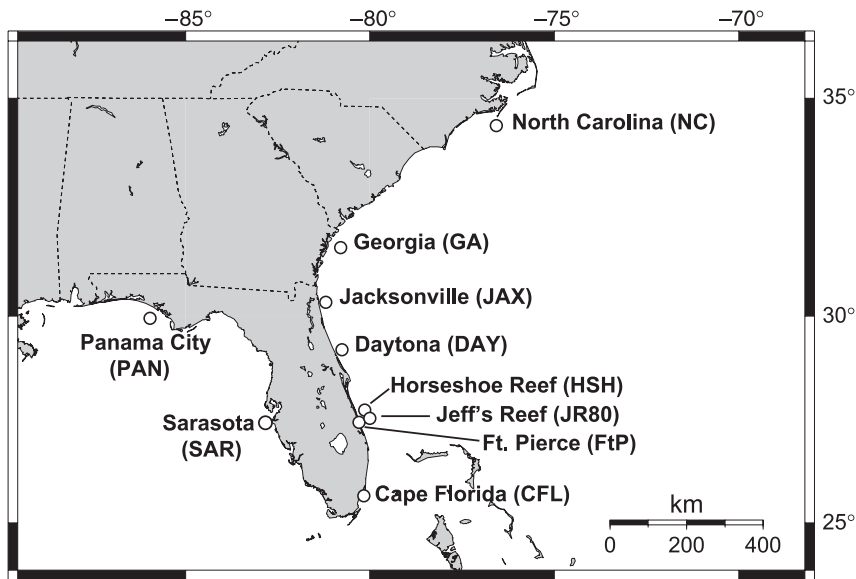
Fig. 1 Geographical locations from which samples of *Oculina* were taken.

Methods

Population sampling

We obtained samples of the four most common nominal species of *Oculina* from the southeastern USA (Fig. 1, Appendix), trying to sample each species from as broad a geographical and bathymetric range as possible. Sample sizes ranged from 8 to 16 colonies per location.

Species identifications were based on colony form, branch thickness, and corallite form. The genus *Oculina* has long been recognized as taxonomically challenging, with original descriptions that are often very sparse on details and virtually every species-level treatment calling for revision going back over 100 years (e.g. Verrill 1902; Zlatarski & Martinez Estalella 1982; Cairns 1991). We based our identifications of *Oculina diffusa* (Lamarck 1816), *O. robusta* (Pourtalès 1871), and *O. varicosa* (LeSueur 1821) on both their original descriptions and on subsequent work and guides (Verrill 1902; Zlatarski & Martinez Estalella 1982; Humann 1993). *Oculina diffusa* has short thin branches and its corallites are clearly raised. Colonies of *O. robusta* are, as the name suggests, more robust, with long thick branches that taper and corallites that are nearly flush with the branch. The branches of *O. varicosa* fall between these extremes: they are generally sturdier than those of *O. diffusa* and extend further between branch points. The corallites of *O. varicosa* extend from a swollen base (Verrill 1902). Individuals from North Carolina and Georgia were designated *Oculina arbuscula* (Agassiz in Verrill 1864; Rupert & Fox 1988), although this appears to be a regional moniker because no characters clearly separate them from *O. varicosa*.

RNA isolation, cDNA library construction and expressed sequence tag sequencing

We chose to use sequences from nuclear gene coding regions to evaluate rates of DNA evolution and patterns of population variability and subdivision. Microsatellite markers have been successfully used to identify regional population isolation in reef corals (Baums *et al*. 2005); however, patterns of nucleotide substitution around these hypervariable regions may be atypical (Stallings 1995; Vowles & Amos 2004). Primers that amplified single-copy nuclear genes that are sufficiently variable for population-level studies in cnidarians were not available when we started this study, so we generated expressed sequence tags (EST) to produce new markers.

RNA was isolated from a single live specimen of *Oculina varicosa* collected at Jeff's Reef, Florida. This deep water (80 m) individual was free of symbiotic algae (zooxanthellae), which might otherwise have contaminated coral tissue (Shearer *et al*. 2005). RNA was isolated through a procedure modified from Chomczyniski & Sacchi (1987). A tissue sample of a live individual was ground in ice cold GIT (4 M guanidine isothiocyanate, 25 mM NaOAc pH 6, 0.82% β-mercaptoethanol) in a Dounce homogenizer. Seven millilitres of the resulting homogenate was layered over 3 mLs of a CsCl cushion (5.7 M CsCl, 25 mM NaOAc pH 6.0) and centrifuged at 115 000 g in a Beckman SW41 rotor 16 h at 20 C. The resulting RNA pellet was resuspended in 150 L RNase-free 0.1 M EDTA. To concentrate this RNA, 1/10 V of RNase-free 5 M ammonium acetate, 5 g RNase-free glycogen and 2.5 V 100% EtOH were added and RNA was precipitated at –20 C overnight. RNA was spun down at 10 000 g at 4 C for 15 min,

Marker	Primers
p14	(Fatty acid elongase)
	Ocp14F: TGTACCACTTGGGATGAACG
	Ocp14R: TCAAGCTTCCAGTCTTGTGAAA
p62	(Elongation factor 1α)
	p62Fb: TGATTGTCCTCAACCATCCA
	p62R: CTCCTGACAGACTTTCGATGG
	p62Rd: ACCACCTTTCTGGGCTTTCT
p302	(Tachylectin-2 motif)
	p302F: TTATACGGCGTCACAAACGA
	p302R: TCGTCATCACCCTTTTATTCC
COI*	(Mitochondrial cytochrome oxidase *c* subunit I)
	HCO2198: TAAACTTCAGGGTGACCAAAAAATCA
	LCO1490: GGTCAACAAATCATAAAGATATTGG

*Primers from Folmer *et al.* (1994).

resuspended in a few microlitres of RNase-free 0.1 M EDTA, then an aliquot was inspected on a gel for degradation. Total RNA from this procedure was accumulated and saved at –80 C. Poly A RNA was isolated from 30 g total RNA using Ambion's Poly(A)Purist-MAG kit. The ultimate mRNA yield was ~739 ng of mRNA.

A cDNA library was constructed from the *O. varicosa* mRNA using Strategene's ZAP-cDNA kit according to the manufacturer's instructions with two exceptions. The kit directions suggest starting with ≥ 1.5 g poly A mRNA. Because we had only half this amount, all reaction volumes were halved as well. To size fractionate cDNAs, we used Pharmacia's SizeSeptember 400 column rather than the Sepharose CL-2B column provided with the kit. The resulting primary library had over 250 000 pfu before amplification. We sequenced 91 random inserts, with an average size of 571 bp. 67 of these contained open reading frame (ORF) of at least 60 residues, and 50 of these returned significant matches to existing sequences (31 March 2008 search) using BLASTp.

DNA isolation and polymerase chain reaction

Genomic DNA was isolated from a small piece of coral using either cetyltrimethyl ammonium bromide extraction protocols (R. J. Toonen, unpublished, available online at http://www2.hawaii.edu/~toonen/files/MsatsV1.pdf) or the MoBio Ultra Clean Soil DNA Isolation Kit.

