Solving the Coral Species Delimitation Conundrum

CATALINA RAMÍREZ-PORTILLA[1,2,*], ANDREW H. BAIRD[3], PETER F. COWMAN[3,4], ANDREA M. QUATTRINI[5], SAKI HARII[6], FREDERIC SINNIGER[6], AND JEAN-FRANÇOIS FLOT[1,7]

[1]*Evolutionary Biology and Ecology (EBE), Département de Biologie des Organismes, Université libre de Bruxelles (ULB), Av. Franklin D. Roosevelt 50, CP 160/12, Brussels B-1050, Belgium;* [2]*Systematics and Biodiversity, Department of Animal Ecology and Systematics, Justus Liebig University, Heinrich-Buff-Ring 26-32 IFZ, Giessen D-35392, Germany;* [3]*Australian Research Council (ARC) Centre of Excellence for Coral Reef Studies, James Cook University, 1 University Drive, Townsville, QLD 4811, Australia;* [4]*Biodiversity and Geosciences Program, Museum of Tropical Queensland, Queensland Museum, Townsville, QLD 4810, Australia;* [5]*Department of Invertebrate Zoology, National Museum of Natural History, Smithsonian Institution, 1000 Madison Drive NW, Washington DC 20560, USA;* [6]*Sesoko Station, Tropical Biosphere Research Center, University of the Ryukyus, 3422 Sesoko, Motobu, Okinawa 905-0227, Japan;* [7]*Interuniversity Institute of Bioinformatics in Brussels - (IB)[2], ULB-VUB, Triomflaan, CP 263, Brussels B-1050, Belgium*
Correspondence to be sent to: Evolutionary Biology and Ecology (EBE), Département de Biologie des Organismes, Université libre de Bruxelles (ULB), Av. Franklin D. Roosevelt 50, CP 160/12, Brussels B-1050, Belgium;
E-mail: catalina.rzpl@gmail.com.

Abstract.—Distinguishing coral species is not only crucial for physiological, ecological, and evolutionary studies but also to enable effective management of threatened reef ecosystems. However, traditional hypotheses that delineate coral species based on morphological traits from the coral skeleton are frequently at odds with tree-based molecular approaches. Additionally, a dearth of species-level molecular markers has made species delimitation particularly challenging in species-rich coral genera, leading to the widespread assumption that interspecific hybridization might be responsible for this apparent conundrum. Here, we used three lines of evidence—morphology, breeding trials, and molecular approaches—to identify species boundaries in a group of ecologically important tabular *Acropora* corals. In contrast to previous studies, our morphological analysis yielded groups that were congruent with experimental crosses as well as with coalescent-based and allele sharing-based multilocus approaches to species delimitation. Our results suggest that species of the genus *Acropora* are reproductively isolated and independently evolving units that can be distinguished morphologically. These findings not only pave the way for a taxonomic revision of coral species but also outline an approach that can provide a solid basis to address species delimitation and provide conservation support to a wide variety of keystone organisms. [*Acropora*; coral reefs; hybridization; reproductive isolation; taxonomy.]

A working coral taxonomy is crucial for meaningful physiological, ecological, and population genetic studies of these keystone organisms, as well as for the effective management and conservation of the ecosystems they support (Knowlton et al. 1992; Knowlton 2001). Even though climate and anthropogenic disturbances represent substantial threats to these ecosystems (Pandolfi et al. 2003; Carpenter et al. 2008; Hughes et al. 2017, 2018a), the taxonomy and systematics of some of the most vulnerable and diverse genera remain obscure (Fukami et al. 2004b; Richards et al. 2016). Such is the case for corals of the genus *Acropora*, the species of which are among those most affected by global warming (Marshall and Baird 2000; Carpenter et al. 2008; Hughes et al. 2018b). The genus is abundant on most reefs throughout the world's tropical and subtropical oceans and with more than 400 nominal species it is the most diverse extant reef-building coral genus (Wallace and Willis 1994; Wallace 1999). Consequently, ascertaining species boundaries in this genus will not only advance approaches to delineate species in corals but is also critical to understand the global response of coral reefs to climate change.

Distinguishing scleractinian coral species has always been a challenge, particularly in species-rich genera (Kitahara et al. 2016). The genus *Acropora* is emblematic of these difficulties: traditional morphological taxonomy

has been mired in confusion, best highlighted by the fact that only 122 of approximately 400 nominal species were considered valid in the most recent revision of the genus (Wallace et al. 2012). Similarly, standard molecular approaches based on genetic distances or species-level monophyly have failed to delineate species. Indeed, closely related *Acropora* morphospecies usually turn out to be interspersed in mitochondrial and nuclear gene trees (Odorico and Miller 1997; van Oppen 2001; Márquez et al. 2002; Suzuki et al. 2016), which has been widely interpreted as evidence for ongoing hybridization between coral species (the "syngameon" concept; Veron 1995).

However, there are alternative interpretations of such widespread nonmonophyly (Wallace 1999; Funk and Omland 2003; Miller and van Oppen 2003). Polyphyletic patterns observed in gene trees can result from the incorrect identification of specimens, which is highly likely in morphologically diverse groups with an intricate taxonomy (Funk and Omland 2003). In addition, the failure of alleles to sort after speciation can produce nonmonophyletic species with intraspecific distances as large as or even larger than interspecific distances (Flot et al. 2010). Such incomplete lineage sorting is more likely in species groups that have recently diversified (e.g., *Acropora* ∼6 Ma; Fukami et al. 2000), as well as in species with large effective population sizes. In such

cases, single-locus species delimitation approaches that require monophyly are bound to fail (Dellicour and Flot 2018).

Tabular morphospecies of *Acropora* have several features that make them an intriguing group on which to propose and validate novel taxonomic approaches (Wallace and Willis 1994): a high overall morphological similarity (Wallace 1999); the occurrence of multiple morphospecies in sympatry (Wallace 1985); and a substantial time overlap in gamete release across multiple described morphospecies (Harrison et al. 1984). *Acropora hyacinthus* (Dana 1846) is considered the epitome of tabular morphospecies and is regarded as the senior synonym for eight other nominal species (e.g., *A. bifurcata* Nemenzo 1971, *A. conferta* (Quelch 1886), *A. pectinata* (Brook 1892), *A. surculosa* (Dana 1846); Wallace 1999; Veron 2000), whereas genetic analyses suggest that it is a complex of several species (Ladner and Palumbi 2012; Suzuki et al. 2016). In addition to incongruence between molecular markers and morphological groupings (Márquez et al. 2002; Suzuki et al. 2016), the possibility of permeable cross-breeding barriers between morphospecies (as documented in *A. hyacinthus* vs. *A. cytherea* (Dana 1846)) casts further doubts on current species boundaries (Willis et al. 1997) and makes this group an exemplar system reflecting the challenges that affect coral taxonomy as a whole.

Traditional coral taxonomy is based on features of the skeleton that can confound species delimitation due to morphological plasticity, potential homoplasy, and cryptic diversity (Fukami et al. 2004b; Budd et al. 2010). However, morphology provides baseline information to identify primary species hypotheses that can be subjected to further analyses (PSHs; Puillandre et al. 2012). In addition, if the evidence supports such groups as independently evolving lineages, morphological analyses can help single out characters that are taxonomically informative (Wolstenholme et al. 2003). In this study, we compared three independent lines of evidence (i.e., morphology, breeding trials, and molecular approaches) to delineate species boundaries and assess hybridization in three sympatric tabular *Acropora* (Supplementary Fig. S1 available on Dryad at https://doi.org/10.5061/dryad.k98sf7m5x), inhabiting the outer reef of Sesoko Island, Okinawa, Japan (Fig. 1a).

Materials and Methods

Colony Sampling

In the days preceding the full moon of May 2018 (29/05/2018), fragments (~8 cm × 8 cm) from living tabular colonies (n=36) of reproductively mature *Acropora* (min. diameter >20 cm) were collected from the outer reef (26° 37′44″ N, 127° 51′44″ E) located south of the Tropical Biosphere Research Center (TBRC) at Sesoko Island (Okinawa, Japan). The reproductive condition of the colonies was assessed by breaking branches to expose developing oocytes (Harrison et al. 1984; Baird and Marshall 2002), and avoiding peripheral areas of the colony and tips of branches, as they usually present no gametes (Wallace 1985). Tissue samples (~2 cm^3) from each colony were preserved in a guanidium thiocyanate solution (4M guanidine thiocyanate, 0.1% N-lauroylsarcosine sodium, 10mM Tris-HCl pH 8, 0.1 M 2-mercaptoethanol; Fukami et al. 2004a) and alternatively in 95% ethanol for subsequent DNA extraction.

Morphological Taxonomy Assessment

Each colony was photographed in the field using an Olympus Tough TG-5 waterproof compact digital camera at the time of fragment collection (Olympus, Japan). After spawning, the fragments were bleached with a commercial solution of ~3–6% sodium hypochlorite (NaOCl) for morphometric assessment and then stored as vouchers at the Sesoko Station (specimen photos available on Morphobank Project 4065 at http://morphobank.org/permalink/?P4065). In addition to the collected specimens, 38 tabular *Acropora* skeletons deposited in the station from other field campaigns (2015, 2019) were also used for morphometric assessment (n=74 colonies in total, Data set S1—Morphological data available on Dryad). Qualitative and quantitative characters adapted from previous studies were recorded and measured from the coral skeletons (Supplementary Table S1 available on Dryad, see Wallace 1999; Wolstenholme et al. 2003; Wallace et al. 2012).

