From Genomics to Integrative Taxonomy? The Case Study of *Pocillopora* Corals

NICOLAS OURY[1,2,*], CYRIL NOËL[3], STEFANO MONA[4,5,6], DIDIER AURELLE[4,7], AND HELENE MAGALON[1,2,6]

[1] *UMR ENTROPIE (UMR 9220 – Université de La Réunion, IRD, IFREMER, Université de Nouvelle-Calédonie, CNRS), Université de La Réunion, St Denis, La Réunion, France*

[2] *Laboratoire Cogitamus, Paris, France*

[3] *IFREMER – IRSI – Service de Bioinformatique (SeBiMER), Plouzané, France*

[4] *ISYEB – Institut de SYstématique, Évolution, Biodiversité (UMR 7205 – CNRS, MNHN, Sorbonne Université, EPHE, Université des Antilles), École Pratique des Hautes Études, Paris, France*

[5] *EPHE, PSL Research University, Paris, France*

[6] *Laboratoire d'Excellence CORAIL, Perpignan, France*

[7] *Aix Marseille Université, Université de Toulon, CNRS, IRD, MIO, Marseille, France*

** Correspondence to be sent to: Nicolas Oury, UMR ENTROPIE, Université de La Réunion, Faculté des Sciences et Technologies, 15 bd René Cassin, CS 92003, 97744 St Denis Cedex 09, La Réunion, France; E-mail: nicolasoury@hotmail.fr.*

ORCID

NO https://orcid.org/0000-0002-5386-4633

CN https://orcid.org/0000-0002-7139-4073

SM https://orcid.org/0000-0001-9087-0656

DA https://orcid.org/0000-0002-3922-7291

HM https://orcid.org/0000-0002-7061-955X

Abstract. — With the advent of genomics, sequencing thousands of loci from hundreds of individuals now appears feasible at reasonable costs, allowing complex phylogenies to be resolved. This is particularly relevant for cnidarians, for which insufficient data due to the small number of currently available markers, coupled with difficulties in inferring gene trees and morphological incongruences, encrypts species boundaries, thereby blurring the study and conservation of these organisms. Yet, can genomics alone be used to delimit species in an integrative taxonomic context? Here, focusing on the coral genus *Pocillopora*, which plays key roles in Indo-Pacific reef ecosystems but has challenged taxonomists for decades, we explored and discussed the usefulness of multiple criteria (genetics, morphology, biogeography and symbiosis ecology) to delimit species of this genus. Phylogenetic inferences, clustering approaches and species delimitation methods based on genome-wide single-nucleotide polymorphisms (SNPs) were first used to resolve *Pocillopora* phylogeny and propose genomic species hypotheses from 356 colonies sampled across the Indo-Pacific (western Indian Ocean, tropical southwestern Pacific and south-east Polynesia). These species hypotheses were then compared to previous genetic evidences, as well as to evidences based on morphology, biogeography and symbiosis. Genomics allowed to delimit 21 species hypotheses where only seven are currently recognised based on current taxonomy. Moreover, 13 species were strongly supported by all approaches, either confirming their currently recognised species status, or supporting the presence of new species that need to be formally described. Some of the other genomic species hypotheses were supported by biogeographic or symbiosis evidences, but additional investigations are needed to state on their species status. Altogether, our results support (1) the obsolescence of macromorphology (i.e., overall colony and branches shape) but the relevance of micromorphology (i.e., corallite structures) to refine *Pocillopora* species limits, (2) the need to identify molecularly species prior to their study, as morphology can blur species identification on the field, (3) the relevance of the mtORF (coupled with other markers in some cases) as a diagnostic marker of most species, and (4) the need for a taxonomical revision in the *Pocillopora* genus. These results give new insights into the usefulness of multiple criteria for resolving *Pocillopora* species limits and will ultimately provide helpful insights for the conservation of the species from this scleractinian genus. [biogeography; cryptic species delimitation; Indo-Pacific; microsatellites; morphology; phylogenetics; single-nucleotide polymorphism (SNP); Symbiodiniaceae]

Efficiently protecting species implies knowing their life history traits and functioning. This requires accurately defining species limits, something that may sound trivial but has long been debated (e.g., Mayden 1997; De Queiroz 2007). Indeed, several species concepts, based on more or less compatible criteria, have previously been proposed (reviewed in De Queiroz 2007). Each concept has its own advantages, but also its own approximations of the biological truth, so it now appears evident to integrate multiple criteria and go towards a unified species concept (De Queiroz 2005). However, it is not always obvious how these criteria should be combined, and some may be more informative than others or give contradictory insights, depending on organisms.

Accurately delimiting species is particularly essential for scleractinian corals, the cornerstone of coral reefs, which are experiencing critical decline worldwide (Hughes et al. 2017, 2018, 2019; Heron et al. 2018), attributable both to local (e.g., coastal development, over-fishing, pollution) and global (e.g., climate change) pressures. Coral taxonomy initially relied on skeleton morphological traits (i.e., *corallum* macromorphology and corallite microstructure; Vaughan and Wells 1943; Wells 1956; Chevalier 1971; Veron 2000), but phenotypic plasticity hampers reliable species delimitation on this basis (see Todd 2008). With the advent of genetics, molecular approaches have been used to explore species boundaries, revealing incongruences of conventional systematics within many scleractinian genera (e.g., Keshavmurthy et al. 2013; Schmidt-Roach et al. 2014; Gélin et al. 2017b; Cunha et al. 2019; Arrigoni et al. 2020). Nuclear internal transcribed spacers (ITS) and mitochondrial markers have been extensively used in phylogenetic inferences (e.g., Benzoni et al. 2007; Gélin et al. 2017b; Nakajima et al. 2017). However, intra-individual and intra-specific variations for the formers (van Oppen et al. 2000; Chen et al. 2004; Vollmer and Palumbi 2004), and relatively slow evolutionary rates for the latter (van Oppen et al. 1999; Shearer et al. 2002; Hellberg 2006), make these markers usually not informative for species delimitation in most genera (e.g., Forsman et al. 2009; Terraneo et al. 2016). Additionally, the small number of currently available markers, coupled with hybridisation (Willis et al. 2006; Combosch et al. 2008; Richards et al. 2008), introgression (Combosch and Vollmer 2015; Hellberg et al. 2016) and incomplete lineage

sorting in gene trees (van Oppen et al. 2001; Fukami et al. 2008) blur phylogenetic relationships between taxa.

The recent development of high-throughput sequencing technologies now enables the cost-effective target of large numbers of loci from hundreds of individuals from virtually any species (Metzker 2010). These methods appear particularly promising to resolve complex phylogenies such as those involving scleractinian corals (e.g., Forsman et al. 2017; Cunha et al. 2019; Arrigoni et al. 2020). In particular, restriction-site associated DNA sequencing (RADseq; Baird et al. 2008) and sequence capture (also called target enrichment; Hodges et al. 2007; Gnirke et al. 2009) are increasingly used, from population genetics to phylogenetic studies (see Narum et al. 2013 for a review). While RADseq typically generates datasets of anonymous loci, sequence capture enables the deep sequencing of previously identified loci of interest, but needs existing genomic resources to design probes (Davey et al. 2011; Harvey et al. 2016). When such genomic resources are unavailable for the species of interest, probes from genomic regions that are conserved across divergent taxa [e.g., ultraconserved elements (UCEs); https://www.ultraconserved.org/] can be used (Faircloth et al. 2012, 2013; McCormack et al. 2012).