From the 50 ESTs with putative matches, 23 primers pairs were designed (using Primer 3, Rozen & Skaletsky 2000) to amplify regions 300–500 bp long that included both parts of the ORF and 3′ untranslated region (UTR). Of the 16 pairs that amplified a single band of the proper size (or larger), three were selected (Table 1) for use as markers based on consistency of amplification and sequencing, variation found in an initial screening of individuals from

the geographical extremes sampled, and single-copy status (based on finding ≤ 2 alleles in cloned heterozygotes). All three of these nuclear gene regions aligned with sequences from the closest animal for which genomic data are available, the anemone *Nematostella vectensis*. The closest BLASTp hits were to sequences from *N. vectensis* (p14, fatty acid elongase, Putnam *et al.* 2007), from the coral *Pocillopora damicornis* (p62, elongation factor 1α, Flot *et al.* 2008), or the coral *Montastraea faveolata* (p302, tachylectin-2, Schwarz *et al.* 2008). No introns were present in the three amplified gene regions.

Products for each of these nuclear markers were amplified from genomic DNA using the same polymerase chain reaction (PCR) profile consisting of an initial denaturation (94 C) for 3 min, initial annealing step (50 C) of 2 min, and initial elongation (at 72 C) of 2 min, followed by 35 cycles of 35 s at 94, 1 min at 50, and 1 min 15 s at 72. A final elongation at 72 for 10 min completed the profile.

Sequencing, alignment and phasing

PCR amplicons were directly sequenced (with ABI BigDye version 3.1) using both of the amplification primers (GenBank accession numbers FJ966395–FJ966875). Many individuals had indels that obscured complete reads in both directions. All of these indels occurred in portions of the amplified region lying in the 3′ UTR of the sequenced gene except for one 3-bp indel in the ORF of tachylectin-2. Sequences containing indels were cloned to resolve constituent allelic sequences using ¼ reactions of the Invitrogen TOPO TA Cloning Kit for Sequencing. Resulting colonies were screened by PCR with the primers M13For(–20) and M13Rev. Eight to 16 clones of the proper size (or more if needed to find both alleles) were sequenced using the M13 primers. The initial direct sequences were always used in determining allelic sequences from cloned DNA to avoid scoring any changes that resulted from errors introduced by the PCR. In total, 945 cloned sequences were generated to resolve all individuals heterozygous for indels.

The COI and EF-1α sequences contained no gaps and were aligned by eye. The fatty acid elongase and tachylectin-2 sequences contained numerous gaps. Most commonly used multiple alignment programs make use of an initial guide tree as a framework for determining the optimal alignment and the placement and length of gaps is determined by a particular set of parameters. However, an incorrect guide tree may introduce bias into the resulting alignment. To avoid this problem, we employed a Bayesian approach to multiple sequence alignment, implemented in BAli-Phy version 2.0.1 (Suchard & Redelings 2006) which does not condition on a single alignment estimate. BAli-Phy finds the multiple alignment with the highest posterior probability by estimating both the alignment and the topology simultaneously, using a Markov chain Monte

Carlo (MCMC) sampler. This approach is computationally intensive, thereby limiting the number of sequences that can be included in the analysis (Redelings & Suchard 2005).

We reduced the full data sets for fatty acid elongase and tachylectin-2 for input into BAli-Phy using a two-step process. First, sequences were grouped by their length and aligned using Muscle (Edgar 2004), implemented in Geneious version 3.6 (Drummond *et al*. 2007). Second, networks were then constructed for each of the different alignments in TCS version 1.21 (Clement *et al*. 2000). For each network, the most common sequence was used to represent all sequences of that length, except when a sequence was more than 10 mutational steps from the most common one, in which case it was represented individually.

Both fatty acid elongase and tachylectin-2 were analysed in BAli-Phy using the GTR substitution model and the default indel model. By default, the MCMC sampler in BAli-Phy collects information after each iteration and runs until stopped by the user. We chose when to stop by first determining when convergence had occurred through visual inspection of output using Tracer version 1.4 (Rambaut & Drummond 2007). After convergence, the Markov chain was then allowed to run until the effective sample size from the Markov chain was equal to or greater than 1000. To ensure that the Markov chain had truly converged, we repeated this process an additional three times, for a total of four independent runs. The final output from each run was separately analysed, with all the samples before convergence discarded as burn-in. The consensus alignment from the run with the highest posterior probability was used for subsequent analyses.

To resolve alleles from sequences with multiple heterozygous single nucleotide polymorphisms (SNP), we employed a Bayesian statistical method implemented in Phase version 2.1.1 (Stephens *et al*. 2001; Stephens & Donnelly 2003; Stephens & Scheet 2005). Each marker data set was split into two inputs for haplotype reconstruction, one containing only individuals from the JR-80 population and the other containing the rest of the data set. This was carried out because preliminary analyses indicated that individuals from the JR-80 population might not be freely interbreeding with those in populations at other depths, which would violate the assumptions of the coalescent model used in Phase (Stephens *et al*. 2001).

Input files were prepared as suggested in the Phase documentation, but with several modifications. First, all variable sites were used. Second, gaps were coded as a fifth allele. The '–d1' option, which specifies a parent-independent mutation model rather than a stepwise one for multi-allelic loci, was used in the analysis of any data set that contained at least one tri-allelic SNP (true for all markers, but not for the data sets with only JR-80 individuals). Alleles determined by cloning heterozygotes were used to create a known file. A default probability threshold of 0.9 was used for all runs. We performed 10 independent runs for each data set analysed, using different random number seeds. The goodness-of-fit measure for each independent run was then plotted and compared to check for consistency between runs.

After the initial Phase runs, all data sets contained some individuals with unresolved SNPs. We cloned a subset of these individuals to directly determine their haplotype phase. The direct haplotype observations were then added to the 'known' file and the data sets were re-analysed. This was carried out iteratively until the phase of all SNPs was recovered with > 0.9 probability or we ran out of tissue. Final data sets contained no more than three individuals for which the phase of a single SNP was not resolved to 0.9 (one for tachylectin-2, three for elongation factor 1α and fatty acid elongase).

After alignment and phasing of heterozygous SNPs, the final nuclear data set contained 122 individuals (244 alleles). The average number of nucleotide differences between haplotypes, k (Tajima 1983; equation A3), was calculated for each marker and population in DNAsp version 4.5.0.2 (Rozas *et al*. 2003), excluding sites with alignment gaps.

Interspecific sequence divergence

We determined relative rates of synonymous and nonsynonymous substitution between *Oculina* and *Solenastrea hyades*, a species for which a congener (*Solenastrea bournii*) has previously shown to be a close relative of *O. diffusa* (Fukami *et al*. 2004). These estimates were made using MEGA version 4.0.1 (Tamura *et al*. 2007) and took into account only the coding regions of the four markers. Appropriate substitution models for calculating genetic distances were chosen by jModelTest (Posada 2008): K80 + Γ for the three nuclear markers and JC for COI under the BIC criteria. For mitochondrial COI, three *Oculina* individuals, representing the unique haplotypes for the genus, and one *Solenastrea* sequence (individuals in this genus were invariant) were used for this comparison. For nuclear genes, between-species means were calculated from the full data set of 122 *Oculina* (with the four nominal species pooled) and two *Solenastrea* individuals. For the nuclear genes, N and S were calculated using the modified Nei–Gojobori method (Nei & Kumar 2000), which accounts for differences in the frequencies of transitions and transversions, because jModelTest had chosen the K80 substitution model. The standard Nei–Gojobori method (Nei & Gojobori 1986) was used to calculate N and S for COI to reflect the JC model selected.

Recombination

Recombination can create DNA sequences with different histories, a violation of the assumptions underlying most

coalescent analyses. We tested for recombination using a combination of haplotype network and population genetic analyses. Networks were constructed for each nuclear marker in TCS version 1.21 (Clement *et al.* 2000), with alignment gaps counted as missing data. Recombination events were inferred if reticulations were present (Crandall 1999). The four-gamete test for recombination (Hudson & Kaplan 1985) was implemented in IMgc (Woerner *et al.* 2007).