In order to provide a quantitative evaluation of the morphological taxonomic units (morphospecies), multivariate analyses of descriptive (qualitative), and morphometric (quantitative) characters were performed in R v3.6.2 (R Core Team 2018) through the Rstudio console v1.2.5033 (RStudio Team 2017). Qualitative characters along with categorized quantitative variables (Supplementary Tables S1 and S2 available on Dryad) were analyzed using hierarchical clustering analysis (HCA) with simple match coefficient distances (nomclust package, v2.1.4) and the Ward clustering method (cluster package, v2.1.0, Fig. 1b). Quantitative variables with a normal distribution and homogeneity of variance (Supplementary Table S2 available on Dryad) were analyzed using linear discriminant analysis (LDA) with the maximum likelihood (ML) estimator method (MASS v 7.3-51.5 and flipMultivariates v1.0.0), and a multivariate analysis of variance (MANOVA, stats package v3.6.2) to test for significant differences (Supplementary Fig. S2 a available on Dryad). Finally, using the complete morphological data set, a factor analysis of mixed data (FAMD) was performed (FactoMineR, v2.3 and factoextra v1.0.7) to identify morphological groupings supported by all features and to determine how much each variable contributed to the differentiation (Fig. 1c and Supplementary Fig. S2 b available on Dryad). Morphospecies (groups) obtained from this morphological assessment were used as primary species hypotheses (PSHs; Puillandre et al. 2012).
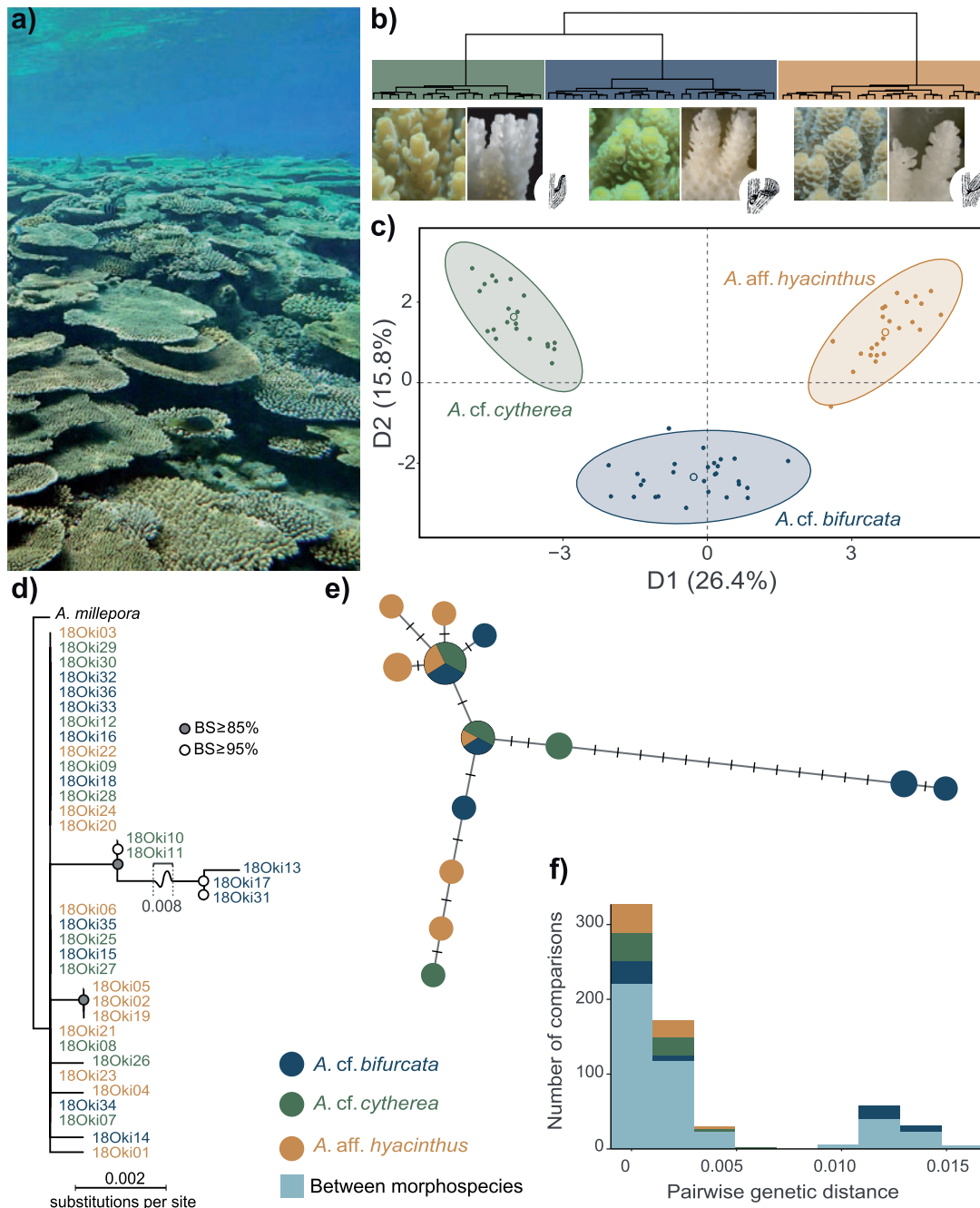
FIGURE 1. Morphology yields primary species hypotheses that are at odds with the mitochondrial phylogeny. a) Tabular *Acropora* at Sesoko Island outer reef (Okinawa, Japan); photo by A.H. Baird. b) Hierarchical clustering analysis (HCA, agglomerative coefficient = 0.95), along with the main morphological features that contributed to the differentiation for each morphospecies: color of colonies in the field (left) and shape and crowding of radial corallites along branches (right). c) Factor analysis of mixed data (FAMD) based on both qualitative and quantitative characters, distinguishing three morphospecies: *A.* cf. *bifurcata*, *A.* cf. *cytherea*, and *A.* aff. *hyacinthus*. See also Figure S2 and Table S1 available on Dryad for additional information. d) Maximum likelihood (ML) phylogeny of the mitochondrial control region (AcroCR) using ultrafast bootstrap. Branches with less than 85% of bootstrap support (BS) were collapsed. e) Haplotype network of the AcroCR region shaded according to morphospecies, with gaps recoded as single base changes. f) Histogram of the pairwise genetic distances of the AcroCR sequences within and between morphospecies.

Field Identification and Taxonomic Identity of the Morphospecies

Acropora bifurcata, *A. cytherea*, and *A. hyacinthus* were identified in the field following Veron (2000). The main field characters for each species are as follows; *A. hyacinthus* has tapered (gradually narrowing) branches with labellate (liplike) radial corallites with a flaring lip and colonies are orange-red; *A. cytherea*

TABLE 1.　A summary of the research into the taxonomic identity of the species used in the study

Nominal species, authority, accepted name (if different), type locality	Type material (ID) and current location	Type material vs. specimens in this study (ON qual)	Ongoing and future perspectives
Madrepora hyacinthus Dana 1846, *Acropora hyacinthus*, Fiji	Lectotype (USNM 246) designated by Wallace (1999), deposited at the NMNH, SI (Washington D.C., US)	Distinctive morphological differences between specimens and type material, particularly in branch shape and width ("aff.", affinity with a known species)	Ongoing molecular and morphological comparison to topotypes and other material from the Indo-Pacific
Madrepora cytherea Dana 1846, *Acropora cytherea*, Tahiti	Lectotype (USNM 423) designated by Wallace (1999), deposited at the NMNH, SI (Washington DC, USA)	Similar morphology, including that of the radial corallites ("cf.", specimens closely resemble type material, but this needs to be confirmed)	Ongoing morphological comparison to lectotypes and molecular comparison to other material from the Indo-Pacific
Acropora bifurcata Nemenzo 1971, Philippines	Holotype (UP C-1295), collected by Nemenzo (1971), deposited at UP, ZD (Quezon City, PH)	Geographical proximity to type location and similar morphology, including radial corallite shape ("cf.", specimens closely resemble type material, but this will require confirmation)	Ongoing molecular and morphological comparison to topotypes and other material from the Indo-Pacific

Catalog numbers from type material (ID) are depicted. Open nomenclature qualifiers (ON qual) were attributed according to the degree of uncertainty in identification following Sigovini et al. (2016) and Cowman et al. (2020): *affinis* (aff.) and *confer* (cf.). Refer to "Field Identification and Taxonomic Identity of the Morphospecies" in Materials and Methods for further information.
USNM = United States National Museum; NMNH = National Museum of Natural History; SI = Smithsonian Institution; UP = University of the Philippines; ZD = Zoology Department. Country codes: United States (US), Philippines (PH).

has terete (cylindrical) branches with labellate radial corallites with an extended outer lip and colonies are dark brown; *A. bifurcata* has terete branches with labellate radial corallites with a square lip and colonies are light brown (see images in Fig. 1b, color of colonies in the field and shape of radial corallites). Most local coral researchers would readily agree with the field identifications of *A. cytherea* and *A. hyacinthus*, however, *A. bifurcata* is not generally accepted as a valid species as it was considered a junior synonym of *A. hyacinthus* in the last major revision of the genus (Wallace 1999). In contrast, Veron (2000) accepted the species as valid but did not record it in Japan. However, field images in Nishihira and Veron (1995, see middle panel at p. 128) indicate that this species does occur in Japan but was identified by these authors as *A. hyacinthus*. Further information regarding type material and ongoing research into the taxonomic status of these species is presented below and summarized in Table 1.

Acropora hyacinthus (Dana 1846) has a type location in Fiji. A comparison of the colonies collected in this study to the lectotype designated by Wallace (1999, USNM 246; see http://n2t.net/ark:/65665/3fdf539df -6f98-4b91-a91c-53aa88a67457) indicates that there are significant differences in morphology. For example, the branches of the colonies at Sesoko are wider with more of a taper, which suggests that the species is distinct from that in Fiji. Consequently, the open nomenclature "aff." is used to indicate that the colonies found in Sesoko have affinities with *A. hyacinthus* but most likely belong to a distinct species.

Acropora cytherea (Dana 1846) has a type location in Tahiti. A comparison of the colonies from Japan to the lectotype designated by Wallace (1999; USNM 423, see http://n2t.net/ark:/65665/367cd18b6-2f69-4451-a32b-6 ae18bacd0ab) suggests that the species is morphologically similar to *A. cytherea*. In particular, colonies of both species have labellate radial corallites with an extended lip. Here we use the open nomenclature qualifier "cf." to suggest that this species is possibly *A. cytherea* but further information is required to confirm this assessment.

Acropora bifurcata Nemenzo 1971 has a type location in the Philippines. Given the proximity of Okinawa to the Philippines and the morphological similarity of the colonies to the holotype (UP C-1295, see http://www. coenomap.org/fact-sheet/acropora-bifurcata/), notably the labellate radial corallites with a squared margin, we used the open nomenclature qualifier "cf." to suggest the species is probably *A. bifurcata* but further information is required to confirm this.