The coral genus *Pocillopora* Lamarck, 1816 (Scleractinia, Pocilloporidae) represents a key component of coral reef ecosystems from the Indo-Pacific and the Red Sea (Veron 2000), as its branching colonies are abundant and sometimes the main bio-constructors (e.g., Benzoni et al. 2003). However, its taxonomy remains challenging, and the extraordinary range of morphological diversity among its colonies has led to the coining of more than 40 species names (Hoeksema and Cairns 2022). Defining morphospecies based on morphological characters (shape and organisation of branches and verrucae), Veron (2000) recognised only 17 of them. Recent genetic studies identified several cryptic species and lineages within those morphospecies (see Gélin et al. 2017b for a review). As an illustration, the so-called *P. damicornis* (Linnaeus, 1758) was disentangled in five genetic lineages: *P. damicornis* types α, β, δ, γ, and ε (Schmidt-Roach et al. 2012), *a posteriori* defined as five distinct species and named *P. damicornis* (Linnaeus, 1758), *P. acuta* Lamarck, 1816,

P. aliciae Schmidt-Roach, Miller & Andreakis 2013, *P. verrucosa* (Ellis & Solander, 1786), and *P. brevicornis* Lamarck, 1816, respectively (Schmidt-Roach et al. 2014). Following this taxonomical revision of the genus, 21 valid *Pocillopora* species are currently accepted (Hoeksema and Cairns 2022). Besides, using species delimitation methods based on sequence data from colonies sampled in three marine provinces (western Indian Ocean, tropical southwestern Pacific and south-east Polynesia), Gélin et al. (2017b) defined within the *Pocillopora* genus 16 primary species hypotheses (PSHs *sensu* Pante et al. 2015). Some of these PSHs correspond to currently accepted species, but others do not and would therefore represent undescribed species. Additionally, using microsatellites, some PSHs were partitioned into several secondary species hypotheses (SSHs *sensu* Pante et al. 2015), themselves partitioned into several divergent but sympatric genetic clusters (Gélin et al. 2017a, 2017b, 2018a, 2018b; Oury et al. 2020a, 2021, 2022). This genetic partitioning questions species limits and shelves taxonomic uncertainties for which traditional genetic markers appear not enough resolutive. So far, only two studies (Johnston et al. 2017, 2022) have inferred phylogenetic relationships among species of the *Pocillopora* genus using high-throughput sequencing data (ezRAD; Toonen et al. 2013). In both cases, they resolved clear monophyletic groups that coincide with previously published mitochondrial clades based on the so-called open reading frame marker (mtORF; a putative protein-coding region of unknown function; Flot and Tillier 2007). However, their samplings were relatively concise (13 and 55 samples from seven morphospecies) and restricted to the Pacific, missing a huge part of the high diversity of this genus.

Here, considering a subset of 356 *Pocillopora* colonies from the same sampling set as in Gélin et al. (2017b), representing the totality of the PSHs, SSHs and clusters previously identified (see Gélin et al. 2017a, 2017b, 2018a, 2018b; Oury et al. 2020a, 2021, 2022), as well as all morphotypes sampled, we used sequence capture of UCEs and exon loci to collect single-nucleotide polymorphisms (SNPs). Maximum-likelihood and Bayesian phylogenetic inferences, clustering approaches and species delimitation methods based on SNP data were applied to resolve the *Pocillopora* phylogeny and define genomic species hypotheses, which were compared to previous

genetic partitionings of the genus (i.e., the PSHs, SSHs and clusters previously defined based on the mtORF marker and microsatellites). Genetic evidences were then confronted to other criteria (macro- and micromorphology, biogeography and associated Symbiodiniaceae communities), to propose species delimitation of *Pocillopora* in an integrative taxonomic context. The usefulness of each criterion and its integration were then discussed.

MATERIALS AND METHODS

Detailed materials and methods, including sampling, sequencing and analytical methods, are available in Appendices 1-4.

Sampling

The sampling was the same as in Gélin et al. (2017b) and comprised ca. 9,000 *Pocillopora* colonies from various habitats and morphotypes, from three marine provinces: the western Indian Ocean (WIO), the tropical southwestern Pacific (TSP) and the south-east Polynesia (SEP). All colonies were previously genotyped with 13 microsatellites and for a subset, we also sequenced the mitochondrial ORF locus (mtORF; see Gélin et al. 2017b for more details). Each colony was thus assigned beforehand a primary and a secondary species hypothesis (PSH and SSH, respectively; *sensu* Gélin et al. 2017b), and a cluster when appropriate, based on these genetic data (see, for example, Oury et al. 2021). From now, to simplify the reading, PSHs that were not subdivided into several SSHs are designated SSHs, keeping their corresponding number (e.g., PSH01 switches to SSH01). These SSHs remain easily recognisable as no lowercase letter follows the number.

In this study, a subset of 356 *Pocillopora* colonies (Table S1 & Fig. S1 in Appendix 1), covering the totality of the localities and morphotypes sampled, as well as all SSHs and clusters, was considered to maximise the genetic diversity explored. Four *Seriatopora hystrix* and four *Stylophora pistillata* colonies were also included as outgroups.

Molecular Analyses

Sequencing and bioinformatics processing. — All 364 colonies, plus eight sequencing replicates, were sequenced following a target enrichment protocol of 1,248 ultraconserved elements (UCEs) and 1,385 exon loci (Quattrini et al. 2018; see Appendix 2 for more details). The bioinformatics pipeline, from demultiplexed reads to final SNP datasets, is detailed in Appendix 2. Three individuals were discarded due to too many missing data (> 60%).

Phylogenomic analyses. — All following analyses (detailed in Appendix 2) were performed on two datasets, one keeping all filtered SNPs and the other keeping one randomly chosen SNP per locus to reduce the effect of linkage disequilibrium. Available *Pocillopora* genomes [i.e., *P. acuta* (Vidal-Dupiol et al. 2019), *P. damicornis* (Cunning et al. 2018) and *P. verrucosa* (Buitrago-López et al. 2020)] were also included by retrieving the genotypes of the SNPs corresponding to each dataset. Phylogenetic relationships were investigated using maximum likelihood (ML) and Bayesian inferences with RAxML-NG v0.9.0 (Kozlov et al. 2019) and BEAST v2.6.3 (Bouckaert et al. 2019), respectively, both using the GTR+G model. To support the phylogenomic analyses and further explore the genetic partitioning of the datasets, several clustering approaches were used. First, assignment tests were performed with STRUCTURE v2.3.4 (Pritchard et al. 2000), sNMF (Frichot et al. 2014) and discriminant analyses of principal components (dAPC; Jombart et al. 2010). Signals of admixture were further investigated with NEWHYBRIDS v1.1 (Anderson and Thompson 2002). Second, Nei (1972) individual genetic distances were computed with the R v4.0.4 (R Core Team 2021) library *'StAMPP'* (Pembleton et al. 2013), and then used to build a minimum spanning tree (MST) and an unrooted equal-angle split network using EDENETWORKS v2.18 (Kivelä et al. 2015) and SplitsTree v4.15.1 (Huson and Bryant 2006), respectively. Finally, SSH and cluster assignments issued from microsatellite data (Gélin et al. 2017a, 2017b, 2018a, 2018b; Oury et al. 2020a, 2021, 2022) were compared to the groups identified with all above analyses, named hereafter genomic species hypotheses (GSHs). F_{ST} (Weir and Cockerham 1984) were

computed with *'StAMPP'* (Pembleton et al. 2013) for each pair of GSHs, and the resulting matrix was clustered using the *heatmap.2* function from the R library *'gplots'* (Warnes et al. 2020).

As some GSHs did not include any individual whose mtORF had previously been sequenced, and in order to retrieve the correspondence with previous studies, we completed the set of mtORF-sequenced colonies following Gélin et al. (2017b), and further sequenced a subset of colonies for the PocHistone, a recently discovered marker partly mapped to partial histone 3 genes from other cnidarians, and allowing to identify *P. grandis* (the senior synonym of *P. eydouxi*) colonies (Johnston et al. 2018). The same laboratory protocol and analyses as for the mtORF in Gélin et al. (2017b) were used (Appendix 2).

Species delimitation analyses and divergence time estimation. — To confirm GSHs, Bayes factor delimitation with genomic data (BFD*; Leaché et al. 2014) was used to test several possible species delimitation models, using the SNAPP package (Bryant et al. 2012) implemented in BEAST v2.6.3 (Bouckaert et al. 2019; details in Appendix 2). To deal with computationally intensive demands from SNAPP, we first tested a batch of species delimitation scenarios to confirm the four main clades from the phylogeny (arbitrarily defined as monophyletic groups of individuals separated by nucleotide substitution per site distances of at least 0.4 on the ML tree). Then, several possible species delimitation models were tested within each clade separately, from one single species to the number of GSHs found for each clade (Table S5 in Appendix 2). Models were compared and ranked using their marginal likelihood estimate (MLE) and by calculating the Bayes factor (BF; Kass and Raftery 1995).

For each of the best-supported model (except for Clade 1, constituted of a single GSH), a coalescent-based tree was calculated with SNAPP. DensiTree v2.2.7 (Bouckaert 2010) was used to visualise the posterior distribution of topologies as cladograms, hence allowing for a clear depiction of uncertainties in the topology. Finally, GSH divergence times were estimated with the BEAST

package SNAPPER v1.0.1 (Stoltz et al. 2021; Appendix 2). The divergence between *Pocillopora* and outgroups was constrained to the middle-end Paleogene (28.4-42.7 Ma; Simpson et al. 2011).