Haplotype networks (not shown) for all three markers contained multiple reticulations, consistent with recombination linking regions with different histories (Crandall 1999). All three nuclear gene regions also failed the four-gamete test for detection of recombination. For this reason, we chose to use an infinite allele model, with each unique haplotype scored as a unique allele for each of the three markers.

Population structure

Identical alleles were collapsed for subsequent analyses. For EF-1α, which did not contain indels, this was carried out using the online implementation of FaBox (Villesen 2007). FaBox ignores indels when collapsing sequences, however, and both fatty acid elongase and tachylectin-2 contained numerous indels. To preserve information from these indels, alignments for these two markers were collapsed using Map (Aylor *et al.* 2006), part of the Snap suite of tools for nucleotide analysis (Price & Carbone 2005).

Hierarchical genetic subdivision, as measured by Φ_{ST}, was analysed using an AMOVA framework (Excoffier *et al.* 1992; Michalakis & Excoffier 1996), implemented in GenoDive version 2.0b11 (Meirmans & Van Tienderen 2004). The categories used for the AMOVA were location, nominal species, and population depth. Three depth ranges were used: shallow (< 12 m), medium (between 20 and 35 m), and deep (> 70 m). Pairwise Φ_{ST} values among populations were also calculated in Genodive using an AMOVA, which for this purpose are exactly equivalent to Weir and Cockerham's θ (Weir & Cockerham 1984). The Φ_{ST} values were plotted against pairwise geographical distances among populations. These were calculated in Google Earth version 4.3.7284.3916 (beta) using the shortest nautical distance among populations. Analyses were repeated with and without the differentiated deep-water population from the Oculina Banks and designating samples from North Carolina and Georgia as either *Oculina arbuscula* or *O. varicosa*. Because high levels of variation within populations necessarily reduce measures of the proportion of variation partitioned among populations (see Hedrick 2005), Φ_{ST} measures were also estimated using a standardizing procedure (Miermans 2006) implemented by GenoDive.

Table 2 Nucleotide sequence divergence between *Oculina* and *Solenastrea*

	Uncorrected *P**	K80 + Γ*	d*S*†	d*N*†
FA elongase	0.0681	0.0752	0.194	0.0208
EF-1α	0.0803	0.109	0.197	0.0456
Tachylectin	0.0582	0.0654	0.070	0.0541
COI	0.0088	0.0089‡	0.025	0.0039

*Full sequence; †coding region only, modified Nei–Gojobori; ‡Jukes–Cantor.

In order to detect significantly differentiated populations (*k*) without the need to define populations a priori, we used the Bayesian clustering analysis implemented by Structure version 2.2 (Pritchard *et al.* 2000). We used the default (and more conservative) admixture model with uncorrelated allele frequencies. Although the default correlated allele model (Falush *et al.* 2003) implemented by Structure is more robust to departures from model assumptions than the uncorrelated model, the correlated model is also more prone to overestimates of *k* (and thus the inference of spurious clusters) than is the uncorrelated model (Pritchard *et al.* 2007). We performed 20 replicates runs for *k* values between 1 (no population differentiation) and 7 (a pragmatic maximum given the number of localities sampled and the relative homogeneity of populations north of Cape Canaveral). Each replicate was run for 10^6 iterations following an initial burn-in of 100 000 iterations. Best estimates of *k* were inferred using Structurama (Huelsenbeck & Andolfatto 2007), which explicitly estimates *k*. Three replicates were each run for 10 000 000 generations, sampling every 100.

Results

Rates of divergence for all three nuclear gene markers were substantially higher than for mitochondrial COI (Table 2). Divergence for the slowest of the three nuclear genes (tachylectin-2) was more than five times faster than for COI, whether rates were corrected for multiple hits or not and whether synonymous or nonsynonymous rates were considered. Corrected divergences for EF-1α were over 10 times greater than those for COI.

Within *Oculina*, levels of variation at the different markers paralleled those for divergence rates. Mitochondrial COI was nearly invariant, regardless of sampling locality or species designation, with 119 of 122 individuals sharing the same haplotype over 681 bp. The three variants within *Oculina* were all singletons and all differed from the dominant haplotype by a single synonymous base pair substitution. In contrast, levels of intraspecific variation at

Table 3 *Oculina* sequence variation

	Full sequence					
	Length (bp)	*S*	No. of haplotypes*	Hap div*	*K*	π
FA elongase	425–510	45	73	0.863	4.027	0.0095
EF-1α	470	10	36	0.806	1.898	0.0054
Tachylectin	429–444	38	53	0.835	5.096	0.0119
	ORF					
	Length (bp)	*S*	No. of haplotypes *	Hap div*	*K*	π
FA elongase	276	24	34	0.819	2.391	0.0087
EF-1α	351	10	23	0.780	1.898	0.0054
Tachylectin	276	25	40	0.809	3.560	0.0130

*Haplotype values calculated under infinite allele model in Arlequin version 3.1.1; *S*, number of segregation sites; *K*, average number of nucleotide differences; π, nucleotide diversity.
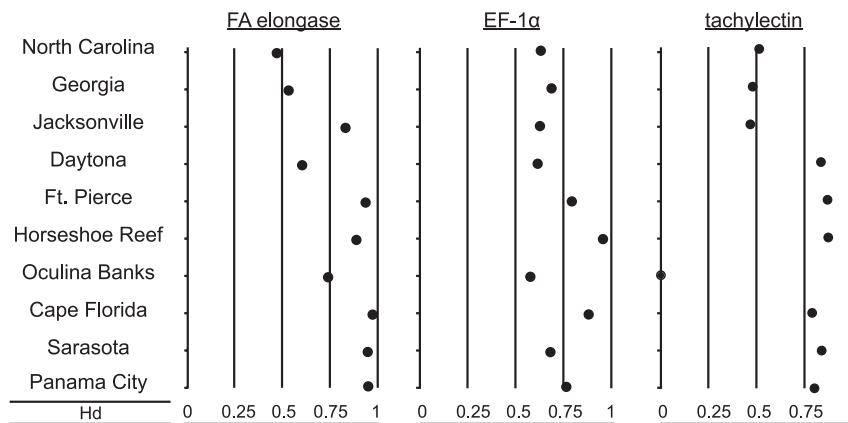
Fig. 2 Variation among populations in haplotype diversity (H_d) for the three nuclear markers used in the study. Populations are arranged contiguously top to bottom beginning with the northernmost population, North Carolina. The biogeographical break at Cape Canaveral occurs between Ft. Pierce and Daytona.

the three nuclear markers were quite high (Table 3). The number of segregating sites varied from 10 (EF-1) to 45 (FA elongase). EF-1 alleles differed by about 2 bp on average, while tachylectin alleles differed by about 5 bp. Nucleotide diversity (π) ran between 0.005 and 0.012. Variation in the ORFs of the three markers was similarly high (Table 3).

Variation in haplotype diversity (Fig. 2) approaches its theoretical maximum, ranging from zero for tachylectin from the Oculina Banks population to near unity (0.95) for EF-1 at Horseshoe Reef (just 29 km away). Two patterns emerged from inspection of these values. First, the four populations north of Cape Canaveral (North Carolina, Georgia, Jacksonville and Daytona) were less variable than populations elsewhere in the range. Second, the Oculina Banks population had the lowest levels of variation among the southern populations for all three nuclear markers.