Breeding Compatibility Experiments

Half of the collected colonies ($n = 18$) were kept in running seawater tanks and separated in individual buckets a few hours before the predicted time of spawning. Immediately after spawning, buoyant gamete bundles containing eggs and sperm were collected at the water surface of each container for the first two colonies that spawned from each morphospecies ($n = 6$). Once the eggs and sperm were separated, eggs were collected and
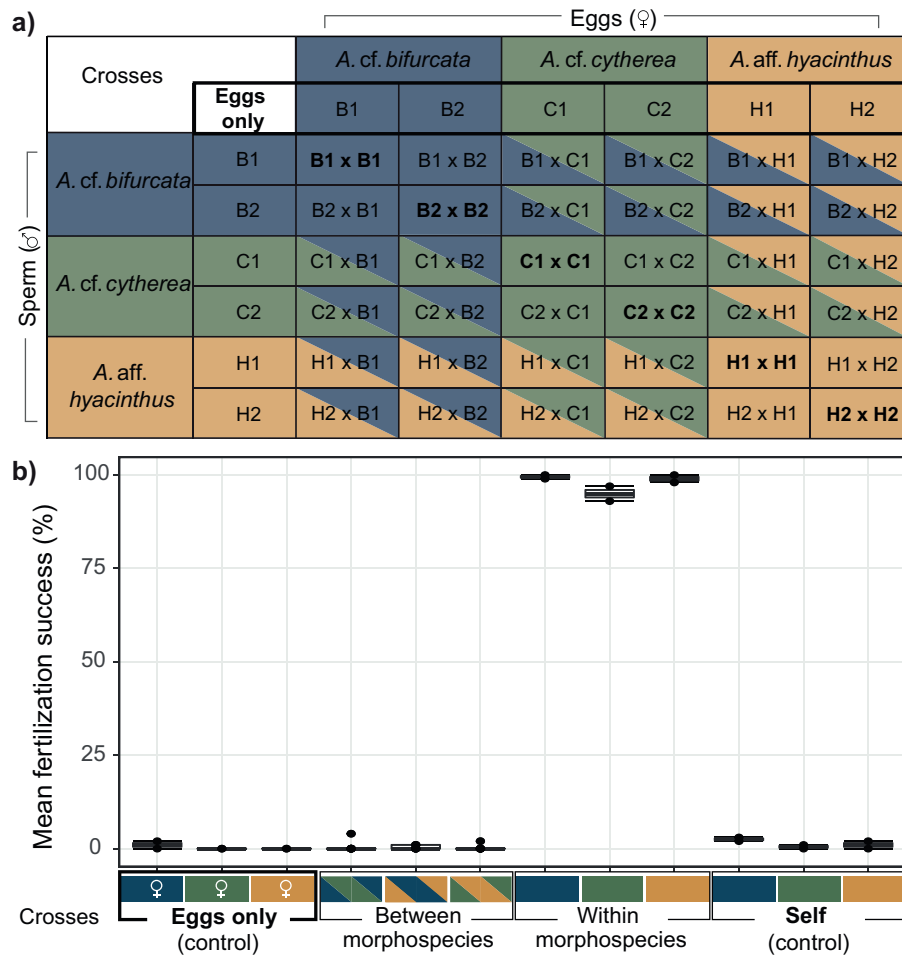
FIGURE 2. Cross-fertilization experiments suggest no hybridization potential. a) Gamete combinations (sperm × eggs) performed between representative colonies of each tabular *Acropora* morphospecies. b) Fertilization success (%) in each category of breeding trial, including the "eggs only" [top cells with bold border in a)] and the "self" fertilization controls [diagonal cells with bold font in a)] to account for sperm contamination and potential self-compatibility, respectively.

serially washed in 0.2 µm-filtered seawater to remove sperm and decrease the potential for self-fertilization. A portion of the eggs ("eggs only"—control) was kept aside in order to control for gamete separation and fertilization that may arise from leftover sperm in the eggs sample (Willis et al. 1997). The concentrated sperm obtained from the bundles was diluted approximately to 1:50 by adding filtered seawater before performing the crosses. In order to evaluate fertilization compatibility between the different morphospecies, approximately 100 washed eggs of each individual were added to each sperm dilution according to the breeding trial matrix (Fig. 2a, Data set S2—Breeding trials data available on Dryad).

Briefly, crosses were performed with gametes from 6 colonies for a total of $n = 6$ eggs only controls and $n = 36$ crosses: 6 self-control, 6 within morphospecies, and 24 between morphospecies, with at least two replicates for each combination. The numbers of regularly shaped embryos (prawn chip stage) and unfertilized eggs were counted under a stereomicroscope approximately 10 h after the breeding trials started. Mean fertilization success (%) was calculated as the average proportion of embryos divided by the number of embryos plus the remaining unfertilized eggs (Data set S2—Breeding trials data available on Dryad). Nonparametric Kruskal–Wallis rank sum test (stats package v3.6.2) was performed to test for significant differences in the mean proportion fertilized, and further posthoc tests (PMCMR v4.3 and PMCMRplus v1.4.4) were implemented in R (RStudio Team 2017; R Core Team 2018) to determine which particular crosses had significantly different fertilization success (Fig. 2b).

Preliminary Screening of Available Molecular Markers

To assess the species-level resolution of previously reported loci (Supplementary Table S3 available on Dryad), DNA was extracted from the 36 tissue samples preserved in guanidium thiocyanate solution using the NucleoSpin Tissue kit (Macherey-Nagel, Germany) and following the manufacturer's protocol. DNA integrity

was assessed on agarose gels (1%) and quality checked with a NanoDrop spectrophotometer (Thermo Fisher Scientific, USA). We used the primers and protocols detailed in Supplementary Table S4 available on Dryad to perform PCR-based amplification then Sanger sequencing of the mitochondrial putative control region (AcroCR) and two nuclear exon-primed intron crossing (EPIC) markers (Ladner and Palumbi 2012): a plasma membrane calcium-transporting ATPase (PMCA) and a frizzled-4 like homolog (FZD or exon 5491). Due to the relatively short span of these markers (545 and 639 bp, respectively), we redesigned primers to extend the product length of the FZD marker. For this purpose, we mapped FZD sequences previously obtained for tabular *Acropora* (Ladner and Palumbi 2012), to the available genome assemblies (see Supplementary Table S5 available on Dryad) of *Acropora digitifera* (Shinzato et al. 2011), *Acropora millepora* (Ying et al. 2019), *A. hyacinthus* (ReFuGe 2020 Consortium 2015; Liew et al. 2016), *Acropora cervicornis* and *Acropora palmata* (Kitchen et al. 2019) using Bowtie2 v2.3.4.3 (Langmead and Salzberg 2012) in local configuration. The unambiguously mapped contigs of each genome were recovered and converted to BAM files using SAMtools v1.9 (Li et al. 2009), then transformed into BED formatted files with BEDtools v2.26.0 (Quinlan and Hall 2010). The mapped regions in the BED files were extended at least 200 bp upstream and downstream, to be then recovered from the contig FASTA files using Seqkt v1.3 (Li 2013). Alignment between the extended mapped regions and FZD original sequences was performed using Mafft (E-INS-i method; Katoh et al. 2008). The consensus sequence for FZD (including ambiguities) was obtained from the alignment using SeaView v4.6.4 (Gouy et al. 2010) and used as a target to design primers using Primer3web v4.1 (Untergasser et al. 2012), by maximizing product length and allowing for a difference of 2° C in melting temperature between primers.

Sanger sequencing of the products was performed at GenoScreen (Lille, France). Sequencher v5.4.6 (GeneCodes, USA) was used to edit the chromatograms (Data set 3—Chromatograms available on Dryad). Multiple sequence alignments for each locus were generated using the E-INS-i method (Katoh et al. 2008) in the online implementation of Mafft v7.471 (available at https://mafft.cbrc.jp/alignment/server/; Katoh et al. 2008). For the mitochondrial putative control region (AcroCR), alignments were used directly for the downstream analyses. For the sequences obtained of the EPIC markers, two different complementary phasing approaches were used. Sequences of heterozygous individuals displaying alleles of the same length (without indel), were phased using SeqPHASE (Steps 1 and 2 available at https://eeg-ebe.github.io/SeqPHASE/; Flot 2010) PHASE v2.1.1 (Stephens et al. 2001; Stephens and Donnelly 2003). When length-variant heterozygotes were found in the data set, Champuru v1.0 (Flot et al. 2006; Flot 2007) was used to phase those sequences in a first step. Subsequently, they were inputted as "known haplotype pairs" during SeqPHASE's step 1, thereby contributing to the phasing of the other individuals. Allele pairs with posterior probability $\geq$ 0.9 were chosen, except when more than one possible pair with similar posterior probabilities was found. In such cases, alleles shared with the highest number of individuals or that were connected with the most frequent haplotypes in the network were selected.

Model-based genetic clustering of the phased EPIC sequences was performed using STRUCTURE v2.3.4 (Pritchard et al. 2000), with admixture model, correlated allele frequencies, and no prior. Implementing StrAuto v1.0 (Chhatre and Emerson 2017), values from 1 to 10 for the inferred number of populations (K) were used (20 runs per K, 250,000 burnin, 1,000,000 MCMC generations) to compute in parallel the probabilities of membership of each individual. Runs were further aligned, combined and finally merged using CLUMPP v1.1.2 (Jakobsson and Rosenberg 2007) and the Pophelper package v2.3.0 in R (Supplementary Fig. S3 a,b available on Dryad; Spöri and Flot 2020; R Core Team 2018). Various species delimitation approaches were performed.

For the allele sharing-based approach (Flot et al. 2010), the EPIC markers phased sequences were input directly into the online program HaplowebMaker (available at https://eeg-ebe.github.io/HaplowebMaker/; Spöri and Flot 2020), from which haplowebs and the corresponding putative species or fields for recombination (FFRs; Doyle 1995) were obtained (Supplementary Fig. S3 c available on Dryad). For the distance-based approach, the best model of evolution was identified using the Bayesian information criterion (BIC) value criterion in ModelFinder (Kalyaanamoorthy et al. 2017). After converting the DNA alignments to bins using *fasta2DNAbin* (adegenet package v2.1.2), pairwise genetic distances using the closest available model to the best BIC score list were computed by *dist.dna* function (ape package v5.3) and histograms were plotted using ggplot2 v3.3.0.9 in R. Further phylogenetic analyses were performed under maximum likelihood (ML) with IQ-TREE v2.0.3 (Nguyen et al. 2015), using 1000 ultrafast bootstrap replicates (-B 1000) and an additional step to optimize trees by nearest-neighbor interchange (-bnni; Minh et al. 2013). Branches of the consensus trees with nodes with less than 85% of bootstrap support were collapsed using *multi2di* function (ape package v5.3). Trees (Supplementary Fig. S3 e available on Dryad) were visualized and formatted using FigTree v1.4.4 (Rambaut 2018). Genomic regions of *A. millepora* (Ying et al. 2019) that mapped to each nuclear loci with Bowtie2 v2.3.4.3 (–local; Langmead and Salzberg 2012), were used as outgroup for the phylogenies. For AcroCR, the closest *A. millepora* match found using megaBLAST (against the nr/nt database, available at https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE_TYPE=BlastSearch; Altschul et al. 1990) with GenBank accession number KY408102.1 was used for that purpose instead (100% query coverage, 99.85% identity and E-value=0).

Overall, the mitochondrial AcroCR (Fig. 1d–f) and the nuclear EPIC markers (Supplementary Fig. S3 available on Dryad) from the literature did not provide enough resolution at the species level. Model-based genetic clustering of the two EPIC markers was only able to recover two clusters (Supplementary Fig. S3 a available on Dryad), neither congruent with the primary species hypotheses (PSHs) inferred from morphological species delimitation (Fig. 1b,c), nor with the fertilization success in breeding trials (Fig. 2b). Similarly, haplowebs obtained from these markers (Supplementary Fig. S3 c available on Dryad) were not able to resolve them. Each morphospecies showed some private alleles but shared alleles connected individuals from different morphospecies into single fields for recombination (FFRs). Pairwise genetic distances and gene trees did not recover groups congruent with the other lines of evidence either (Supplementary Fig. S3 d,e available on Dryad). For these reasons, we explored target enrichment followed by high-throughput sequencing to assess more accurately the species boundaries in this case study and to target for loci with enough resolution at the species level that could be amplified in a larger data set.