Macro- and Micromorphological Analyses

In order to compare previously described morphospecies with GSHs (defined above), each colony was attributed a morphotype (or several when morphology was unclear), determined only by its *corallum* macromorphology [branch shape and thickness, size and uniformity of verrucae, and overall growth form as described in Veron (2000) and Schmidt-Roach et al. (2014)].

A subset of 10 colonies per GSH were also randomly selected for micromorphological observations of the bleached skeletons (particularly of the corallite structures) using scanning electron microscopy (SEM). A collection of skeleton images was thus obtained for each specimen, and multiple measurements of seven quantitative variables (e.g., corallite and columella diameters; see Appendix 3 for details) were done with ImageJ2 (Rueden et al. 2017; https://imagej.nih.gov/ij/). A non-parametric permutational multivariate anova (PERMANOVA) was then performed using the R library *'RVAideMemoire'* (Hervé 2021) with the GSHs as factor. Each metric was analysed separately using a non-parametric permutational anova. Two additional categorical variables were also considered, and a factorial analysis of mixed data (FAMD) was performed for all nine variables using the R library *'FactoMineR'* (Lê et al. 2008). A reference specimen representative of each species enclosed in the latest *Pocillopora* taxonomic revision (Schmidt-Roach et al. 2014) was included by measuring the variables on the images incorporated.

Characterisation of Associated Symbiodiniaceae

Symbiodiniaceae communities were characterised for a subset of colonies (ca. 15 per GSH, when available; including three replicates) by high-throughput sequencing the ribosomal RNA internal transcribed spacer 2 (ITS2; see Appendix 4 for details). Reads were processed with the SAMBA v3.0.1 workflow (https://github.com/ifremer-bioinformatics/samba). Resulting operational

taxonomic units (OTUs) were taxonomically assigned by querying a custom reference database of Symbiodiniaceae ITS2 adapted from the one available in SymPortal (downloaded on 13/01/2022; Hume et al. 2019). Taxonomic affiliations of the OTUs were confirmed by reconstructing the phylogenetic relationships among them using MAFFT v7.713 (Katoh and Standley 2013) and FastTree v2.1.11 (GTR+CAT model; Price et al. 2009). OTUs and individuals with less than 10 and 500 sequences, respectively, were then removed to reduce possible sequencing errors. Alpha diversity metrics (Chao1 and Shannon) were computed at the OTU level with the R library *'vegan'* (Oksanen et al. 2020) and compared using non-parametric permutational ANOVA performed with the R library *'RVAideMemoire'* (Hervé 2021), with the GSHs or the localities as factor. Finally, a nonmetric multidimensional scaling (NMDS) using Bray and Curtis (1957) dissimilarity index was performed to assess community similarity.

Species hypotheses delimited with each criterion (genomics, genetics, macro- and micromorphology, and symbiosis ecology) were then compared in an integrative species delimitation context. Sampling sites were also integrated to identify sympatric or allopatric GSHs. We then discussed the usefulness of each criterion for the delimitation of *Pocillopora* species.

RESULTS

Molecular Analyses

Sequencing and bioinformatics processing. — A total of 1.6×10^9 reads (2.5×10^{11} bp) were produced with a highly variable number of reads per individual [varying from 9.1×10^3 to 8.2×10^6 reads; mean ± s.e. = $(4.4 \pm 0.1) \times 10^6$ reads], but only three individuals (*a posteriori* removed) had less than a million reads. Quality controls and adapter trims then led to the removal of 3.0% of the bases. From the resulting trimmed reads, between 41.0% and 86.2% reads per individual were successfully mapped on the reference sequences (mean ± s.e. = 78.3 ± 0.4%), with a mean coverage

depth (± s.e.) of 60.2× (± 0.1). Finally, SNPs calling and filtering (Table S4 in Appendix 2) led to two datasets: one including all SNPs (361 individuals × 17,465 SNPs; 5.8% missing data) and the other keeping randomly one SNP per locus (361 individuals × 1,559 SNPs; 6.0% missing data), with mean SNP coverage depths (± s.e.) of 85.8× (± 0.4) and 76.1× (± 1.3), respectively.

Phylogenomic analyses. — All results were very consistent between both datasets (i.e., with one or all SNPs per locus). Thus, only results with one SNP per locus are presented below, but results keeping all SNPs are provided in Appendix 2. The phylogenetic trees inferred both with RAxML and BEAST gave similar tree topologies and recovered four strongly supported clades (Clades 1-4; i.e., monophyletic groups of individuals chosen here as separated by at least 0.4 nucleotide substitution per site on the ML tree; Fig. 1), themselves split (except Clade 1) into a total of 21 genomic species hypotheses (GSHs). Each GSH (except three) was restricted to a single marine province (Table 1 & Fig. 1) and several GSHs were thus sympatric (12 in the WIO, 11 in the TSP and 2 in the SEP; Table S6 in Appendix 2), supporting evolutionary rather than geographic reproductive isolations. Moreover, most of the GSHs (see below for the exceptions) roughly corresponded to previously defined secondary species hypotheses on the basis of microsatellites (SSHs *sensu* Gélin et al. 2017b). Therefore, to avoid introducing a new nomenclature and to ease correspondence with earlier works, the GSHs were named according to the corresponding SSHs (e.g., the GSH corresponding to SSH01 was named GSH01; Table 1 & Fig. 1).

SSH06, SSH07, SSH08 and SSH16, previously defined from few individuals, were not retrieved here as the corresponding individuals were grouped with those from SSH09a, SSH09b-1 or SSH13c. Similarly, SSH09b-2 was grouped with SSH10 and was far apart from the rest of SSH09 *sensu lato*, suggesting that individuals from SSH09b-2 correspond to GSH10 (as observed with the mtORF). GSH09b, therefore, corresponds to only SSH09b-1. SSH12 and SSH15, previously grouped with SSH13a and SSH13c, respectively, using microsatellites (Gélin et al. 2017b), were retrieved, confirming the distinction between them. Conversely, the over-partitioning

previously found with microsatellites inside SSH04a (Oury et al. 2020a), SSH05d (Clusters 1 and 4 in Gélin et al. 2018b), SSH09a, SSH09c (Gélin et al. 2018a) and SSH13c (Oury et al. 2021) was not retrieved, while the one found into SSH05c (Clusters 2 and 3 in Gélin et al. 2018b; *a posteriori* named SSH05c-1 and SSH05c-2 in Oury et al. 2020b) was. Finally, SSH05a was split into three new groups (GSH05a-1, GSH05a-2 and GSH05a-3), and SSH09c split into two new groups ($GSH09c_{WIO}$ and $GSH09c_{TSP}$) restricted to the WIO or the TSP, respectively (Table 1 & Fig. 1).

The three assignment methods, although estimating different admixture rates and suggesting different optimal *K* according to their respective criterion, gave similar results, retrieving almost all 21 GSHs (Fig. 1 & S3-S4 in Appendix 2). In particular, sNMF and STRUCTURE highlighted introgression signals among several GSHs, compatible with allopatrism, that were further investigated with NEWHYBRIDS. This was notably the case within GSH05 *sensu lato*, but also with GSH12 as hybrids between GSH13a and GSH13c, GSH15 between GSH13c and GSH14, or $GSH09c_{WIO}$ between GSH09a and $GSH09c_{TSP}$ (Fig. 1 & S3-S4). No individual was assigned to a hybrid class (i.e., F1, F2 or backcrosses), except the $GSH09c_{WIO}$ ones that were assigned as F2 hybrids from GSH09a and $GSH09c_{TSP}$ (data not shown). The two networks clustering also retrieved the 21 GSHs (Fig. S5 in Appendix 2). Thus, published genomes were assigned to the same GSHs with all datasets and analyses (Fig. 1 & Appendix 2): two [*P. acuta* (Vidal-Dupiol et al. 2019) and *P. verrucosa* (Buitrago-López et al. 2020)] were assigned to GSH13a (currently considered as *P. verrucosa*) and the third [*P. damicornis* (Cunning et al. 2018)] to $GSH09c_{TSP}$ (*P. grandis*).

Finally, all pairwise F_{ST} were significantly positive ($P < 0.001$***) and the dendrogram topology obtained from the clustering of F_{ST} values was comparable to phylogenies (Fig. S6 in Appendix 2). Intra-clade F_{ST} ranged from 0.092*** to 0.689*** [mean (± s.e.) = 0.332 ± 0.011], while inter-clade ones ranged from 0.420*** to 0.795*** [mean (± s.e.) = 0.551 ± 0.004; Table S7 in Appendix 2].