Analysis of molecular variation revealed Φ_{ST} values that were significant at the $P = 0.05$ level when individuals were partitioned by location, nominal species, and depth (Table 4). Location had the highest values, with the values for nominal species and depth both dropping when location

was accounted for. AMOVA results were roughly similar across loci (with the exception of the effect of depth on tachylectin subdivision). Overall, about 16% of all variation could be traced to subdivision among all 10 sampled populations. Depth accounted for about 10% of variation overall. Genotypes from shallow (Radio Island and Piver's Island, 1.5–4 m) and mid-depth (38 km Reef and ISO5, 23–26 m) sites off North Carolina were shared and similar (data not shown). Nominal species designations meant even less than depth, accounting for about 8.5% of variation. Proportions became higher once Φ_{ST} was standardized for levels of variation within populations (Table 4), but the rank order of importance for the three sources of variation remained the same. When the potentially differentiated population from the Oculina Banks was removed from the analysis, nominal species and depth had a further diminished impact, failing even to reach significance over all three loci (Table 4). Pooling the *Oculina arbuscula* samples with *Oculina varicosa* had little impact on the proportion of overall variance explained by species or the other factors (not shown).

Table 4 Analysis of molecular variation among locations, nominal species, and collection depths

		Φ_{ST}*				Standardized Φ_{ST}*			
	Source of variation	FAelo	EF-1α	Tachy	Overall	FAelo	EF-1α	Tachy	Overall
w/Oculina Banks	Location	0.119	0.136	0.233	0.163	0.552	0.464	0.677	0.554
	'Species'	0.079	0.077	0.098	0.085	0.421	0.320	0.431	0.391
	Depth	0.055	0.057	0.189	0.102	0.338	0.251	0.663	0.417
	'Species' — Location	0.050	0.038	0.019	0.036	0.310	0.183	0.120	0.204
	Depth – Location	0.025^n	0.021^n	0.152	0.088	0.181^n	0.104^n	0.601	0.295
w/o Oculina Banks	Location	0.094	0.118	0.134	0.115	0.418	0.422	0.470	0.437
	'Species'	0.081	0.080	0.088	0.083	0.398	0.326	0.366	0.383
	Depth	0.001^n	0.019	0.025	0.015	0.008^n	0.096	0.132	0.079
	'Species' — Location	0.066	0.049	0.055^n	0.057^n	0.347	0.224	0.258^n	0.265^n
	Depth — Location	-0.020^n	-0.005^n	-0.001^n	-0.012^n	-0.155^n	-0.030^n	-0.017^n	-0.067^n

*All Φ_{ST} values significant at the $P = 0.05$ level unless marked with n.
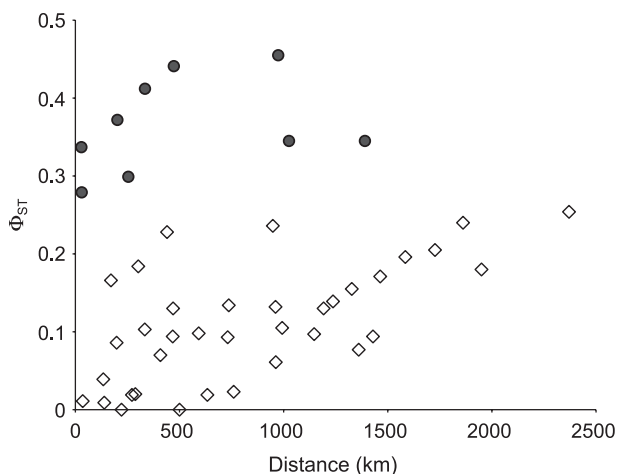
Fig. 3 Pairwise Φ_{ST} values for the concatenated nuclear gene data set, plotted against nautical distance. Closed circles indicate comparisons involving the 80 m Oculina Banks (Jeff's Reef) population, open diamonds indicate comparisons for all other populations.

Pairwise values of Φ_{ST} were plotted against distance of population separation to see whether overall measures of subdivision masked any population-specific patterns (Fig. 3). The relationship between Φ_{ST} and distance was weak ($r^2 = 0.021$). However, this analysis identified the Oculina Banks population as a strong outlier: the values of Φ_{ST} for every pairwise comparison involving the Oculina Banks were higher than for Φ_{ST} involving all other pairs of populations (Fig. 3). This difference held true even though the three populations in the Ft. Pierce area (Jeff's Reef, Horseshoe Reef, and Ft. Pierce) were all within 36 km of each other, while some of the other (genetically closer) populations were separated by up to 2370 km. When the Oculina Banks population was removed from the analysis, distance then explained a significant proportion of the variation in Φ_{ST} ($r^2 = 0.39$).

Results from the Bayesian clustering analyses further supported the conclusion that the Oculina Banks population is genetically isolated from all other populations sampled. Using Structurama, $k = 3$ had the highest posterior probability for both the full length data and the ORF-only data. The full-length data identify one of the three multilocus clusters as strongly associated with the Oculina Banks population (Fig. 4). All individuals from Oculina Banks have at least 93% of their genome assigned to the same cluster, while no individuals from outside the Oculina Banks have > 68% of their genome assigned to this (the red in Fig. 4a) cluster, with no more than one individual per locality greater than 45%. Truncating the full-length sequences to just ORFs reduced the number of distinguishable alleles for all three loci (e.g. from 73 to 34 for FA elongase, from 38 to 23 for EF-1, and from 53 to 40 for tachylectin), but the Oculina Banks population remains distinct using the ORF data (Fig. 4b).

The two other clusters (aside from the Oculina Banks) were partitioned among populations as well. The genomes of all individuals from the four populations north of Cape Canaveral fell largely into one of these clusters (blue in Fig. 4), which was also prevalent in several individuals from the distant Sarasota population. This clustering appears to be driven in large part by the presence of the most common northern alleles at FA elongase. This allele is frequent (at 62.7%) in the four populations north of Cape Canaveral, and next most common in Sarasota (20.8%), the southern population with the highest northern component. The genotypes of exceptional individuals (those that cluster differently from others in their same population based on the full sequence analysis) are also instructive here: 9 of 13 individuals with a high (> 60%) proportion of north (blue) in the Ft. Pierce, Cape Florida and Sarasota
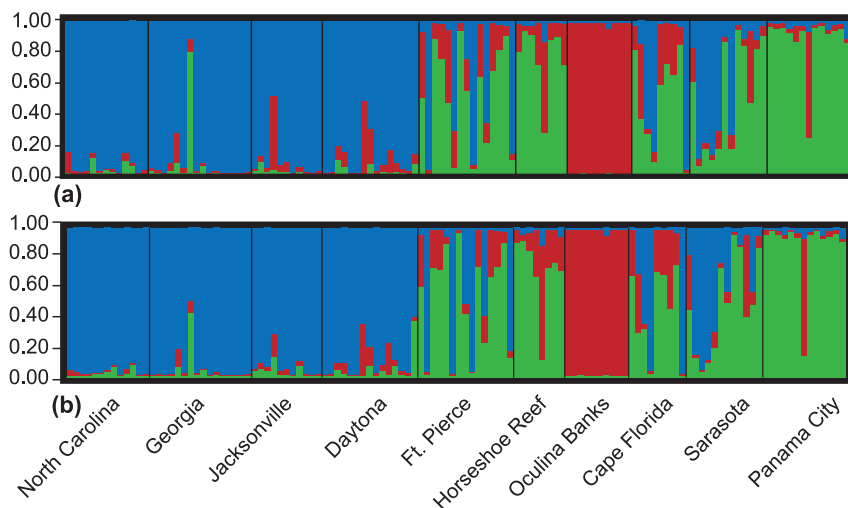
Fig. 4 Graphical summary of the results from the Structure analysis for *k* = 3 for (a) full and (b) coding region only data sets. Each individual is represented by a vertical line broken into three segments to represent the estimated proportions of that individual's genome originating from each of the three inferred clusters.

populations possessed at least one copy of the most common FA elongase northern allele (otherwise rare in the south), while the single individual from Georgia with a more southern (green) genome did not.