Target-Enrichment Using the Scleractinian Bait Set

To find molecular markers that provide better resolution at the species level (Supplementary Table S3 available on Dryad), we performed target-capture sequencing for nine of the samples ($n=9$, three from each morphospecies) preserved in 95% ethanol (Supplementary Table S6 available on Dryad). DNA was extracted, its quality assessed and then sent to Arbor Biosciences (Ann Arbor, USA) for library preparation (following Quattrini et al. 2018) and target capture sequencing (detailed in Cowman et al. 2020). For target enrichment of conserved elements (derived from exonic loci and ultraconserved elements (UCEs); Faircloth et al. 2012), we implemented a new set of baits (Cowman et al. 2020), that was redesigned from a set that originally targeted anthozoans (Quattrini et al. 2018). The new bait set targets hexacorallians (hexacoral-v2 bait set, scleractinian subset—2,476 target loci) and has been successfully tested in a comprehensive sample of acroporids (Cowman et al. 2020). Demultiplexing, trimming, and assembly were performed according to the parameters and software previously tested in *Acropora* (Cowman et al. 2020). Subsequently, the contigs assembled for the nine tabular samples (Supplementary Table S4 available on Dryad) were matched to the baits employing PHYLUCE (Faircloth 2016) with default parameters (*phyluce_assembly_match_contigs_to_probes*). As a result, 2060 loci (1026 exons and 1034 UCEs) were extracted into FASTA (*phyluce_assembly _get_match_counts* & *phyluce_assembly_get_fastas_from_match_counts*) to proceed with allele phasing using two different pipelines, described in the following sections (see Supplementary Table S3 available on Dryad for a summary).

Genetic Clustering and Preliminary Species Trees Using the Target-Enrichment Data Set

First, to generate a broad subset of loci that could be used to evaluate genetic clustering and estimate a preliminary species tree, loci were aligned (*phyluce_align_ seqcap_align –incomplete -matrix –no-trim –aligner mafft*) and globally trimmed using Gblocks (Castresana 2000; Talavera and Castresana 2007) with default parameters (*phyluce_align_get_gblocks_trimmed_alignments_from_ untrimmed*). Phasing of the aligned loci was performed following the phase_everyone v0.1 or "Laninsky" pipeline (Baca et al. 2017; Alexander 2018a). Once alleles were obtained, they were aligned and processed following Steps 5–8 from the pipeline reference_ aligning_to_established_loci v0.0.3 (Baca et al. 2017; Alexander 2018b). Then, single nucleotide polymorphisms (SNPs) were pulled out from each locus and filtered to ensure that only loci with data for at least one individual per morphospecies were included in the downstream analyses. SNPs for the resulting 1889 loci (1022 exons and 867 UCEs) were used to perform a STRUCTURE analysis (Pritchard et al. 2000) parallelized through StrAuto (Chhatre and Emerson 2017), with K values from 1 to 9, admixture model, correlated allele frequencies and no prior (20 runs per K, 250,000 burnin, 1,000,000 MCMC generations). CLUMPP (Jakobsson and Rosenberg 2007) and pophelper package tutorial (available at http://www.royfrancis.com/pophelper/articles) were used to align, combine and merge the runs. Evanno ΔK plots (Evanno et al. 2005) were used to determine the most likely number of clusters (K, Fig. 3a). Using ggplot2 v3.3.0.9 in R, the corresponding bar plots depicting the probability of individual membership to each cluster were obtained for the suggested K values ($K=3$ or $K=5$), from which K=3 depicted better stratification of the samples according to their allele frequencies and suggested that there does not seem to be population structure within the putative species (Fig. 3b).

In addition, the most likely species tree was estimated with SNAPP v1.5.1 (Bryant et al. 2012) through the CIPRES gateway (Miller et al. 2010). SNPs were extracted from the concatenated FASTA of a subset of 210 loci present in all the samples (128 UCEs and 82 exons) using *fasta2DNAbin* (adegenet package v2.1.2) and storing them in a Nexus file using the *write.nexus.data* function (ape package v5.3). This file was used to create the XML input file in the Bayesian Evolutionary Analysis Utility (BEAUti) v2.6.3 (Bouckaert et al. 2014, 2019). Five independent runs of BEAST were performed with MCMC length of 10,000,000, preburnin of 100,000, sampling frequency of 1000, and default model parameters. Output trees and log files were combined using LogCombiner v2.6.3 (Bouckaert et al. 2014, 2019). After 10% burnin, combined logs were input into Tracer v1.7.1 (Rambaut et al. 2018) to check MCMC convergence and effective sample sizes (ESS) > 200. TreeAnnotator v2.6.2 (Bouckaert et al. 2014, 2019) was used to generate maximum clade credibility trees and DensiTree v2.2.7
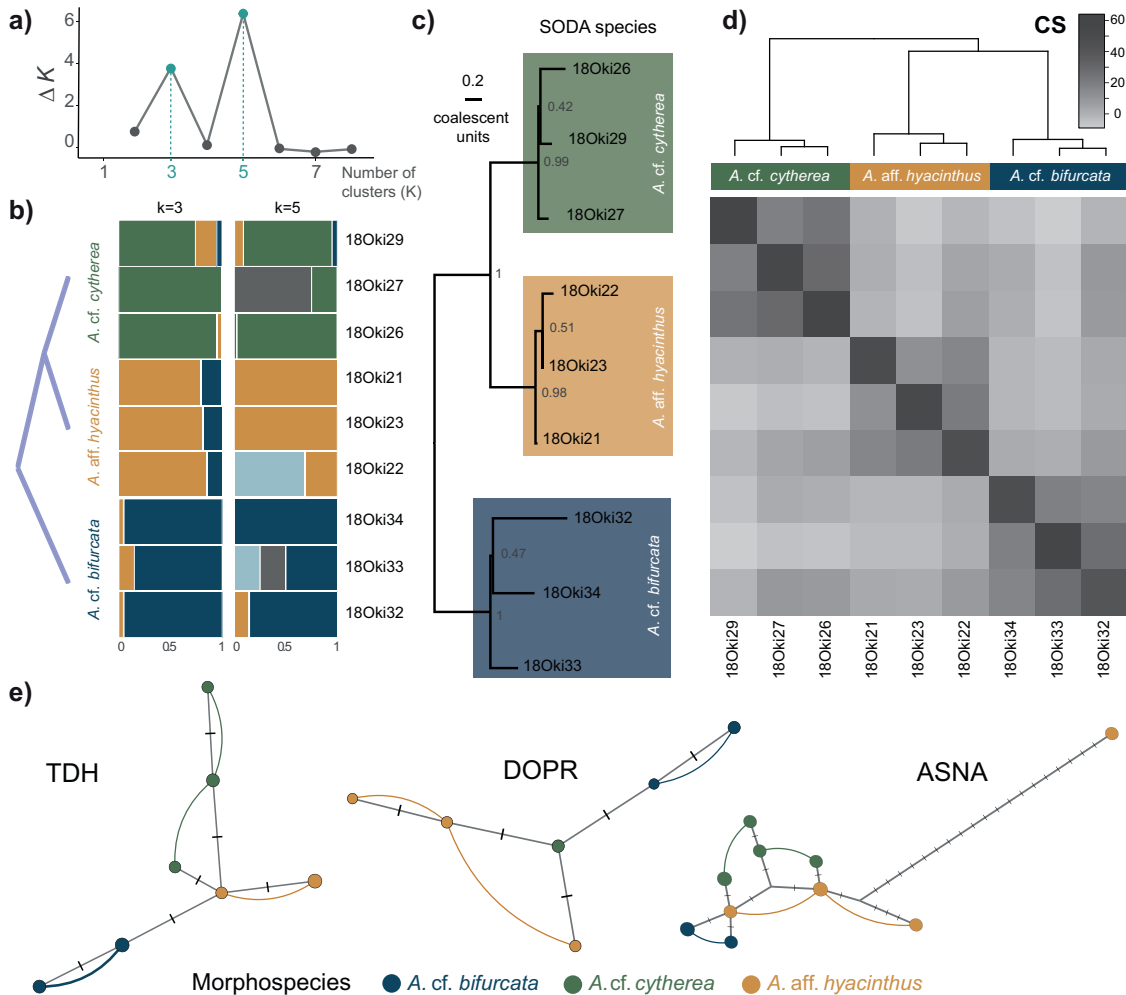
FIGURE 3. Screening of target capture-derived markers. a) Evanno ΔK plot (above) depicting two possible optimal cluster (K) values (dashed lines) and bar plots (below) displaying the individual probability of membership assigned using model-based clustering for each K value (1889 loci). The most frequent SNAPP tree (using 210 loci present in all samples) is depicted on the left side of the plots. b) ASTRAL resolved extended species tree with phased sequences according to the molecular species delineated by SODA, where alleles were mapped to individuals and nodes with less than 10% of local posterior probability (LPP) or low branch support were collapsed. c) Conspecificity score (CS) matrix for a subset of 79 target-enrichment sequenced loci used to perform a preliminary allele sharing-based species delimitation. d) Haplowebs of three loci displaying putative species delimitation under mutual allelic exclusivity criterion, congruent with model-based genetic clustering, species trees and the primary species hypotheses (PSHs) based on morphology and breeding trials.

(Bouckaert 2010) to plot the corresponding consensus tree (Fig. 3b, left).

Estimation of a Resolved Extended Species Tree Using the Target-Enrichment Data Set

Loci were first aligned and edge trimmed (*phyluce_align_seqcap_align* –taxa 9 –incomplete-matrix) using PHYLUCE (Faircloth 2016). Subsequently, following the phasing tutorial (available at https://phyluce.readthedocs.io/en/latest/tutorial-two.html; Andermann et al. 2019), loci were phased into alleles for each individual. Allelic sequences were aligned (*phyluce_align_seqcap_align* –no-trim –ambiguous – incomplete-matrix) and globally trimmed (*phyluce_ align_get_gblocks_trimmed_alignments_from_ untrimmed*).

To remove sequences with unphased bases (N) that could cause problems in downstream analyses, loci alignments were further screened and filtered (*phyluce_align_screen_alignments_for_problems*). The resulting subset of 79 loci (TC79loci hereafter) was used to perform species delimitation from the data available for the target-enriched samples. To achieve this, we estimated a resolved extended species tree using the frequency of the quartet topologies of the individual gene trees build from the phased loci alignments. However, instead of mapping individuals to species as in an extended species tree (Rabiee et al. 2019), the resulting guide tree was obtained by mapping alleles to individuals.