For the mtORF, 59 additional colonies were sequenced, but no new haplotype was found. Each haplotype (except three) was restricted to a single GSH (Table 1 & Fig. 1), confirming

315 previous results from Gélin et al. (2017b). In particular, ORF27 was found in GSH09 *sensu lato*,
316 thus corresponding to *P. grandis* and/or *P. meandrina*. To distinguish both species, we sequenced
317 10 colonies of each of the five GSHs from Clade 4 for the PocHistone. Among the 43 successfully
318 sequenced colonies, no heterozygote was found and eight novel 588 bp-haplotypes were identified
319 (Hist01-08; GenBank accession numbers ON155826-ON155833; Table S9 in Appendix 2), to
320 which we added the two available in GenBank (MG587096 and MG587097, corresponding to
321 *P. grandis* and *P. meandrina*, respectively; Johnston et al. 2018; Table S9). All, but one haplotype
322 (Hist07), were restricted to a single GSH (Table 1 & Fig. 1). Hist07 and Hist08 had the *P. grandis*
323 diagnostic SNP, suggesting that GSH09c corresponds to *P. grandis* (Table S9). The reconstructed
324 PocHistone phylogeny consistently regrouped these two haplotypes with the *P. grandis* one from
325 Johnston et al. (2018), but all other haplotypes were grouped inconsistently with the defined GSHs
326 (Fig. S7 in Appendix 2).
327

328 **Fig. 1** *Pocillopora* phylogeny reconstructed with one SNP per locus (361 individuals × 1,559 SNPs).
329 (a) maximum likelihood (ML) phylogenetic tree. Branches are coloured according to marine provinces
330 [black: western Indian Ocean (WIO); light grey: tropical southwestern Pacific (TSP); dark grey: south-east
331 Polynesia (SEP)], and branch support, based on ML bootstrap analyses (first number) and Bayesian posterior
332 probabilities (second number), is indicated for branches supporting the genomic species hypotheses (GSHs;
333 delimited by dashed lines; full lines delimit the clades indicated alongside). Published genomes are indicated
334 by lowercase letters [a: *P. verrucosa* (Buitrago-López et al. 2020); b: *P. acuta* (Vidal-Dupiol et al. 2019); c:
335 *P. damicornis* (Cunning et al. 2018)].
336 (b) sNMF assignments at K = 15, K = 21 and K = 25, mitochondrial open reading frame (mtORF) and
337 PocHistone haplotypes repartition [number of colonies in parentheses; haplotypes in bold are found in
338 several GSHs; *: ORF30 (3) & ORF31(4)], corresponding secondary species hypotheses (SSHs) and clusters
339 (defined in Gélin et al. 2017a, 2017b, 2018a, 2018b; Oury et al. 2020a, 2021, 2022), genomic species
340 hypotheses (GSHs; number of colonies in parentheses), clades and colour code retained throughout this
341 study.

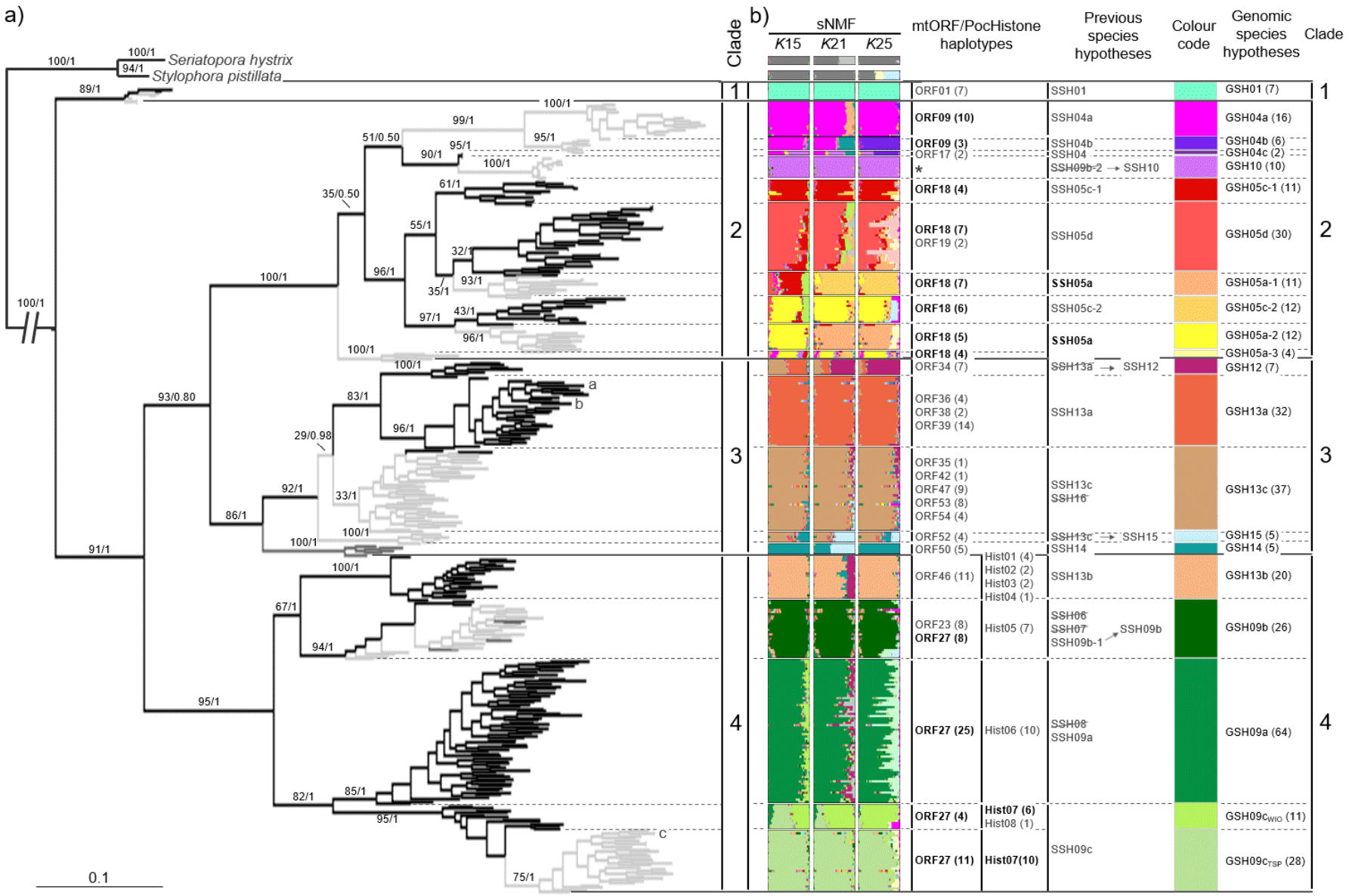


Table 1 Summary of the different approaches exploring *Pocillopora* species limits: genetics [genomic and corresponding corrected secondary species hypotheses (GSHs and SSHs *sensu* Gélin et al. 2017b, respectively), mtORF (mitochondrial open reading frame) and PocHist. (PocHistone) haplotypes; values in bold are retrieved in several GSHs], micro- and macromorphological, symbiosis (S.; each colour denotes distinct dominant Symbiodiniaceae; see Fig. S14 in Appendix 4) and geographical (WIO: western Indian Ocean; TSP: tropical southwestern Pacific; SEP: south-east Polynesia; values in bold highlight sympatric GSHs within a species complex) evidences. Corresponding lineages from previous studies are also indicated (arabic numerals correspond to types from Pinzón et al. 2013, roman numerals to clades from Marti-Puig et al. 2014 and greek letters to types from Schmidt-Roach et al. 2014; lineages in parentheses were extrapolated from SSHs).*: *P. villosa nomen nudum* was proposed by Gélin et al. (2017b) but does not correspond to a currently valid species.