The northern and southern clusters did not correlate with nominal species (compare Fig. 4 and the Appendix). For example, nominal *O. arbuscula* (North Carolina and Georgia) falls into the northern cluster with *O. varicosa* from Jacksonville and Daytona, although nominal individuals of *O. varicosa* fall into the other two clusters as well. Individuals from the Sarasota population are all *O. robusta* by morphology, but genetically appear to be mixed between the northern and southern clusters.

Discussion

Nuclear sequence markers for taxa with slow mtDNA: possibilities and problems

Previous studies on plants, fungi, sponges, and anthozoans, including corals, have reported extremely low levels of mtDNA variation among populations of the same nominal species (references in Hellberg 2006). When divergence rates have been estimated, these appear to be 50–100× slower than for bilateral animals. In plants, these slow rates of mitochondrial sequence evolution are not paralleled by relatively slow rates for nuclear genes (Wolfe *et al*. 1987). This same pattern holds for corals: rates of divergence for *Oculina* and *Solenastrea* were 6.6–9.1 times faster (uncorrected *p*) for nDNA than for mtDNA (Table 2).

Levels of nucleotide diversity were also far higher for nuclear markers examined here (Fig. 2) than for mtDNA (which was nearly fixed). Such high levels of nDNA sequence variation are similar to those seen previously in plants (Moeller & Tiffin 2005) and marine animals (Taylor & Hellberg 2006), including corals (Nunes and Knowlton,

unpublished data). These higher rates offer hope for revealing population isolation within coral species; however, nuclear markers in these taxa will still offer challenges beyond those commonly seen for mtDNA in bilateral animals.

The two nuclear loci with the highest levels of nucleotide sequence variation (fatty acid elongase and tachylectin-2, Table 3) also showed high frequencies of indels. Biologically, it may be that nucleotide sequence variation and indel variation are linked mechanistically (Tian *et al*. 2008). Practically, for the population geneticist looking to score both alleles at multiple individuals, resolving indels by cloning can (and, for us, did) prove costly and time-consuming. We found most indels in the 3′ UTR regions we sequenced, which we initially targeted in the belief that they would be richer sources of informative variation. Recent work and our results suggest this need not be the case. Andolfatto (2005) found silent sites within open reading frames are at least three times as variable as noncoding sites elsewhere in the genome, compensating for their threefold lower frequency within exons. Here we found that ORF-only sequences were nearly as variable as those including the 3′ UTR (Table 3), and that Structure could still identify the Oculina Banks population as isolated using the reduced, ORF-only data (Fig. 4). These results suggest that sequencing markers set in ORFs may reveal ample power to resolve population differences while avoiding the practical problems of resolving indel heterozygotes.

Indel heterozygotes also complicate analysis of another feature associated with high levels of nucleotide variation: recombination (Begun & Aquadro 1992). Four-gamete tests found recombination at all three nuclear loci, although these tests were complicated by problems with coding indels. High levels of recombination are not rare for population surveys of nuclear sequences, even for sequences shorter than those surveyed here (e.g. Ibrahim *et al*. 2002). Inspection of recombination patterns can reveal stretches

of sequence that have maintained their integrity, thus indicating a recent shared history and an appropriate basis for coalescent analyses. Indels can complicate these analyses because some programs don't allow them as input. At loci with high rates of recombination, remaining stretches may have few variable sites and thus limited statistical power. Our results thus represent a worst-case scenario in which recombination restricted data analysis; however, analyses based on infinite-allele assumptions nevertheless revealed patterns consistent with population isolation in *Oculina*.

Subdivision, population isolation and species status within *Oculina*

Populations of *Oculina* from the southeastern coast of the USA are genetically subdivided (Table 4). Limited larval dispersal may underlie some of this pattern: the larvae of *Oculina varicosa* swim actively and are negatively geotactic for 1–2 weeks after hatching and become negatively phototactic after about 14 days (Brooke & Young 2005). The high proportion ($\approx$ 40%) of the variation in Φ_{ST} between populations that is due to geographical separation also suggests most dispersal occurs between neighbouring populations, as in other coastal corals (Hellberg 1995).

Characteristics shared by more than one population generally explain little of the overall genetic variation in *Oculina*. Species designations have been considered problematic in the genus, and those used here do not designate genetically meaningful entities. Much of the variation attributable to species in the AMOVA analysis (Table 4) stems from the geographical nature of existing species definitions: all *Oculina* from North Carolina and Georgia have been called *O. arbuscula*, while *O. robusta* has been largely restricted to the Gulf of Mexico. The possibility remains, however, that the genetically distinguishable clusters identified here, while not coincident with existing species definitions, nonetheless represent species or populations on a course towards reproductive isolation.

For northern (blue, Fig. 4) and southern (green) clusters, this does not seem to be the case. Populations to the north of Cape Canaveral were largely united by the Structure analysis. These same northern populations also show reduced variation (Fig. 2). The alleles present in these populations are a subset of those found over the rest of the sampled range, not in any way phylogenetically distinct, and there is no indication of differentiation among these northernmost populations. In combination, these patterns are consistent with a relatively recent range expansion north of Cape Canaveral, a long-recognized marine phylogeographical break (Avise 2000). Unlike a traditional phylogeographical break, which separates reciprocally monophyletic clades, the break here marks a decline in heterozygosity beyond a barrier. Similar patterns have been seen for other marine animals, including an intertidal snail moving poleward past a historical barrier at Point Conception in California (e.g. Hellberg *et al.* 2001) and a tropical goby returning to habitat denuded by recent sea level changes (Thompson *et al.* 2005).

The major source of subdivision that we found in *Oculina* involved the deep-water corals from the Oculina Banks. The combination of AMOVA, pairwise Φ_{ST}, and Structure analyses all suggest that the Oculina Banks population is genetically isolated from all others and perhaps already a separate reproductively isolated species. Multilocus clustering singled out this population as distinct (Fig. 4), and every pairwise value of Φ_{ST} was greater for comparisons involving the Oculina Banks population than for all other comparisons (Fig. 3). These results strongly suggest that the Oculina Banks population is genetically isolated from all shallower (*c.* 30 m or less) populations. While larvae from deep and shallow populations have similar broad temperature tolerances (Brooke & Young 2005), colony growth rates appear to be faster for the deep population (Reed 1981). Furthermore, Brooke (2002) found that shallow populations in the Ft. Pierce area spawn 2 or 3 weeks before those on the Oculina Banks. Such a difference in reproductive timing may result from responding to similar seasonal cues that differ with depth, or could indicate species-specific breeding seasonality. Whichever the reason, these differences should facilitate the continued isolation and divergence of populations. That two closely related but genetically isolated populations should be segregated by depth is not unusual for marine organisms. Geographically sympatric sister species that live at different depths have been reported many times (see Knowlton 1993; Hellberg 1998; Hyde *et al.* 2008), including for corals and other anthozoans (Knowlton *et al.* 1992; Carlon & Budd 2002; Prada *et al.* 2008).