Similar to the preliminary screening, IQ-TREE (Nguyen et al. 2015) was implemented to obtain individual ML trees from the phased FASTA alignments

obtained from the TC79loci data set. Those trees were used as input to run ASTRAL-III v5.7.3 (Zhang et al. 2018; Rabiee et al. 2019) and to estimate a resolved extended species tree following the ASTRAL tutorial (available at https://github.com/smirarab/ASTRAL/blob/master/astral-tutorial.md#running-astral). After pruning branches with low support or local posterior probability (LPP) < 10% (Junier and Zdobnov 2010), the gene trees were used to generate a resolved extended species tree without constraining each morphospecies to be monophyletic and incorporating a mapping file that assigned each allele sequence to an individual (-a option, Fig. 3c). Moreover, we used the same data set to perform multilocus species delimitation using quartet frequencies implementing Species bOundary Delimitation using Astral (SODA) v1.0.1 (Rabiee and Mirarab 2020) with the default alpha (α) threshold of 0.05 (Fig. 3c, right).

Screening for Loci with Species-Level Resolution in the Target-Enrichment Data Set

To screen for markers providing resolution at species-level, we used the mutual allelic exclusivity criterion to define species boundaries. This criterion is always met before or at the same time as reciprocal monophyly; thereby it provides a more sensitive criterion to delineate species (Flot et al. 2010). Consequently, allele sharing-based species delimitation was performed on the TC79loci data set using both haplowebs and the corresponding conspecificity matrix (Debortoli et al. 2016) obtained using the online programs HaplowebMaker and CoMa (Spöri and Flot 2020). There, a conspecificity score (CS) was calculated for each pair of individuals by subtracting the number of markers/loci that do not support them being conspecific (H, different species or heterospecific) from the number of markers/loci for which they are considered conspecific (C, same species or partition) [$CS = C\text{-}H$]. According to these scores, the matrix was then clustered and plotted (Fig. 3d) using the R package heatmap3 v1.1.7 with the Ward agglomeration method from the *hclust* function (stats package v3.6.2). Loci with at least one individual per genetic cluster (as identified in STRUCTURE) were kept, and their corresponding haplowebs were individually explored to assess their congruence with the conspecificity matrix, and the primary species hypotheses (PSHs) inferred from the morphological assessment and supported by the breeding trials. We selected loci with haplowebs depicting partitions (FFRs) congruent with the PSHs, and that provided resolution (genetic clusters containing different PSHs did not lump in the same FFR) even when gaps were considered as missing data. From this reduced subset, three loci with different variability degrees were chosen as candidate regions to develop markers at species level (2 exons and 1 UCE loci, Fig. 3e). GenBank megaBLAST (Altschul et al. 1990) searches were implemented (nr/nt database) to find the closest annotated match for each locus and code them accordingly (TDH, DOPR, and ASNA, see Supplementary Table S7 available on Dryad).

Developing Target-Enrichment Derived Markers for Larger Data Sets

To delineate species boundaries without resorting to high-throughput techniques, we used an identical approach to that previously employed to extend the length of the EPIC markers (see primer redesign performed in the section *Preliminary Screening of Available Molecular Markers* and Supplementary Table S5 available on Dryad). From the DNA extracted of the 36 tissue samples of *Acropora* preserved in the guanidium thiocyanate solution (Fukami et al. 2004a), PCR-based amplification followed by Sanger sequencing of the three target-enrichment derived loci was performed (Supplementary Tables S3 and S4 available on Dryad). Sequences obtained from GenoScreen (Lille, France) were processed and phased as for the preliminary screened EPIC markers (PMCA and FZD).

In a first step, genetic clustering, potential population structure, and admixture within the sympatric putative species was assessed using model-based genetic clustering for the derived target-capture markers (TDH, DOPR, and ASNA). The corresponding ΔK plot and the bar plots to evaluate individual probability membership were performed using K = 1 – 10 and the same parameters as before (Fig. 4a,b). Additionally, to detect clusters based on genetic similarity and without relying in evolution models, a discriminant analysis of principal components (DAPC; Jombart et al. 2010) was completed using the package adegenet v2.1.2 in R (following Quattrini et al. 2019; Fig. 4c,d).

Molecular Delineation of Species Boundaries Using the Target-Enrichment Derived Markers

Sampling pattern, speciation rate, species richness, mutation rate, and effective population size tend to exert widely different effects and biases onto species delimitation methods (Dellicour and Flot 2018). To overcome these issues, we performed different approaches to delineate species boundaries in this tabular *Acropora* study case. Distributions of pairwise genetic distances were first evaluated in search of a barcode gap (see Supplementary Fig. S4 available on Dryad, right). As such, distance-based approaches might not work for recently diverged species, on which intraspecific distances may not be substantially smaller than interspecific ones. Consequently, we also used haplowebs and their corresponding conspecificity matrices to delineate species under the mutual allelic exclusivity criterion (Fig. 4e, Supplementary Table S8 available on Dryad). Instead of taking into account the genetic distances, such allele sharing-based approaches aggregate individuals based on the haplotypes they share, providing a more sensitive criterion to delineate closely related species (Flot et al. 2010).
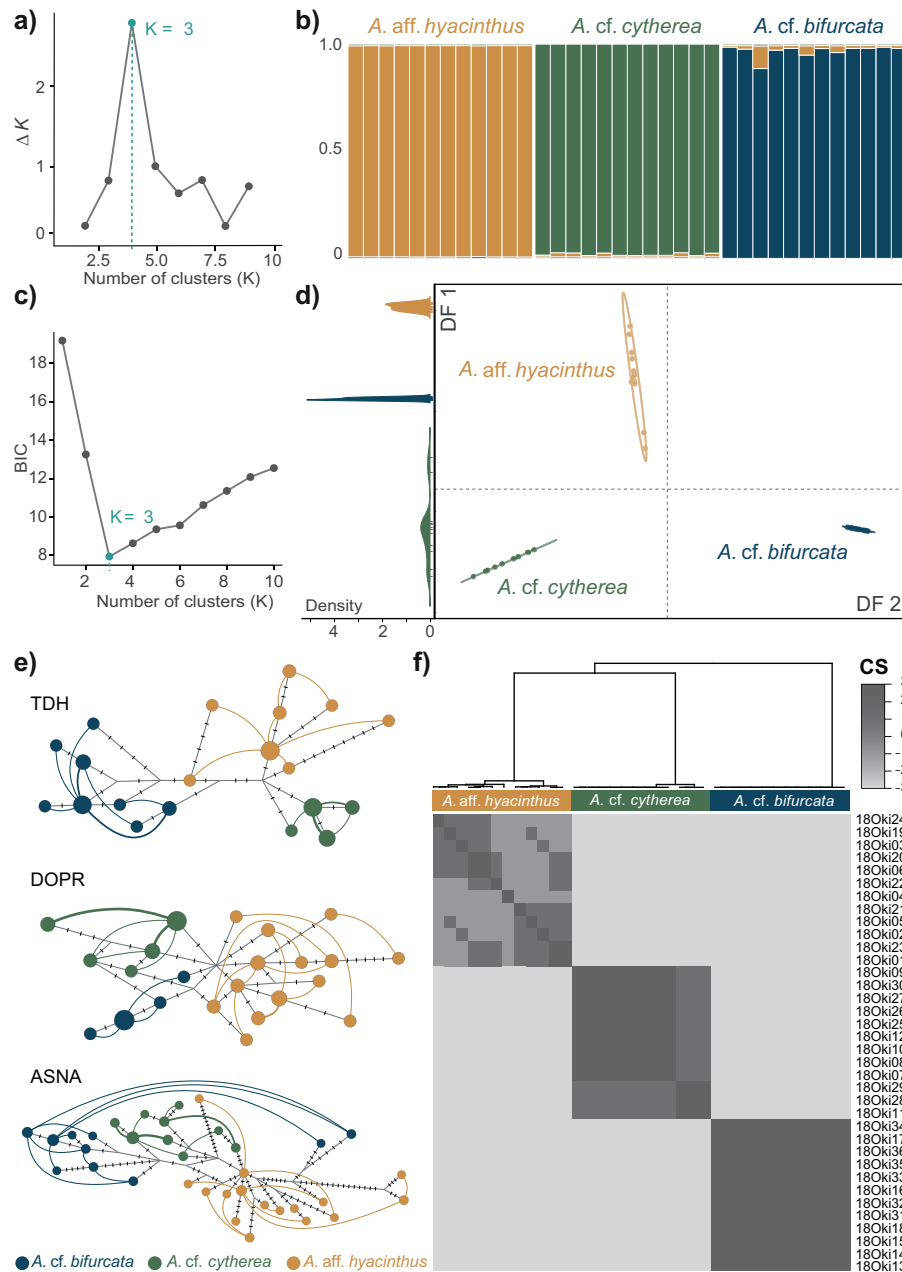
FIGURE 4.   Molecular evidence supports the primary species hypotheses in *Acropora* corals. a) Evanno ΔK plot highlighting the most likely number of genetic clusters. b) Bayesian model-based genetic structure plot depicting the probability of individual membership to each cluster when $K = 3$. c) Optimal cluster number for the Discriminant analysis of principal components (DAPC) according to the Bayesian Information Criterion (BIC) statistic with the most likely K value highlighted ($K = 3$). d) DAPC scatterplot depicting clustering based on genetic similarity among the individuals using two discriminant functions (DF). e) Haplowebs delineating putative species based on the co-occurrence of alleles for each one of the nuclear markers defined from target-capture sequencing, coded according to morphospecies. f) Conspecificity score (CS) matrix summarizing the fields for recombination (FFRs) found using the allele sharing-based approach to delineate species with the three target-capture defined loci. The conspecific groups delineated by the FFRs of the three markers are congruent with the morphospecies and with the results from breeding trials (see Figs. 1 and 2).

To evaluate species boundaries under the reciprocal monophyly criterion, maximum likelihood phylogenies of individual (left in Supplementary Fig. S4 available on Dryad) and concatenated genes (Supplementary Fig. S5 a available on Dryad) were performed on the target-capture derived loci as described for the preliminary

screening of available molecular markers. Additionally, the CIPRES gateway (Miller et al. 2010) was used to perform SNAPP and estimate the posterior distribution of trees from the SNPs extracted from the three loci. Independent runs of BEAST v2.6.3 (Bouckaert et al. 2014, 2019) were performed with MCMC length of

10,000,000, preburnin of 100,000, sampling frequency of 1000, and default model parameters. After 10% burnin, the output trees and log files were combined and examined for MCMC convergence. A cloudogram depicting the most frequently recovered species trees with individuals as terminal tips (Supplementary Fig. S5 b available on Dryad) was generated from this analysis. In addition, individual ML trees obtained from the phylogenetic analyses performed with IQ-TREE, were used in a resolved extended species tree estimation using ASTRAL on the three loci, both when constraining each morphospecies to be monophyletic (top left inset in Supplementary Fig. S5 c available on Dryad), and without such constraint (main resolved extended species tree in Supplementary Fig. S7 c available on Dryad). In both cases, alleles were mapped to individuals to obtain the final tree (using ASTRAL's -a option). Due to the small number of loci, species delimitation with SODA was not performed on this data set.