Clade/GSH		SSH	mtORF	Poc Hist.	Micromorphology		Macromorphology	S.	Biogeo-graphy	Current taxonomy	Proposed taxonomy	Correspon-ding lineages
					Columella	Septa						
1	01	01	01		Styliform	Robust, with lobes	Robust and encrusting		WIO+TSP	*P. effusa*	***P. effusa***	2, IIIa
2	04a	04a	09		Flat and spinulate	Absent to rudimentary, indicated by small septal teeth	Slender branches, round to flattened with more or less pointed ends		**TSP**	*P. damicornis*	***P. damicornis*** **species complex?**	4a, Ib (4b, 4c, α)
	04b	04b							**TSP**			
	04c	n/a	17						WIO	n/a	**n/a**	n/a
	10	10	30,31				Short verrucose branches		TSP	*P. brevicornis*	***P. brevicornis***	(ε)
	05c-1	05c-1	**18**				Slender and bushy branches, ramified towards terminal ends		**WIO**	*P. acuta*	***P. acuta*** **species complex?**	5, Ia, β
	05d	05d	**18**,19						**WIO**			
	05a-1	**05a**	**18**						**TSP**			
	05c-2	05c-2							**WIO**			
	05a-2	**05a**							**TSP**			
	05a-3								**TSP**			
3	12	12	34		Weakly developped to flat, covered with spinulae	Developped (+ long teeth)	Robust, short and verrucose branches, with a cauliflower aspect		WIO	*P. verrucosa*	**12**	7
	13a	13a	36,38,39			Often developped, with small (≈ 50 µm) teeth			WIO		**13a**	3c, 3e (3g, 3j)
	13c	13c	35,42,47, 53,54			Well developped, with long (≈ 130 µm) teeth			WIO+TSP		**13c** ***P. verrucosa*?**	3b, 3d, 3f, 3h, IIa, γ (3i, χ)
	15	15	52		Oval-convex	Long and thin teeth			TSP		**15**	n/a
	14	14	50		Variable	Thin teeth (80-120 µm)			SEP		**14**	n/a
4	13b	13b	46	01-04	Flat and spinulate	Variable	Robust branches, with a velvety aspect		WIO	*P. villosa**	***P. villosa****	3a, γ
	09b	09b	23,**27**	05	Oval-convex, spinulate	Long (80-100 µm) and thin septal teeth	Robust long branches, often meandering		TSP+SEP	*P. meandrina*	**09b** ***P. meandrina*?**	1a, 8a, IIb, e/m (1b)
	09a	09a	**27**	06					WIO		**09a**	
	$09c_{WIO}$	09c		**07**-08	Styliform, with 1-3 stylae	Long (80-100 µm) and thin teeth, two cycles	Robust, long and cylindrical branches		WIO	*P. grandis*	***P. grandis*** **species complex?**	
	$09c_{TSP}$			**07**					TSP			

Species delimitation analyses. — Among the scenarios tested to delimit the four main clades (Clades 1-4; Table S5 in Appendix 2), the best-supported model was the one separating those four clades (model 4: MLE = -11,868.07, BF = –). The three other models were ranked with a decreasing number of clades (i.e., from three clades to a single one; 1,294.87 < BF < 4,706.03; Table S5). Within each clade, the model with the lowest MLE was the one separating colonies according to the different GSHs previously identified based on phylogenomic and clustering analyses. The best-supported model for Clade 1 was therefore the 1-species model (GSH01; model 1.1: MLE = -1,112.45, BF = –), followed by the 1-species-per-ocean model (model 1.2: MLE = -1,381.70, BF = 538.49). For Clade 2, analyses supported the 10-species model (model 2.17: MLE = -20,955.08, BF = –), followed by the models lumping GSH05a-1 and GSH05d (model 2.16: MLE = -21098.78, BF = 287.39) or GSH05a-2 and GSH05c-2 (model 2.14: MLE = -21,408.68, BF = 907.18). Finally, for Clades 3 and 4, the best supported models were the 5-species ones (model 3.8: MLE = -10,529.96, BF = –; model 4.11: MLE = -12,717.20, BF = –). However, Clade 4 5-species model was closely followed by the model lumping GSH09c$_{WIO}$ and GSH09c$_{TSP}$ (model 4.10: MLE = -12,763.15, BF = 91.89; Table S5). In summary, BFD* supported the 21 GSHs identified with the phylogenomic analyses.

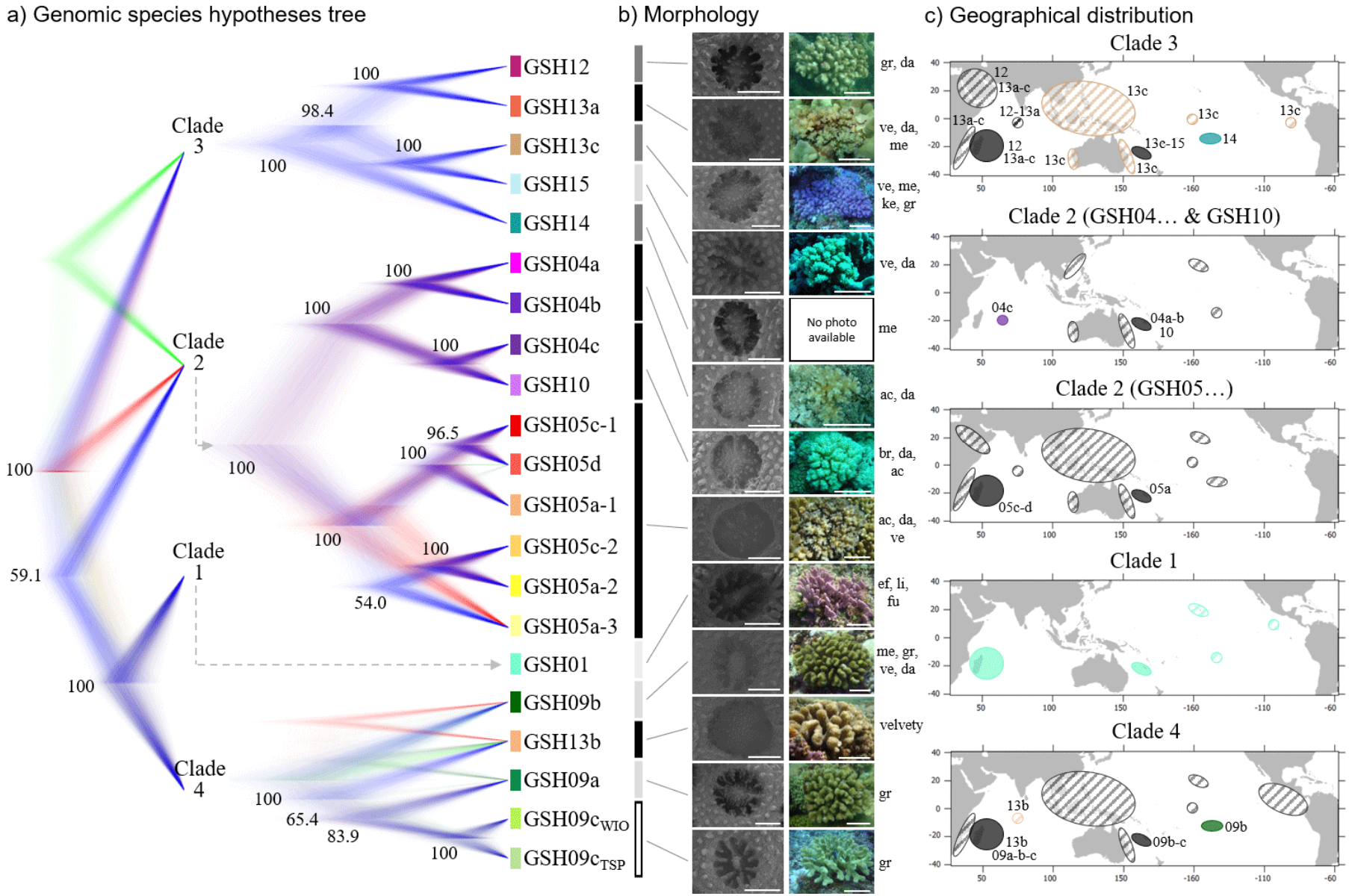
A total of four species trees were estimated (i.e., one for the best-supported model in the initial batch of scenarios, and then one for each best-supported model for scenarios within Clades 2-4 separately; Fig. 2a). For the initial batch of scenarios, three (out of three) consensus tree topologies were identified in the 95% HPD set, and all grouped Clades 1 and 4 together, whereas Clade 1 was the most distant group according to all previous analyses. The three topologies differed in whether Clades 2 or 3 shared a direct common ancestor with Clades 1 and 4 (59.1% and 22.0% of the trees, respectively), or together (18.9%; Fig. S8 in Appendix 2). For Clade 2, two (out of nine) consensus tree topologies were identified in the 95% HPD set. Both topologies were very similar and consistent with previous analyses, except that GSH05a-3 was alternatively grouped with or without GSH05a-2 and GSH05c-2 (52.1% and 44.3% of the trees, respectively; Fig. 2 & S8).