The isolated Oculina Banks population occurs in an ecologically different habitat below 50 m, a bathymetric line that has been drawn between deep and shallow water corals (Cairns 2007). Alleles from the deep-water population nest phylogenetically within the more broadly distributed (and paraphyletic) shallow form, consistent with the notion that deep sea species are often derived from shallow water ones (Jablonski *et al.* 1983), although hydrocorals provide a counterexample (Lindner *et al.* 2008). More unusual is the geographical nesting of its range: the deep-water population occurs near the centre of *Oculina*'s continental geographical range (Fig. 1) and only a short distance (< 50 km) from shallow-water populations. High relief reef habitat at the depth of the Oculina Banks is presently rare along Florida's eastern coast (Parker *et al.* 1983), and the Oculina Banks population may represent a geographically restricted relic of a formerly more broadly distributed form. Genetic analysis of newly discovered deep-water populations of *O. varicosa* from the northeastern Gulf of Mexico (Barnette 2006), as well as populations from further south in *Oculina*'s range, may help resolve the origins of this curious population.

Whatever that history, the corals of the Oculina Banks have created an ecosystem that harbours exceptionally high diversity (Reed 2002) and provides a nursery and feeding grounds to several commercially harvested fish (Koenig *et al.* 2000). Our results suggest that any efforts to preserve and restore this ecosystem will have to be based on the recognition that the population of *Oculina* at the Oculina Banks are genetically isolated from shallow water populations of the genus.

Acknowledgements

References

Agassiz L (1864) Pp. 46–47 in AE Verrill: List of the polyps and corals sent by the Museum of Comparative Zoology to other institutions in exchange, with annotations. *Bulletin of the Museum of Comparative Zoology*, **1**, 29–60.

Andolfatto P (2005) Adaptive evolution of non-coding DNA in Drosophila. *Nature*, **437**, 1149–1152.

Avise JC (2000) *Phylogeography: the History and Formation of Species.* Harvard University Press, Cambridge, Massachusetts.

Aylor DL, Price EW, Carbone I (2006) Snap: Combine and MAP modules for multilocus population genetic analysis. *Bioinformatics*, **22**, 1399–1401.

Baker AM, Bartlett C, Bunn SE *et al.* (2003) Cryptic species and morphological plasticity in long-lived bivalves (Unionoida: Hyriidae) from inland Australia. *Molecular Ecology*, **12**, 2707–2717.

Barnette MC (2006) *Observations of the deep-water coral* Oculina varicosa *in the Gulf of Mexico*. NOAA Technical Memorandum NMFS-SEFSC-535.

Baums IB (2008) A restoration guide for coral reef conservation. *Molecular Ecology*, **17**, 2796–2811.

Baums IB, Miller MW, Hellberg ME (2005) Regionally isolated populations of an imperiled Caribbean coral, *Acropora palmata*. *Molecular Ecology*, **14**, 1377–1390.

Begun DJ, Aquadro CF (1992) Levels of naturally occurring DNA polymorphism correlate with recombination rates in *Drosophila melanogaster*. *Nature*, **356**, 519–520.

Brooke SD (2002) *Reproductive ecology of a deep-water scleractinian coral Oculina varicosa from the southeast Florida shelf*. PhD diss. University of Southampton, Southampton, UK.

Brooke S, Young CM (2003) Reproductive ecology of a deep-water scleractinian coral, *Oculina varicosa*, from the southeast Florida shelf. *Continental Shelf Research*, **23**, 847–858.

Brooke S, Young CM (2005) Embryogenesis and larval biology of the ahermatypic scleractinian *Oculina varicosa*. *Marine Biology*, **146**, 665–675.

Bruno JF, Edmunds PJ (1997) Clonal variation for phenotypic plasticity in the coral *Madracis mirabilis*. *Ecology*, **78**, 2177–2190.

Cairns SD (1991) A revision of the ahermatypic scleractinia of the Galapagos and Cocos Islands. *Smithsonian Contributions to Zoology*, **504**, 32.

Cairns SD (2007) Deep-water corals: an overview with special reference to diversity and distribution of deep-water scleractinian corals. *Bulletin of Marine Science*, **81**, 311–322.

Carling MD, Brumfield RT (2008) Integrating phylogenetic and population genetic analyses of multiple loci to test species divergence hypotheses in *Passerina* buntings. *Genetics*, **178**, 363–377.

Carlon DB, Budd AF (2002) Incipient speciation across a depth gradient in a scleractinian coral? *Evolution*, **56**, 2227–2242.

Chomczyniski P, Sacchi N (1987) Single-step method of RNA isolation by acid guanidinium thiocyanate phenol chloroform extraction. *Analytical Biochemistry*, **162**, 156–159.

Clement M, Posada D, Crandall KA (2000) TCS: a computer program to estimate gene genealogies. *Molecular Ecology*, **9**, 1657–1659.

Combosch DJ, Guzman HM, Schuhmacher H, Vollmer SV (2008) Interspecific hybridization and restricted trans-Pacific gene flow in the tropical eastern Pacific. *Molecular Ecology*, **17**, 1304–1312.

Crandall KA (1999) Uses of statistical parsimony in HIV analyses. In: *Statistics in Molecular Biology and Genetics* (ed. Seillier-Moiseiwitsch F). Institute of Mathematical Statistics and the American Mathematical Society, Hayward, California.

Drummond AJ, Ashton B, Cheung M *et al.* (2007) *Geneious version 3.6*. Available from URL: http://www.geneious.com.

Edgar RC (2004) Muscle: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**, 1792–1797.

Erwin PM, Thacker RW (2007) Phylogenetic analyses of marine sponges within the order Verongida: a comparison of morphological and molecular data. *Invertebrate Biology*, **126**, 220–234.

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes — application to human mitochondrial-DNA restriction data. *Genetics*, **131**, 479–491.

Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, **164**, 1567–1587.

Fine M, Zibrowius H, Loya Y (2001) *Oculina patagonica*: a non-lessepsian scleractinian coral invading the Mediterranean Sea. *Marine Biology*, **138**, 1195–1203.

Flot J-F, Tillier S (2007) The mitochondrial genome of *Pocillopora* (Cnidaria: Scleractinia) contains two variable regions: the putative D-loop and a novel ORF of unknown function. *Gene*, **401**, 80–87.

Flot J-F, Magalon H, Cruaud C, Couloux A, Tillier S (2008) Patterns of genetic structure among Hawaiian corals of the genus *Pocillopora* yield clusters of individuals that are compatible with morphology. *CR Biologies*, **331**, 239–247.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology*, **3**, 294–299.

Foster AB (1979) Phenotypic plasticity in the reef coral *Montastraea annularis* (Ellis and Solander) and *Siderastrea siderea* (Ellis and Solander). *Journal of Experimental Marine Biology and Ecology*, **39**, 25–54.

Fukami H, Knowlton N (2005) Analysis of complete mitochondrial DNA sequences of three members of the *Montastrea annularis* coral species complex (Cnidaria, Anthozoa, Scleractinia). *Coral Reefs*, **24**, 410–417.

Fukami H, Budd AF, Pauley G *et al.* (2004) Conventional taxonomy obscures deep divergence between Pacific and Atlantic corals. *Nature*, **427**, 832–835.

Fukami H, Chen CA, Budd AF *et al.* (2008) Mitochondrial and nuclear genes suggest that stony corals are monophyletic but most families of stony corals are not (Order Scleractinia, Class Anthozoa, Phylum Cnidaria). *PLoS One*, **3**, e3222.

Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, **270**, 313–321.

Hedrick P (2005) A standardized genetic differentiation measure. *Evolution*, **59**, 1633–1638.