To test for alternative species models a SNAPP coalescence-based analysis was performed (Supplementary Table S9 available on Dryad). The alternative models tested were: 1) a single species-model that includes individuals from the three morphospecies, in one complex; 2) the two species-model supported by the current taxonomy in which *A. hyacinthus* and *A. cytherea* are considered different species but *A. bifurcata* is a synonym of the former; and 3) a three species-model (*A. hyacinthus*, *A. cytherea*, and *A. bifurcata*), supported by the morphological and breeding trial approaches from this study. Five runs of SNAPP were performed using BEAST, with 48 path sampling steps, 100,000 MCMC and 10,000 of preburnin (following Herrera and Shank 2016; Quattrini et al. 2019). Finally, ranking of the models was performed using Bayes factor delimitation (BFD; Grummer et al. 2014; Leaché et al. 2014) by comparing the marginal likelihood estimates (MLE) obtained for each model by calculating the Bayes factor (BF; Kass and Raftery 1995) between the current taxonomy model (model 1, i.e., two accepted species) and the alternative species models (model x) , as suggested in the tutorial (BF= 2 * [model 1 - model x]; Leaché and Bouckaert 2018).

Additionally, a joint Bayesian analysis of species delimitation and species tree estimation was performed using Bayesian Phylogenetics and Phylogeography v4.2 (BPP; Yang 2015). We performed the A11-type analysis (Flouri et al. 2020), using Phylip alignments for each target-capture derived loci obtained with the *fas2phy* function of the R package chopper v0.1.8. BPP was run for 200,000 generations, with a burnin of 20,000, and a sample frequency of 1 (following McFadden et al. 2017). Comparison of replicate runs performed with each rjMCM algorithm, different starting tree topologies and initial seeds was performed to assess overall convergence. The influence of prior distributions of the ancestral population size (θ) and root age (τ_0),

was evaluated under three scenarios (similar to Leaché and Fujita 2010): 1) large ancestral population size and deep divergence, 2) small ancestral population size and shallow divergence, and 3) large ancestral population size and shallow divergence among species (Supplementary Table S10 available on Dryad).

RESULTS AND DISCUSSION

Morphology Yields Primary Species Hypotheses

We first examined our collected specimens for a series of morphological characters (Tables S1, S2 and Data set S1 available on Dryad; Wallace 1999; Wolstenholme et al. 2003; Wallace et al. 2012). Multivariate analysis clearly distinguished three morphospecies (Fig. 1b,c and Supplementary Fig. S2a available on Dryad; $n = 74$, $P \le 0.001$), tentatively identified by comparison with the relevant type material as *A.* cf. *cytherea*, *A.* aff. *hyacinthus* and *A.* cf. *bifurcata* (see Table 1, Wallace 1999; Veron 2000; Wallace et al. 2012). The main features that contributed to the discrimination achieved by this analysis were the color of the colonies in the field, the shape, and extent of crowding of the radial corallites (Fig. 1b) and the median length and width of the branches (Supplementary Fig. S2b available on Dryad).

Mitochondrial Marker Analyses are at Odds with Morphology

As in previous studies of the genus *Acropora* (van Oppen 2001; Márquez et al. 2002), neither maximum-likelihood phylogeny (Fig. 1d) nor pairwise genetic distances (Fig. 1e) obtained from the mitochondrial putative control region (AcroCR) recovered groups congruent with the morphological analyses. Instead, specimens from the three morphospecies were scattered throughout the tree, a pattern that may result from incorrect identification of the colonies (caused for instance by morphological stasis or by phenotypic plasticity), incomplete lineage sorting, or hybridization (Funk and Omland 2003).

Due to considerable overlap in the time of spawning among *Acropora* species (Harrison et al. 1984; Baird et al. 2009) and their high rates of interspecific breeding in vitro (Willis et al. 1997), hybridization has often been evoked as the most likely cause for the lack of species-level monophyly in this genus (Miller and van Oppen 2003; Ying et al. 2019). However, in groups with relatively recent diversification and significant population size, such as the *A. hyacinthus* species group (~2.58 Ma; Wallace 1999), shared ancestral polymorphisms caused by large expected coalescent time should be considered as an alternative explanation. Distinguishing among these competing hypotheses requires several independent markers, which is impossible using only mitochondrial sequences (Sang and Zhong 2000).

Cross-fertilization Experiments Suggest no Hybridization Potential

In such a situation, breeding trials not only supply an important layer of biologically relevant information for delimiting sympatric species but also provide a litmus test to assess hybridization potential based on *in vitro* fertilization success (Wallace and Willis 1994). Consequently, we evaluated mating compatibility by performing cross-fertilization experiments using representative colonies from each of the three morphospecies (Fig. 2a). Significant fertilization success only occurred in crosses performed within morphospecies (Kruskal–Wallis *chi-squared* $= 23.26$, $df = 3$, $P = 3.565\text{e-}05$), whereas all the other crosses resulted in almost no fertilization (Fig. 2b). The reproductively isolated groups delineated using this approach comprised only individuals of the same morphospecies, thereby supporting the boundaries inferred from morphology.

Molecular Evidence Supports the Primary Species Hypotheses

Since breeding compatibility experiments can only be performed between colonies that reproduce synchronously or within a few hours of difference (Willis et al. 1997), we extended the scope of the cross-fertilization trials by looking at patterns of genetic clustering and allele sharing, i.e. using genetic similarity and mutual allelic exclusivity as indirect evidence for reproductive isolation (Supplementary Table S3 available on Dryad). Molecular approaches stemming from high-throughput techniques have recently overcome long-standing methodological limitations of molecular studies such as the small number of markers available and lack of species-level resolution (Cowman et al. 2020; Erickson et al. 2021). Here, three individuals per morphospecies ($n = 9$) were analyzed applying an enrichment procedure designed to capture conserved elements (derived from UCEs and exonic loci) with a set of baits targeting hexacorals (Quattrini et al. 2018; Cowman et al. 2020). Using this approach, more than two thousand phased loci were recovered (1026 exons and 1034 UCEs, Supplementary Table S6 available on Dryad).

Model-based genetic clustering using STRUCTURE (Fig. 3a,b), as well as an ASTRAL resolved extended species tree (Fig. 3c) of subsets of these loci (1889 and 79 loci, respectively), identified groups that were consistent with both morphology and breeding trials. To verify this across a larger number of specimens, we screened the captured loci for candidate markers displaying allelic exclusivity for each cluster (79 loci, Fig. 3d). As a result, three nuclear loci—L-threonine 3-dehydrogenase (TDH), dopamine receptor 2 (DOPR), and ATPase ASNA-1 (ASNA) (Fig. 3e)—were selected for PCR-based amplification of the 36 individuals in the tabular *Acropora* data set followed by various molecular species delimitation approaches (Supplementary Tables S3, S4, and S7 available on Dryad).

Genetic clustering of the specimens ($n = 36$) differentiated three groups that were congruent with both morphospecies hypotheses and breeding compatibility results (Fig. 4a–d). As previously observed with the mitochondrial control region, the pairwise genetic distances between and within morphospecies overlapped for each marker (Supplementary Fig. S4 available on Dryad, right), and neither the individual gene phylogenies (Supplementary Fig. S4 available on Dryad, left), a concatenated tree (Supplementary Fig. S5 a available on Dryad) nor a cloudogram (Supplementary Fig. S5 b available on Dryad) inferred from these loci supported the reciprocal monophyly of the three species.

By contrast, each of the three species was recovered as monophyletic in the resolved extended species tree obtained using ASTRAL (Supplementary Fig. S5 c available on Dryad), albeit with uncertain topology and low support for some clades. In addition, the haplowebs inferred from these three loci (Fig. 4e) and the conspecificity matrix summarizing them (Fig. 4f and Supplementary Table S8 available on Dryad) all unequivocally supported the grouping of our samples into three reproductively isolated units. Similarly, coalescence-based (Supplementary Table S9 available on Dryad) and Bayesian species delimitation analyses (Supplementary Table S10 available on Dryad) supported the three-species model with decisive values (Bayes factor >10 and posterior probability > 0.95, respectively).

These results challenge the generally accepted idea that morphospecies of *Acropora* cannot be distinguished using molecular approaches because of hybridization. On the contrary, despite being closely related these species appear to be reproductively isolated. It was possible to delineate them using target-enrichment followed by genomic sequencing (which probes thousands of markers but can yield incomplete data matrices) as well as using traditional PCR amplification followed by Sanger sequencing (which targets only one marker/individual at a time but yields high-quality, complete data sets). Hence, our results are different from other examples of successful molecular species delimitation based exclusively on high-throughput genomic sequencing (Quattrini et al. 2019; Erickson et al. 2021).

CONCLUSIONS

By using approaches sensitive enough to detect divergence at both the morphological and molecular levels, congruence between the three lines of evidence (i.e., morphology, breeding trials, and molecular approaches) demonstrates that it is possible to develop a robust coral taxonomy, thus helping to solve one of the greatest taxonomical conundrums since Linnaeus (Kitahara et al. 2016). Comparing evidence from multiple independent sources improved confidence in coral species boundaries by illustrating that *Acropora* species, once considered a taxonomic nightmare, are actually reproductively isolated and independently evolving units that can be distinguished morphologically.

Our findings show that allele sharing-based and coalescence-based multilocus approaches to species delimitation outperform mainstream methodologies relying on the monophyly and genetic distance as the criteria to delineate boundaries, particularly between closely related species. Although our methodology was focused on the taxonomic revision of coral species, the approaches outlined here are in principle applicable to a wide variety of plant and animal taxa.

SUPPLEMENTARY MATERIAL

Data available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.k98sf7m5x

The main sequence data sets generated for this study have been placed in GenBank and SRA repositories (see Supplementary Tables S4 and S5 available in Dryad). All photographical records of the specimens used for this study have been deposited in MorphoBank (Project 4065, http://morphobank.org/permalink/?P4065). Alignments, trees and examples of scripts and commands used can be found in the GitHub repository (https://github.com/catalinarp/SpeciesDelimitationTabularAcropora).

REFERENCES

Alexander A. 2018a. phase_everyone v0.1. Available from: https://github.com/laninsky/reference_aligning_to_established_loci/edit/master/phase_everyone.