Only one (out of four) consensus tree topology was identified in the 95% HPD set for Clade 3 (representing 98.4% of the trees). This topology was consistent with previous analyses, i.e. grouping GSH12 and GSH13a on one side, GSH13c and GSH15 on the other side, and GSH14 being the most distant species (Fig. 2). Finally, for Clade 4, a total of 15 consensus tree topologies were found, of which five were in the 95% HPD set. All topologies identified GSH09c$_{WIO}$ and GSH09c$_{TSP}$ as sister species, but then differed in whether GSH09a shared a direct common ancestor with them (83.9% of the trees) and whether GSH09b and GSH13b were sister species (17.8%) or progressive outgroups (65.4%; Fig. 2 & S8).

The time-calibrated phylogeny indicated a first divergence within the *Pocillopora* genus 20.4 Ma, separating on one side Clades 1 and 4, and on the other side Clades 2 and 3. Clade pairs then diverged 17.4 Ma and 16.0 Ma, respectively (Fig. 3). Each clade then went through a first diversification period in the late Miocene (6.5-7.5 Ma), followed by a second period in the Pliocene and the Quaternary (i.e., from 4.5 Ma). Thus, almost all *Pocillopora* GSHs appeared relatively recently (Fig. 3).

Fig. 2 Species tree estimation for the 21 delimited *Pocillopora* genomic species hypotheses (GSHs). (a) complete set of consensus trees visualised with DensiTree for each best-supported model (i.e., for the initial batch of scenarios, and then for scenarios within Clades 2-4 separately). Higher density areas indicate greater topology agreement and different colours represent different topologies (trees with the highest clade credibility in blue). Node supports (Bayesian posterior probabilities) > 50% are indicated. (b) micro- (scale ≈ 500 µm) and macromorphological (scale ≈ 10 cm) overview of the GSHs (characteristic features only; see Appendix 5 for more illustrations). Bar colours symbolise separate micromorphological groups on the factorial analysis of mixed data (FAMD; see also Fig. S10 in Appendix 3) and morphotypes encountered in this study (sorted by occurrence) are indicated alongside photographs (ac: acuta, br: brevicornis, da: damicornis, ef: effusa, fu: fungiformis, gr: grandis, ke: kelleheli, li: ligulata, me: meandrina and ve: verrucosa). (c) geographical distribution of each GSH. Filled circles represent data from this study while hashed ones were taken from the literature [based on mtORF identifications; colours refer to the GSH and black denotes multiple sympatric GSHs (indicated alongside) or ambiguous identifications (no GSH indicated)].

406



18

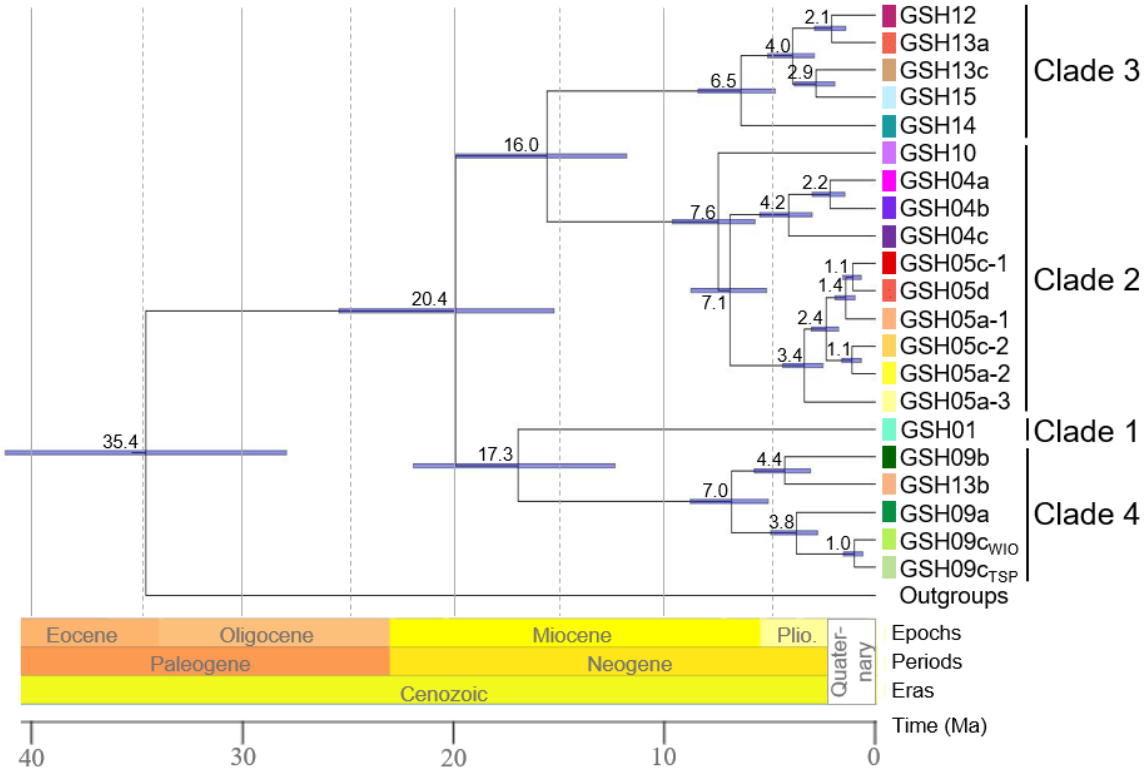


Fig. 3 Time-calibrated phylogeny of *Pocillopora* genomic species hypotheses (GSHs). Values above nodes indicate median node ages and blue bars represent the 95% highest posterior density (HPD) interval. Plio.: Pliocene.

Macro- and Micromorphological Analyses

Morphotypes based on macromorphology were not exclusive of a single GSH. Indeed, each GSH usually grouped colonies with a dominant morphotype, but also included several other morphotypes (e.g., GSH09b mostly grouped *P. meandrina*-like colonies, but also *P. damicornis*-like, *P. grandis*-like or *P. verrucosa*-like). Reciprocally, colonies from different GSHs can share the same morphotype (e.g., *P. damicornis*-like colonies were found in 14 GSHs; Fig. 2). Clades 1 and 4 were mostly characterised by robust morphs with large branches, while Clades 2 and 3 grouped more stunted colonies (Fig. 2 & Appendix 5).

Concerning micromorphology, intraspecific variations were smaller (Fig. 2 & S9 in Appendix 3). In particular, all species from Clades 1 and 4 (except GSH13b) and GSH15 were characterised by a styliform (GSH01, GSH09c$_{WIO}$ and GSH09c$_{TSP}$) or oval-convex (GSH09a, GSH09b and GSH15) columella, while all other species had a flat, more or less spinulate columella.

Accordingly, significant differences among GSHs were found for the columella diameter variables (v6 and v7; non-parametric permutational anovas; v6: $F_{(5,46)} = 92.95$, $P < 10^{-3}$***; v7: $F_{(5,46)} = 98.20$, $P < 10^{-3}$***), distinguishing three groups: GSH01 + GSH15, GSH09a + GSH09b and GSH09c$_{WIO}$ + GSH09c$_{TSP}$ (pairwise permutational t tests; $P < 0.05^{*}$; Fig. S9). Significant differences among GSHs were also found for all other five numeric morphological variables (v1-v5; non-parametric permutational anovas; $4.53 \leq F_{(19,150)} \leq 18.96$; $P < 10^{-3}$***), but no particular pattern was identified, except that GSHs from Clade 2, GSH13a and GSH13b had poorly developed septa (Fig. 2 & S9). The PERMANOVA and FAMD (Fig. S9 & S10 in Appendix 3) also highlighted these differences. Five micromorphological groups were thus distinguished on the first three principal components of the FAMD (explaining 68.4% of the variability): GSH01, Clade 2 + GSH13a + GSH13b, GSH12 + GSH13c + GSH14, GSH09a + GSH09b + GSH15 and GSH09c$_{WIO}$ + GSH09c$_{TSP}$ (Fig. 2 & S10). Detailed macromorphological and micromorphological illustrations of the GSHs are provided in Appendix 5.