Hellberg ME (1994) Relationships between inferred levels of gene flow and geographic distance in a philopatric coral, *Balanophyllia elegans*. *Evolution*, **48**, 1829–1854.

Hellberg ME (1995) Stepping-stone gene flow in the solitary coral *Balanophyllia elegans:* equilibrium and nonequilibrium at different spatial scales. *Marine Biology*, **123**, 573–581.

Hellberg ME (1998) Sympatric species along the sea's shore: the geography of speciation in the marine gastropod *Tegula*. *Evolution*, **52**, 1311–1324.

Hellberg ME (2006) No variation and low synonymous substitution rates in coral mtDNA despite high nuclear variation. *BMC Evolutionary Biology*, **6**, 24.

Hellberg ME, Balch DP, Roy K (2001) Climate-driven range expansion and morphological evolution in a marine gastropod. *Science*, **292**, 1707–1710.

Hoegh-Guldberg O, Mumby PJ, Hooten AJ *et al.* (2007) Coral reefs under rapid climate change and ocean acidification. *Science*, **318**, 1737–1742.

Huang DW, 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.

Hudson RR, Kaplan NL (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics*, **111**, 147–164.

Huelsenbeck JP, Andolfatto P (2007) Inference of population structure under a Dirichlet process model. *Genetics*, **175**, 1787–1802.

Humann P (1993) *Reef Coral Identification*. New World Publications, Inc, Jacksonville, Florida.

Hyde JR, Kimbrell CA, Budrick JE, Lynn EA, Vetter RD (2008) Cryptic speciation in the vermilion rockfish (*Sebastes miniatus*) and the role of bathymetry in the speciation process. *Molecular Ecology*, **17**, 1122–1136.

Ibrahim KM, Cooper SJB, Hewitt GM (2002) Testing for recombination in a short nuclear DNA sequence of the European meadow grasshopper, *Chorthippus parallelus*. *Molecular Ecology*, **11**, 583–590.

Jablonski D, Sepkoski JJ, Bottjer DJ, Sheehan PM (1983) Onshore–offshore patterns in the evolution of Phanerozoic shelf communities. *Science*, **222**, 1123–1125.

King RA, Tibble AL, Symondson WOC (2008) Opening a can of worms: unprecedented sympatric cryptic diversity within British lumbricid earthworms. *Molecular Ecology*, **17**, 4684–4698.

Klautau M, Russo CAM, Lazoski C *et al.* (2003) Does cosmopolitanism result from overconservative systematics? A case study using the marine sponge *Chondrilla nucula*. *Evolution*, **53**, 1414–1422.

Knowlton N (1993) Sibling species in the sea. *Annual Review of Ecology and Systematics*, **24**, 189–216.

Knowlton N (2001) Who are the players on coral reefs and does it matter? The importance of coral taxonomy for coral reef management. *Bulletin of Marine Science*, **69**, 305–308.

Knowlton N, Weil E, Weight LA, Guzman HM (1992) Sibling species in *Montastraea annularis*, coral bleaching, and the coral climate record. *Science*, **255**, 330–333.

Koenig CC, Coleman FC, Grimes CB *et al.* (2000) Protection of fish spawning habitat for the conservation of warm-temperate reef-fish fisheries of shelf-edge reefs of Florida. *Bulletin of Marine Science*, **66**, 593–616.

Lamarck JBPA (1816) *Histoire Naturelle des Animaux sans Vertèbres*, 2. Verdière. Paris.

Leache AD, McGuire JA (2006) Phylogenetic relationships of horned lizards (*Phrynosoma*) based on nuclear and mitochondrial data: evidence for a misleading mitochondrial gene tree. *Molecular Phylogenetics and Evolution*, **39**, 628–644.

LeSueur CP (1821) Description de plusieurs animaux appartement aux polypiers lamellifères de M le Chev. de Lamarck. *Mémoires du Museum Histoire Naturelle Paris*, **6**, 271–298.

Lindner A, Cairns SD, Cunningham CW (2008) From offshore to onshore: multiple origins of shallow-water corals from deep-sea ancestors. *PLoS One*, **3**, e2429.

Lopez JV, Kersanach R, Rehner SA, Knowlton N (1999) Molecular determination of species boundaries in corals: genetic analysis of the *Montastrea annularis* complex using amplified fragment length polymorphisms and a microsatellite marker. *Biology Bulletin*, **196**, 80–93.

Marquez LM, van Oppen MJH, Willis BL, Miller DJ (2002) Sympatric populations of the highly cross-fertile coral species *Acropora hyacinthus* and *Acropora cytherea* are genetically distinct. *Proceedings of the Royal Society B: Biological Sciences*, **269**, 1289–1294.

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.

Meirmans PG, Van Tienderen PH (2004) GenoType and GenoDive: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, **4**, 792–794.

Michalakis Y, Excoffier L (1996) A generic estimation of population subdivision using distances between alleles with special reference for microsatellite loci. *Genetics*, **142**, 1061–1064.

Miermans PG (2006) Using the AMOVA framework to estimate a standardized genetic differentiation measure. *Evolution*, **60**, 2399–2402.

Miller MW, Hay ME (1996) Coral-seaweed–grazer–nutrient interactions on temperate reefs. *Ecological Monographs*, **66**, 323–344.

Moeller DA, Tiffin P (2005) Genetic diversity and the evolutionary history of plant immunity genes in two species of *Zea*. *Molecular Biology and Evolution*, **22**, 2480–2490.

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution*, **3**, 418–426.

Nei M, Kumar S (2000) *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.

Niemiller ML, Fitzpatrick BM, Miller BT (2008) Recent divergence with gene flow in Tennessee cave salamanders (Plethodontidae: Gyrinophilus) inferred from gene genealogies. *Molecular Ecology*, **17**, 2258–2275.

van Oppen MJH, McDonald BJ, Willis B, Miller DJ (2001) The evolutionary history of the coral genus *Acropora* (Scleractinia, Cnidaria) based on a mitochondrial and a nuclear marker: reticulation, incomplete lineage sorting, or morphological convergence? *Molecular Biology and Evolution*, **18**, 1315–1329.

Parker RO, Colby DR, Willis TD (1983) Estimated amount of reef habitat on a portion of the United-States South-Atlantic and Gulf of Mexico continental shelf. *Bulletin of Marine Science*, **33**, 935–940.

Posada D (2008) jModelTest: Phylogenetic model averaging. *Molecular Biology and Evolution*, **25**, 1253–1256.

Pourtalès LF (1871) Deep-sea corals. *Illustrated Catalogue of the Museum of Comparative Zoology*, **4**, 93.

Prada C, Schizas NV, Yoshioka PM (2008) Phenotypic plasticity or speciation? A case from a clonal marine organism. *BMC Evolutionary Biology*, **8**, 47.

Price EW, Carbone I (2005) Snap: workbench management tool for evolutionary population genetic analysis. *Bioinformatics*, **21**, 402–404.

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

Pritchard JK, Wen X, Falush D (2007) *Documentation for Structure Software. Version 2.2*. Available from URL: http://pritch.bsd.uchicago.edu.

Putnam NH, Srivastava M, Hellsten U *et al.* (2007) Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science*, **317**, 86–94.

Rambaut A, Drummond AJ (2007) *Tracer version 1.4*. Available from URL: http://beast.bio.ed.ac.uk/Tracer.

Redelings BD, Suchard MA (2005) Joint Bayesian estimation of alignment and phylogeny. *Systematic Biology*, **54**, 401–418.

Reed JK (1981) In situ growth rates of the scleractinian coral *Oculina varicosa* occurring with zooxanthellae on 6 m reefs and without on 80 m banks. *Proceedings of the 4th International Coral Reef Symposium*, **2**, 201–206.