Alexander A. 2018b. Reference_aligning_to_established_loci v0.0.3. Available from: https://github.com/laninsky/reference_aligning_to_established_loci.

Altschul S.F.S.F., Gish W., Miller W., Myers E.W.E.W., Lipman D.J.D.J. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403–410.

Andermann T., Fernandes A.M., Olsson U., Töpel M., Pfeil B., Oxelman B., Aleixo A., Faircloth B.C., Antonelli A. 2019. Allele phasing greatly improves the phylogenetic utility of ultraconserved elements. Syst. Biol. 68:32–46.

Baca S.M., Alexander A., Gustafson G.T., Short A.E.Z. 2017. Ultraconserved elements show utility in phylogenetic inference of Adephaga (Coleoptera) and suggest paraphyly of 'Hydradephaga.' Syst. Entomol. 42:786–795.

Baird A.H., Guest J.R., Willis B.L. 2009. Systematic and biogeographical patterns in the reproductive biology of scleractinian corals. Annu. Rev. Ecol. Evol. Syst. 40:551–571.

Baird A.H., Marshall P.A. 2002. Mortality, growth and reproduction in scleractinian corals following bleaching on the Great Barrier Reef. Mar. Ecol. Prog. Ser. 237:133–141.

Bouckaert R., Heled J., Kühnert D., Vaughan T., Wu C.H., Xie D., Suchard M.A., Rambaut A., Drummond A.J. 2014. BEAST 2: a software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 10:1–6.

Bouckaert R., Vaughan T.G., Barido-Sottani J., Duchêne S., Fourment M., Gavryushkina A., Heled J., Jones G., Kühnert D., De Maio N., Matschiner M., Mendes F.K., Müller N.F., Ogilvie H.A., Du Plessis L., Popinga A., Rambaut A., Rasmussen D., Siveroni I., Suchard M.A., Wu C.H., Xie D., Zhang C., Stadler T., Drummond A.J. 2019. BEAST 2.5: an advanced software platform for Bayesian evolutionary analysis. PLoS Comput. Biol. 15:1–28.

Bouckaert R.R. 2010. DensiTree: making sense of sets of phylogenetic trees. Bioinformatics 26:1372–1373.

Brook G. 1892. Preliminary descriptions of new species of *Madrepora* in the collections of the British Museum. Part II. Ann. Mag. Nat. Hist. 10:451–465.

Bryant D., Bouckaert R., Felsenstein J., Rosenberg N.A., Roychoudhury A. 2012. Inferring species trees directly from biallelic genetic markers: bypassing gene trees in a full coalescent analysis. Mol. Biol. Evol. 29:1917–1932.

Budd A.F., Romano S.L., Smith N.D., Barbeitos M.S. 2010. Rethinking the phylogeny of scleractinian corals: a review of morphological and molecular data. Integr. Comp. Biol. 50:411–427.

Carpenter K.E., Abrar M., Aeby G., Aronson R.B., Banks S., Bruckner A., Chiriboga A., Cortés J., Delbeek J.C., DeVantier L., Edgar G.J., Edwards A.J., Fenner D., Guzmán H.M., Hoeksema B.W., Hodgson G., Johan O., Licuanan W.Y., Livingstone S.R., Lovell E.R., Moore J.A., Obura D.O., Ochavillo D., Polidoro B.A., Precht W.F., Quibilan M.C., Reboton C., Richards Z.T., Rogers A.D., Sanciangco J., Sheppard A., Sheppard C., Smith J., Stuart S., Turak E., Veron J.E.N., Wallace C., Weil E., Wood E. 2008. One-third of reef-building corals face elevated extinction risk from climate change and local impacts. Science 321:560–563.

Castresana J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17:540–552.

Chhatre V.E., Emerson K.J. 2017. StrAuto: Automation and parallelization of STRUCTURE analysis. BMC Bioinformatics 18:1–5.

Cowman P.F., Quattrini A.M., Bridge T.C., Watkins-Colwell G.J., Fadli N., Grinblat M., Roberts T.E., McFadden C.S., Miller D.J., Baird A.H. 2020. An enhanced target-enrichment bait set for Hexacorallia provides phylogenomic resolution of the staghorn corals (Acroporidae) and close relatives. Mol. Phylogenetics Evol. 153:106944.

Dana J.D. 1846. Zoophytes. United States Exploring Expedition during the years 1838-1842. Philadelphia: Lea and Blanchard. p. 740.

Debortoli N., Li X., Eyres I., Fontaneto D., Hespeels B., Tang C.Q., Flot J.-F.F., Van Doninck K. 2016. Genetic exchange among bdelloid rotifers

is more likely due to horizontal gene transfer than to meiotic sex. Curr. Biol. 26:723–732.

Dellicour S., Flot J.F. 2018. The hitchhiker's guide to single-locus species delimitation. Mol. Ecol. Resour. 18:1234–1246.

Doyle J.J. 1995. The irrelevance of allele tree topologies for species delimitation, and a non-topological alternative. Syst. Bot. 20:574–588.

Erickson K.L., Pentico A., Quattrini A.M., McFadden C.S. 2021. New approaches to species delimitation and population structure of anthozoans: two case studies of octocorals using ultraconserved elements and exons. Mol. Ecol. Resour. 21:78–92.

Evanno G., Regnaut S., Goudet J. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. Mol. Ecol. 14:2611–2620.

Faircloth B.C. 2016. PHYLUCE is a software package for the analysis of conserved genomic loci. Bioinformatics 32:786–788.

Faircloth B.C., McCormack J.E., Crawford N.G., Harvey M.G., Brumfield R.T., Glenn T.C. 2012. Ultraconserved elements anchor thousands of genetic markers spanning multiple evolutionary timescales. Syst. Biol. 61:717–726.

Flot J.-F. 2010. Seqphase: a web tool for interconverting phase input/output files and fasta sequence alignments. Mol. Ecol. Resour. 10:162–166.

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 Evol. Biol. 10:372.

Flot J.-F. 2007. Champuru 1.0: a computer software for unraveling mixtures of two DNA sequences of unequal lengths. Mol. Ecol. Notes. 7:974–977.

Flot J.-F., Tillier A., Samadi S., Tillier S. 2006. Phase determination from direct sequencing of length-variable DNA regions. Mol. Ecol. Notes 6:627–630.

Flouri T., Rannala B., Yang Z. 2020. A tutorial on the use of BPP for species tree estimation and species delimitation. In: Scornavacca C., Delsuc F., Galtier N., editors. Phylogenetics in the genomic era. HAL archives-ouvertes. p. 5.6:1–5.6:16.

Fukami H., Budd A.F., Levitan D.R., Jara J., Kersanach R., Knowlton N. 2004a. Geographic differences in species boundaries among members of the *Montastraea annularis* complex based on molecular and morphological markers. Evolution 58:324–337.

Fukami H., Budd A.F., Paulay G., Solé-Cava A., Chen C.A., Iwao K., Knowlton N. 2004b. Conventional taxonomy obscures deep divergence between Pacific and Atlantic corals. Nature 427:832–835.

Fukami H., Omori M., Hatta M. 2000. Phylogenetic relationships in the coral family Acroporidae, reassessed by inference from mitochondrial genes. Zool. Sci. 17:689–696.

Funk D.J., Omland K.E. 2003. Species-level paraphyly and polyphyly: frequency, causes, and consequences, with insights from animal mitochondrial DNA. Annu. Rev. Ecol. Evol. Syst. 34:397–423.

Gouy M., Guindon S., Gascuel O. 2010. SeaView version 4: a multiplatform graphical user interface for sequence alignment and phylogenetic tree building. Mol. Biol. Evol. 27:221–224.

Grummer J.A., Bryson R.W., Reeder T.W. 2014. Species delimitation using Bayes factors: Simulations and application to the *Sceloporus scalaris* species group (Squamata: Phrynosomatidae). Syst. Biol. 63:119–133.

Harrison P.L., Babcock R.C., Bull G.D., Oliver J.K., Wallace C.C., Willis B.L. 1984. Mass spawning in tropical reef corals. Science 223:1186–1189.

Herrera S., Shank T.M. 2016. RAD sequencing enables unprecedented phylogenetic resolution and objective species delimitation in recalcitrant divergent taxa. Mol. Phylogenet. Evol. 100:70–79.

Hughes T.P., Anderson K.D., Connolly S.R., Heron S.F., Kerry J.T., Lough J.M., Baird A.H., Baum J.K., Berumen M.L., Bridge T.C., Claar D.C., Eakin C.M., Gilmour J.P., Graham N.A.J., Harrison H., Hobbs J.P.A., Hoey A.S., Hoogenboom M., Lowe R.J., McCulloch M.T., Pandolfi J.M., Pratchett M., Schoepf V., Torda G., Wilson S.K. 2018a. Spatial and temporal patterns of mass bleaching of corals in the Anthropocene. Science 359:80–83.

Hughes T.P., Barnes M.L., Bellwood D.R., Cinner J.E., Cumming G.S., Jackson J.B.C., Kleypas J., Van De Leemput I.A., Lough J.M.,

Morrison T.H., Palumbi S.R., Van Nes E.H., Scheffer M. 2017. Coral reefs in the Anthropocene. Nature 546:82–90.

Hughes T.P., Kerry J.T., Baird A.H., Connolly S.R., Dietzel A., Eakin C.M., Heron S.F., Hoey A.S., Hoogenboom M.O., Liu G., McWilliam M.J., Pears R.J., Pratchett M.S., Skirving W.J., Stella J.S., Torda G. 2018b. Global warming transforms coral reef assemblages. Nature 556:492–496.

Jakobsson M., Rosenberg N.A. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806.

Jombart T., Devillard S., Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genet. 11:94.

Junier T., Zdobnov E.M. 2010. The Newick utilities: high-throughput phylogenetic tree processing in the UNIX shell. Bioinformatics 26:1669–1670.

Kalyaanamoorthy S., Minh B.Q., Wong T.K.F., Von Haeseler A., Jermiin L.S. 2017. ModelFinder: fast model selection for accurate phylogenetic estimates. Nat. Methods 14:587–589.

Kass R.E., Raftery A.E. 1995. Bayes factors. J. Am. Stat. Assoc. 90:773–795.

Katoh K., Misawa K., Kuma K.I., Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res. 30:3059–3066.

Katoh K., Toh H. 2008. Recent developments in the MAFFT multiple sequence alignment program. Brief. Bioinform. 9:286–298.

Kitahara M. V., Fukami H., Benzoni F., Huang D. 2016. The new systematics of Scleractinia: integrating molecular and morphological evidence. In: Goffredo S., Dubinsky Z., editors. The Cnidaria, past, present and future. Cham: Springer. p. 41–59.

Kitchen S.A., Ratan A., Bedoya-Reina O.C., Burhans R., Fogarty N.D., Miller W., Baums I.B. 2019. Genomic variants among threatened *Acropora* corals. G3 Genes, Genomes, Genet. 9:1633–1646.