Characterisation of Associated Symbiodiniaceae

ITS2 amplicon sequencing yielded a total of 1.6×10^7 reads (4.0×10^9 bp) with between 1.7×10^4 to 1.2×10^5 reads per individual [mean ± s.e. = $(6.1 \pm 0.1) \times 10^4$ reads]. After merging paired reads and removing chimeras, 9.0×10^6 sequences were retained [with between 0 and 7.2×10^4 sequences per individual; mean ± s.e. = $(3.5 \pm 0.0) \times 10^4$ sequences], corresponding to 1,014 amplicon sequence variants that were clustered in 590 operational taxonomic units (OTUs; represented by 1 to 6.1×10^5 sequences). Finally, 534 OTUs (90.5%) were taxonomically assigned, with a majority (511 OTUs, representing 97.6% of the sequences) belonging to *Cladocopium* (formerly *Symbiodinium* clade C), and mostly to clades C1 (173 OTUs and 35.6% of the sequences), C40 (267 OTUs; 56.6% of the sequences) and C42 (45 OTUs; 3.6% of the sequences). The other OTUs were assigned to *Symbiodinium* (clade A1; 6 OTUs; 0.2% of the sequences), *Durisdinium* (clade D1; 2 OTUs; < 0.1% of the sequences), *Gerakladium* (clade G3; 12 OTUs;

0.1% of the sequences) and Symbiodiniaceae clade I (clades I1 and I3; 3 OTUs; < 0.1% of the sequences). The reconstructed phylogeny based on these OTUs retrieved the five genera, with largely unresolved polytomies within *Cladocopium*, as previously observed (LaJeunesse 2005; Brener-Raffalli et al. 2018; Fig. S11 in Appendix 4). Nevertheless, such polytomies should not affect subsequent analyses, as performed at the OTU level.

From the remaining 252 individuals and 552 OTUs that passed the filtration steps, OTU richness within colonies varied from 0.14 to 2.67 for Shannon diversity index, and from 2 to 39 for Chao1 index. Both indexes were significantly different among GSHs (non-parametric permutational anova; Shannon: $F_{(20,231)} = 3.89$, $P < 10^{-3}$***; Chao1: $F_{(20,231)} = 2.96$, $P < 10^{-3}$***; Fig. S12 in Appendix 4), but no significant difference was found in Chao1 post-hoc tests (pairwise permutational t tests; $P > 0.05^{NS}$), and no obvious pattern was found for Shannon (Fig. S12). Differences were clearer when looking at the proportion of each taxon within samples (Fig. S13 in Appendix 4). For example, individuals from GSH04c, GSH09a and GSH13a displayed mainly C1ky [38.8 ± 2.4% on average (± s.e.)], while it was almost absent in other GSHs. Similarly, GSH05c-2, GSH13c and GSH15 contained mainly C1ag (56.5 ± 4.6%) and GSH05c-1, GSH05d, GSH12 and GSH14 mainly C1d (50.0 ± 4.4%). Except few individuals, other GSHs contained almost exclusively C40c (Fig. S13). Accordingly, the NMDS based on Bray and Curtis (1957) dissimilarity index followed this partitioning with three groups on the first two principal components (explaining 41% of the variability): (1) individuals mostly composed of C1ky, (2) those mostly composed of C1ag or C1d (separated with the third principal component) and (3) individuals mostly composed of C40c (Table 1 & Fig. S14 in Appendix 4).

Concerning localities, significant differences were found for Chao1 (again, without any obvious pattern; non-parametric permutational anova; $F_{(13,238)} = 5.94$, $P < 10^{-3}$***), but neither for Shannon (non-parametric permutational anova; $F_{(13,238)} = 1.69$, $P = 0.07^{NS}$; Fig. S12), nor by looking the individual proportions (Fig. S13) or the NMDS (Fig. S14).

Summary of all Evidences

Genomic analyses allowed the definition of 21 GSHs, while only four species hypotheses were distinguished if based only on Symbiodiniaceae communities, and up to 10 based only on qualitative micromorphology. Thus, combining evidences from all approaches (i.e., genomics, genetics, macro- and micromorphology, and symbiosis ecology) and considering the existence of two species only if supported by all criteria, only one single unambiguous species (corresponding to the entire genus) could be delimited (i.e., no separation appears fully supported; Table 1). Removing the Symbiodiniaceae criterion, five species, corresponding to Clades 1-3, GSH13b and GSH09 *sensu lato*, could be delimited. Then, sequentially removing the macro- and micromorphology criteria (i.e., considering all genetic evidences alone) could lead to nine and 12 species (Table 1). However, three out of the four GSHs within GSH09 *sensu lato* are split by other criteria (morphology and sometimes Symbiodiniaceae), supporting some genetic and genomic evidences. Conversely, considering two species once a single criterion separates them, genomics alone allows to distinguish all partitions. Consequently, we discuss below the usefulness of each criterion and propose a parsimonious consensus of 13 species strongly supported by most approaches (Table 1). Among them, three (*P. acuta*, *P. damicornis* and *P. grandis*) could represent species complexes according genetic evidences, and six were not attributed to a currently valid species, potentially representing new species (Table 1).

DISCUSSION

Although accurately delimiting species remains of particular importance and requires integrating multiple criteria, all investigated criteria do not provide the same resolution nor congruent insights. Not all criteria should therefore obviously be considered equally in order to define a consensus of the species limits as parsimonious as possible. As an illustration, both pigs and humans have four limbs, udders, etc., but that does not mean they belong to the same species. Conversely, eye colour does not distinguish different species in humans. Thus, in this study,

focusing on the scleractinian genus *Pocillopora* across a wide range of sampled localities (18 islands or regions from three marine provinces), genetic, morphological, geographical and symbiosis data were collected and compared to define robust species limits and assess the usefulness of each criterion. The different genetic approaches allowed to delimit 21 genomic species hypotheses (GSHs) where only seven are currently recognised based on current taxonomy. Moreover, 13 species appear strongly supported by all approaches, supporting the presence of six potentially new species that need to be formally described. Some of the other GSHs were supported by biogeographic or symbiosis evidences, but additional investigations are needed to state on their species status. In any case, a taxonomical revision of the *Pocillopora* genus, taking into account evidences brought by these results and previous ones, becomes urgent. This will allow to give formal names to the new species and thus throw off the multitude of current nomenclatures based on genetic lineages which can be difficult to follow, even for specialists.

On the (Ir)Relevance of Symbiosis Ecology to Define Species

As many scleractinian genera, *Pocillopora* species host diverse communities of symbionts (Cunning et al. 2017; Brener-Raffalli et al. 2018; Li et al. 2021; Rabbani et al. 2021). In this genus, Symbiodiniaceae are expected to be maternally transmitted (vertical transmission), as they are already present in oocytes before spawning (Sier and Olive 1994; Hirose et al. 2001; Harii et al. 2002). Such symbiont inheritance could result in species-specific associations and co-evolutions (Pinzón and LaJeunesse 2011; Schmidt-Roach et al. 2012; Johnston et al. 2022), that could also be responsible for habitat specialisations (driven by symbionts thermotolerance and photosynthetic needs; Jokiel and York 1982; Baker et al. 2013; Brener-Raffalli et al. 2018; Ros et al. 2021). Characterising associated Symbiodiniaceae communities can therefore bring additional elements to the delimitation of *Pocillopora* species, as in other scleractinian genera (Bongaerts et al. 2010; Keshavmurthy et al. 2013; Warner et al. 2015; Arrigoni et al. 2016; Forsman et al. 2020), but this does not guarantee a self-sufficient criterion.

Indeed, symbiosis ecology alone does not appear informative enough to delimit species, as evidenced by our results. We found a high prevalence of *Cladocopium* C1 (*C. goreaui*) and C40, both host-generalists, consistently with other studies on *Pocillopora* (e.g., Magalon et al. 2007; Pinzón and LaJeunesse 2011; Schmidt-Roach et al. 2012; Brener-Raffalli et al. 2018; Armstrong et al. 2021; Johnston et al. 2022). C1 variants allowed to distinguish five groups of colonies with distinct Symbiodiniaceae communities, but colonies within those groups were very distinct morphologically and genetically. Conversely, colonies from a single GSH generally shared the same communities.