Reed JK (2002) Deep-water *Oculina* coral reefs of Florida: biology, impacts, and management. *Hydrobiologia*, **471**, 43–55.

Reed JK, Koenig CC, Shepard AN (2007) Impacts of bottom trawling on a deepwater Oculina coral ecosystem off Florida. *Bulletin of Marine Science*, **81**, 481–496.

Ridgway T, Gates RD (2006) Why are there so few genetic markers available for coral population analyses? *Symbiosis*, **41**, 1–7.

Rozas J, Sanchez-De I, Barrio JC, Messeguer X, Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics*, **19**, 2496–2497.

Rozen S, Skaletsky H (2000) Primer 3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology*, **132**, 365–386.

Rupert EE, Fox RS (1988) *Seashore Animals of the Southeast: A Guide to Common Shallow-Water Invertebrates of the Southeastern Atlantic Coast*. University Of South Carolina Press, Columbia, South Carolina.

Schwarz JA, Brokstein PB, Voolstra C *et al.* (2008) Coral life history and symbiosis: functional genomic resources for two reef building Caribbean corals, *Acropora palmata* and *Montastraea faveolata*. *BMC Genomics*, **9**, 97.

Shearer TL, Gutiérrez-Rodríguez C, Coffroth MA (2005) Generating molecular markers from zooxanthellate cnidarians. *Coral Reefs*, **24**, 57–66.

Shearer T, Oppen MV, Romano S, Worheide G (2002) Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). *Molecular Ecology*, **11**, 2475–2487.

Stallings RL (1995) Conservation and evolution of (Ct)(N)/ (Ga)(N) microsatellite sequences at orthologous positions in diverse mammalian genomes. *Genomics*, **25**, 107–113.

Steinauer ML, Nickol BB, Orti G (2007) Cryptic speciation and patterns of phenotypic variation of a highly variable acanthocephalan parasite. *Molecular Ecology*, **16**, 4097–4109.

Stephens M, Donnelly P (2003) A comparison of Bayesian methods for haplotype reconstruction from population genotype data. *American Journal of Human Genetics*, **73**, 1162–1169.

Stephens M, Scheet P (2005) Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. *American Journal of Human Genetics*, **76**, 449–462.

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.

Suchard MA, Redelings BD (2006) BAli-Phy: simultaneous Bayesian inference of alignment and phylogeny. *Bioinformatics*, **22**, 2047–2048.

Tajima F (1983) Evolutionary relationship of DNA sequences in finite populations. *Genetics*, **105**, 437–460.

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA 4: molecular evolutionary genetics analysis (MEGA) software version 4.0. *Molecular Biology and Evolution*, **24**, 1596–1599.

Taylor MS, Hellberg ME (2006) Comparative phylogeography of a genus of coral reef fishes: biogeographical and genetical concordance in the Caribbean. *Molecular Ecology*, **15**, 695–707.

Thompson AR, Thacker CE, Shaw EY (2005) Phylogeography of marine mutualists: parallel patterns of genetic structure between obligate goby and shrimp partners. *Molecular Ecology*, **14**, 3557–3572.

Tian DC, Wang Q, Zhang PF *et al.* (2008) Single-nucleotide mutation rate increases close to insertions/deletions in eukaryotes. *Nature*, **455**, 105–108.

Todd PA (2008) Morphological plasticity in scleractinian corals. *Biology Reviews*, **83**, 315–337.

Verrill AE (1902) Papers on corals. *Transactions of the Connecticut Academy of Arts and Sciences*, **XI**, 63–266.

Villesen P (2007) FaBox: an online toolbox for FASTA sequences. *Molecular Ecology Notes*, **7**, 965–968.

Vowles EJ, Amos W (2004) Evidence for widespread convergent evolution around human microsatellites. *PLoS Biology*, **2**, e199.

Weir BS, Cockerham CC (1984) Estimating *F*-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Willis BL (1985) Phenotypic plasticity versus phenotypic stability in the reef corals *Turbinaria mesenterina* and *Pavona cactus*. *Proceedings of the 5th International Coral Reef Symposium*, **6**, 107–112.

Willis BL, van Oppen MJH, Miller DJ, Vollmer SV, Ayre DJ (2006) The role of hybridization in the evolution of reef corals. *Annual Review of Ecology, Evolution and Systematics*, **37**, 489–537.

Woerner AE, Cox MP, Hammer MF (2007) Recombination-filtered genomic datasets by information maximization. *Bioinformatics*, **23**, 1851–1853.

Wolfe KH, Li WH, Sharp PM (1987) Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proceedings of the National Academy of Sciences, USA*, **84**, 9054–9058.

Zlatarski VN, Martinez Estalella N (1982) *Les Scléractiniaires de Cuba avec des données sur les organismes associés*. Editions de l'Académie bulgare des sciences, Sofia, Bulgaria.

RIE is interested in the origin and maintenance of marine biodiversity, particularly in coral reef taxa. MH is a Research Associate who applies molecular genetics, functional genomics and biochemical approaches to the study of host microbe interactions, using plants, invertebrates and bacteria as model organisms. His current research focuses on secreted bacterial virulence proteins and their role in plant disease. PAR is a Research Associate currently working on fat cells and canine restoration. MM is a NOAA scientist working on coral population ecology and restoration. MEH is broadly interested in population isolation, speciation and molecular evolution in marine animals, especially anthozoans and gastropods.

Appendix

Oculina *sampling localities*

Location information for sublocalities included when individuals came from more than one place, with numbers from each sublocality indicated parenthetically. Average collecting depth or range indicated, along with depth class: shallow: < 12 m; medium: 20–30 m; deep: > 70 m

North Carolina *O. arbuscula* (13)	
Radio Island Jetty (9)	34 42.58′N, 76 40.85′W (S, 2–4 m)
38 km Reef (2)	34 19.99′N, 76 53.90′W (M, 26 m)
ISO5 (2)	34 23.29′N, 76 34.23′W (M, 23 m)
Georgia *O. arbuscula* (16)	
J Reef (6)	31 36 06′N, 80 47.43′W (M, 21 m)
R2 tower (10)	31 22.10′N, 80 35.03′W (M, 27 m)
Jacksonville, FL *O. varicosa* (11)	
Paul Mains (6)	30 19.81′N, 81 10.98′W (M, 23 m)
Pablo G Culverts (5)	30 20.09′N, 81 11.74′W (M, 21 m)
Daytona, FL *O. varicosa* (15)	
Mindinao (9)	29 11.97′N, 80 44.85′W (M, 21 m)
Culverts (6)	29 19.27′N, 80 44.67′W (M, 23 m)
Fort Pierce Inlet, FL (15) *O. diffusa* (5), *O. robusta* (4), *O. varicosa* (6)	27 27.61′N, 80 16.99′W (S, < 2 m)
Horseshoe Reef, FL *O. varicosa* (8)	27 45.22′N, 80 07.86′W (M, 29 m)
Jeff's Reef, FL *O. varicosa* (10)	27 31.86′N, 79 58.81′W (D, 80 m)
Cape Florida, FL *O. diffusa* (9)	25 39.99′N, 80 09.34′W (S, 2 m)
Sarasota, FL *O. robusta* (12)	27 26.64′N, 82 49.20′W (S, 11 m)
Panama City, FL *O. diffusa* (13)	
Site 1 (8)	30 03.26′N, 85 51.99′W (M, 29 m)
Site 2 (5)	30 02.09′N, 85 51.12′W (M, 28 m)