Knowlton N. 2001. Who are the players on coral reefs and does it matter? The importance of coral taxonomy for coral reef management. Bull. Mar. Sci. 69:305–308.

Knowlton N., Weil E., Weight L.A., Guzman H.M. 1992. Sibling species in *Montastraea annularis*, coral bleaching, and the coral climate record. Science 255:330–333.

Ladner J.T., Palumbi S.R. 2012. Extensive sympatry, cryptic diversity and introgression throughout the geographic distribution of two coral species complexes. Mol. Ecol. 21:2224–2238.

Langmead B., Salzberg S.L. 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods 9:357–359.

Leaché A.D., Bouckaert R.R. 2018. Species trees and species delimitation with SNAPP: a tutorial and worked example. Workshop on Population and Speciation Genomics; Český Krumlov, Czech Republic. p. 1–17.

Leaché A.D., Fujita M.K. 2010. Bayesian species delimitation in West African forest geckos (*Hemidactylus fasciatus*). Proc. R. Soc. B Biol. Sci. 277:3071–3077.

Leaché A.D., Fujita M.K., Minin V.N., Bouckaert R.R. 2014. Species delimitation using genome-wide SNP Data. Syst. Biol. 63:534–542.

Li H. 2013. Seqtk: a fast and lightweight tool for processing FASTA or FASTQ sequences. Available from: https://github.com/lh3/seqtk.

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079.

Liew Y.J., Aranda M., Voolstra C.R. 2016. Reefgenomics.Org – a repository for marine genomics data. Database 2016:baw152.

Márquez L.M., Van Oppen M.J.H., Willis B.L., Reyes A., Miller D.J. 2002. The highly cross-fertile coral species, *Acropora hyacinthus* and *Acropora cytherea*, constitute statistically distinguishable lineages. Mol. Ecol. 11:1339–1349.

Marshall P.A., Baird A.H. 2000. Bleaching of corals on the Great Barrier Reef: differential susceptibilities among taxa. Coral Reefs 19:155–163.

McFadden C.S., Haverkort-Yeh R., Reynolds A.M., Halàsz A., Quattrini A.M., Forsman Z.H., Benayahu Y., Toonen R.J. 2017. Species

boundaries in the absence of morphological, ecological or geographical differentiation in the Red Sea octocoral genus *Ovabunda* (Alcyonacea: Xeniidae). Mol. Phylogenet. Evol. 112:174–184.

Miller D.J., van Oppen M.J.H. 2003. A 'fair go' for coral hybridization. Mol. Ecol. 12:805–807.

Miller M.A., Pfeiffer W., Schwartz T. 2010. Creating the CIPRES Science Gateway for inference of large phylogenetic trees. 2010 Gateway Computing Environments Workshop (GCE); Institute of Electrical and Electronics Engineers (IEE). p. 1–8. doi: 10.1109/GCE.2010.5676129.

Minh B.Q., Nguyen M.A.T., Von Haeseler A. 2013. Ultrafast approximation for phylogenetic bootstrap. Mol. Biol. Evol. 30:1188–1195.

Nemenzo F. 1971. Systematic studies on Philippine shallow-water scleractinians: VII. Additional forms. Nat. Appl. Sci. Bull. Univ. Philipp. 23:141–209.

Nguyen L.T., Schmidt H.A., Von Haeseler A., Minh B.Q. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Mol. Biol. Evol. 32:268–274.

Nishihira M., Veron J.E.N. 1995. Hermatypic corals of Japan. Tokyo: Kaiyusha.

Odorico D.M., Miller D.J. 1997. Variation in the ribosomal internal transcribed spacers and 5.8S rDNA among five species of Acropora (Cnidaria; Scleractinia): patterns of variation consistent with reticulate evolution. Mol. Biol. Evol. 14:465–473.

van Oppen M.J.H., McDonald B.J., Willis B., Miller D.J. 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? Mol. Biol. Evol. 18:1315–1329.

Pandolfi J.M., Bradbury R.H., Sala E., Hughes T.P., Bjorndal K.A., Cooke R.G., McArdle D., McClenachan L., Newman J.H. M., Paredes G., Warner R.R., Jackson J.B.C. 2003. Global trajectories of the long-term decline of coral reef ecosystem. Science 301:955–958.

Pritchard J.K., Stephens M., Donnelly P. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Puillandre N., Modica M. V., Zhang Y., Sirovich L., Boisselier M.C., Cruaud C., Holford M., Samadi S. 2012. Large-scale species delimitation method for hyperdiverse groups. Mol. Ecol. 21:2671–2691.

Quattrini A.M., Faircloth B.C., Dueñas L.F., Bridge T.C.L., Brugler M.R., Calixto-Botía I.F., DeLeo D.M., Forêt S., Herrera S., Lee S.M.Y., Miller D.J., Prada C., Rádis-Baptista G., Ramírez-Portilla C., Sánchez J.A., Rodríguez E., McFadden C.S. 2018. Universal target-enrichment baits for anthozoan (Cnidaria) phylogenomics: new approaches to long-standing problems. Mol. Ecol. Resour. 18:281–295.

Quattrini A.M., Wu T., Soong K., Jeng M.S., Benayahu Y., McFadden C.S. 2019. A next generation approach to species delimitation reveals the role of hybridization in a cryptic species complex of corals. BMC Evol. Biol. 19:1–19.

Quelch J.J. 1886. Report on the Reef-corals collected by H.M.S. Challenger during the years 1873–1876. Zoology 16:1–203.

Quinlan A.R., Hall I.M. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842.

R Core Team. 2018. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available from: https://www.R-project.org/.

Rabiee M., Mirarab S. 2020. SODA: multi-locus species delimitation using quartet frequencies. Bioinformatics 36:5623–5631.

Rabiee M., Sayyari E., Mirarab S. 2019. Multi-allele species reconstruction using ASTRAL. Mol. Phylogenet. Evol. 130:286–296.

Rambaut A. 2018. FigTree: a graphical viewer of phylogenetic trees. Available from: http://tree.bio.ed.ac.uk/software/figtree/.

Rambaut A., Drummond A.J., Xie D., Baele G., Suchard M.A. 2018. Posterior summarization in Bayesian phylogenetics using Tracer 1.7. Syst. Biol. 67:901–904.

ReFuGe 2020 Consortium. 2015. The ReFuGe 2020 Consortium – using "omics" approaches to explore the adaptability and resilience of coral holobionts to environmental change. Front. Mar. Sci. 2:68.

Richards Z.T., Berry O., van Oppen M.J.H. 2016. Cryptic genetic divergence within threatened species of *Acropora* coral from the Indian and Pacific Oceans. Conserv. Genet. 17:577–591.

RStudio Team. 2017. RStudio: integrated development for R. Boston (MA): RStudio, PBC. Available from: http://www.rstudio.com/.

Sang T., Zhong Y. 2000. Testing hybridization hypotheses based on incongruent gene trees. Syst. Biol. 49:422–434.

Shinzato C., Shoguchi E., Kawashima T., Hamada M., Hisata K., Tanaka M., Fujie M., Fujiwara M., Koyanagi R., Ikuta T., Fujiyama A., Miller D.J., Satoh N. 2011. Using the *Acropora digitifera* genome to understand coral responses to environmental change. Nature 476:320–323.

Spöri Y., Flot J.-F. 2020. HaplowebMaker and CoMa: two web tools to delimit species using haplowebs and conspecificity matrices. Methods Ecol. Evol. 11:1434–1438.

Stephens M., Donnelly P. 2003. A comparison of bayesian methods for haplotype reconstruction from population genotype data. Am. J. Hum. Genet. 73:1162–1169.

Stephens M., Smith N.J., Donnelly P. 2001. A new statistical method for haplotype reconstruction from population data. Am. J. Hum. Genet. 68:978–989.

Suzuki G., Keshavmurthy S., Hayashibara T., Wallace C.C., Shirayama Y., Chen C.A., Fukami H. 2016. Genetic evidence of peripheral isolation and low diversity in marginal populations of the *Acropora hyacinthus* complex. Coral Reefs 35:1419–1432.

Talavera G., Castresana J. 2007. Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. Syst. Biol. 56:564–577.

Untergasser A., Cutcutache I., Koressaar T., Ye J., Faircloth B.C., Remm M., Rozen S.G. 2012. Primer3 - new capabilities and interfaces. Nucleic Acids Res. 40:1–12.

Veron J.E.N. 1995. Corals in space and time: the biogeography and evolution of the Scleractinia. Ithaca, NY: Cornell University Press.

Veron J.E.N. 2000. Corals of the world. Townsville: Australian Institute of Marine Science (AIMS).

Vollmer S. V, Palumbi S.R. 2002. Hybridization and the evolution of reef coral diversity. Science 296:2023–2025.

Wallace C. 1999. Staghorn corals of the genus *Acropora*. Collibgwood, VIC: CSIRO Publishing.

Wallace C.C. 1985. Reproduction, recruitment and fragmentation in nine sympatric species of the coral genus *Acropora*. Mar. Biol. 88:217–233.

Wallace C.C., Done B.J., Muir P.R. 2012. Revision and catalogue of worldwide staghorn corals *Acropora* and *Isopora* (Scleractinia: Acroporidae) in the Museum of Tropical Queensland. Mem. Queensl. Museum. 52:1–22.

Wallace C.C., Willis B.L. 1994. Systematics of the coral genus *Acropora*: implications of new biological findings for species concepts. Annu. Rev. Ecol. Syst. 25:237–262.

Willis B.L., Babcock R.C., Harrison P.L., Wallace C.C. 1997. Experimental hybridization and breeding incompatibilities within the mating systems of mass spawning reef corals. Coral Reefs 16:S53–S65.

Willis B.L., van Oppen M.J.H., Miller D.J., Vollmer S. V, Ayre D.J. 2006. The role of hybridization in the evolution of reef corals. Annu. Rev. Ecol. Evol. Syst. 37:489–517.

Wolstenholme J.K., Wallace C.C., Chen C.A. 2003. Species boundaries within the *Acropora humilis* species group (Cnidaria; Scleractinia): a morphological and molecular interpretation of evolution. Coral Reefs. 22:155–166.

Yang Z. 2015. The BPP program for species tree estimation and species delimitation. Curr. Zool. 61:854–865.

Ying H., Hayward D.C., Cooke I., Wang W., Moya A., Siemering K.R., Sprungala S., Ball E.E., Forêt S., Miller D.J. 2019. The whole-genome sequence of the coral *Acropora millepora*. Genome Biol. Evol. 11:1374–1379.

Zhang C., Rabiee M., Sayyari E., Mirarab S. 2018. ASTRAL-III: Polynomial time species tree reconstruction from partially resolved gene trees. BMC Bioinformatics 19:15–30.