These results should nevertheless be considered cautiously as (1) host-symbiont associations may vary over time and depth (Cunning et al. 2013), and (2) quantitative interpretation of metabarcoding results can be misleading (Lamb et al. 2019). First, Pinzón and LaJeunesse (2011) found that *Pocillopora* type 1 (ORF27; probably GSH09b or GSH09c$_{TSP}$) was the only type associated to the thermotolerant *Durusdinium glynnii* (D1; Wham et al. 2017) in the tropical eastern Pacific. But it was later found in *Pocillopora* types 3a and 3b (ORF46 and ORF47; GSH13b and GSH13c, respectively), with different prevalence among sites (Cunning et al. 2013), suggesting variable host-symbiont associations. In particular, *Durusdinium* would represent an opportunist genus, replacing specialist symbionts in health-compromised (e.g., bleached) corals (Stat and Gates 2010), potentially explaining these results. In our study, *Durusdinium* was rare (representing ca. 0.5% of the sequences in a single individual), suggesting no recent bleaching event prior to sampling, and thus mature host-symbiont associations. However, horizontal (i.e., from the water column) acquisition of Symbiodiniaceae remains possible, potentially corrupting species-specific associations. Second, PCR inherent biases (reviewed in Lamb et al. 2019) can result in differential sequence amplifications, either quantitatively or qualitatively. This can result in artificial differences in Symbiodiniaceae compositions among individuals and GSHs. Conversely, rare or specific Symbiodiniaceae taxa that could be diagnostic of a GSH might not be amplified, sequenced or detected.

Species limits evidences from symbiosis ecology inferred with metabarcoding data should therefore be taken cautiously, and rather used in support of other criteria in an integrative context. Besides, this criterion has not been systematically explored in previous taxonomic revisions of scleractinian genera (e.g., Benzoni et al. 2010; Arrigoni et al. 2020, 2021; Wepfer et al. 2020), demonstrating that it is not the most relevant criterion.

Should we Trust Morphology?

While most of the delimited GSHs grouped colonies with one major morphotype (which could be shared between GSHs), they also harboured high morphotype diversities. This demonstrates, once again (e.g., Pinzón et al. 2013; Marti-Puig et al. 2014; Gélin et al. 2017b), the obsolescence of *corallum* macromorphology to define *Pocillopora* species limits, as in other scleractinian genera (e.g., Warner et al. 2015; Shimpi et al. 2019; Bongaerts et al. 2021; Terraneo et al. 2021). Indeed, *Pocillopora* corals can display great morphological plasticity mostly driven by light and currents (Todd 2008). As an illustration, in the Gulf of California, five morphospecies have been reported (Glynn and Ault 2000), all belonging to mtORF type 1a (= ORF27; Pinzón et al. 2013). Switches from one morphospecies to another have also been demonstrated following shifts in environmental conditions (Paz-García et al. 2015a, 2015b).

Contrary to macromorphology, micromorphology brought additional insights to the refining of *Pocillopora* species limits, as in other scleractinian genera (e.g., Benzoni et al. 2007; Forsman et al. 2010; Budd and Stolarski 2011; Stefani et al. 2011; Budd et al. 2012; Arrigoni et al. 2020). Intraspecific variations were smaller, and several differences, either qualitative or quantitative, allowed to distinguish almost all GSHs. The GSHs within Clade 2 were not separated, but Schmidt-Roach et al. (2014) raised several differences that we could not recover. It is also possible that the morphological characters investigated here were not the most relevant to distinguish these GSHs (e.g., as for the number of limbs to distinguish humans and pigs).

Morphology-based criteria are thus questionable and subject to interpretation (particularly for the presence/absence of subtle characters) which, coupled with morphological plasticity, makes them unsuitable for identifying *Pocillopora* species. The misidentification of two out of the three currently available *Pocillopora* genomes perfectly illustrates this point. While the *P. verrucosa* genome (Buitrago-López et al. 2020) has been assigned to a GSH consistent with this identification (GSH13a), the two others were not [the *P. acuta* genome (Vidal-Dupiol et al. 2019) was assigned to GSH13a (currently considered as *P. verrucosa*) too and the *P. damicornis* genome (Cunning et al. 2018) to GSH09c$_{TSP}$ (*P. grandis*)]. Surprisingly, the colony sequenced for the *P. acuta* genome has been identified molecularly using the mtORF, but the haplotype was not provided (Vidal-Dupiol et al. 2019), so we could not verify the identification.

Exploring Species Limits: Lessons from Genomics

Pocillopora species limits have been extensively studied using genetic markers over the past decades (e.g., Schmidt-Roach et al. 2012; Pinzón et al. 2013; Marti-Puig et al. 2014; Gélin et al. 2017b), revealing a great diversity within some morphospecies (e.g., *P. damicornis*; Schmidt-Roach et al. 2012). Most of these previous studies used mtDNA and microsatellites to explore species limits. Only Johnston et al. (2017, 2022) inferred genetic relationships among few tens of *Pocillopora* colonies from the Pacific using genomic data. Consequently, our study represents the most extensive investigation to date of the taxonomy of the *Pocillopora* genus using genomics.

Our genomic analyses based on SNPs collected from the sequence capture of UCEs and exon loci provided very congruent results among methods and allowed the robust definition of four main clades comprising 21 GSHs. However, despite thousands of SNPs and loci analysed, we were not able to fully resolve GSH relationships, and multiple species tree topologies were inferred (Fig. 2 & S8 in Appendix 2). Recent species divergences and the presence of several closely related sister species, as well as introgression, could explain unresolved topologies. Indeed, most GSHs appeared less than 5 Ma, with a substantial number in the Quaternary (i.e., 0-2.6 Ma; Fig. 3). This suggests a

recent radiation, probably linked to major geological and climatic events during the Pliocene or the Pleistocene [e.g., changes in currents (Philander and Fedorov 2003), glacial-interglacial cycles (Adams et al. 1999; Lambeck et al. 2002) and formation of the Isthmus of Panama (O'Dea et al. 2016)], as already suggested in this genus (Johnston et al. 2017). Recent divergences also suggest that some sister GSHs might still be in speciation and are in the grey zone (*sensu* De Queiroz 2007) where distinctive characters are set up and gene flow are still possible. Not all investigated criteria can therefore distinguish them and the question of their validity as two distinct species arises. However, since they harbour distinct allelic states for the SNPs used, and since some SNPs are coding, differential characters between these GSHs are expected. The question is whether these characters allow to distinguish species (e.g., eye colour in humans is encoded by over 150 genes, resulting in many SNPs, and yet it is still a single species). Therefore, parsimoniously, these GSHs should be considered as a single species that potentially represents a species complex (e.g., *P. damicornis* with two GSHs or *P. acuta* with six GSHs), waiting for further (e.g., ecological or reproductive) evidences to separate them.

Interestingly, almost all 21 GSHs corresponded to previously defined genetic species hypotheses or clusters (based on the mtORF marker and microsatellites). Several GSHs had their own mtORF or PocHistone haplotypes, confirming that both can be used as diagnostic markers for some (but not all) *Pocillopora* species. Conversely, the over-partitioning previously found in several SSHs using microsatellites (e.g., Gélin et al. 2018a; Oury et al. 2021) was not retrieved. This could be an effect either of the limited numbers of loci in microsatellite inferences, or of genus-level phylogenetic inferences masking such genetic patterns. Genetic criteria therefore appear robust to define species limits but present a risk of overestimating their number. BFD*, as other molecular species delimitation methods, has already been suggested to overestimate the number of species (Grummer et al. 2014; Hundsdoerfer et al. 2019; Derkarabetian et al. 2022). This supports the need for integrative approaches, where molecular criteria should be the first criteria to robustly and objectively explore species limits and define genetic species hypotheses that are then confirmed

with other criteria (as previously suggested by Pante et al. 2015). In particular, genomics, although not systematically necessary to molecularly identify species, appears fundamental to set robust species limits in such taxa whose phylogenetic reconstructions are complex. So for biodiversity monitoring (e.g., the global coral reef monitoring network), for each region, an exhaustive inventory of *Pocillopora* species using genomics/genetics is precognised to first identify the species present in the field.

From Multiple Criteria to Integrative Taxonomy: Towards a Revision of the Pocillopora *Genus*

Putting together evidences from all approaches (i.e., genetics, morphology, geography and symbiosis ecology), 13 species appeared strongly supported, where only seven are currently recognised based on current taxonomy. Six species thus need formal taxonomic descriptions (Table 1). Clades 1 and 2 support current taxonomy, the first consisting of a single species (*P. effusa*, corresponding to GSH01), and the second being consistent with Schmidt-Roach et al. (2014) taxonomic revision [i.e., three species: *P. damicornis* (GSH04 *sensu lato*), *P. brevicornis* (GSH10) and *P. acuta* (GSH05 *sensu lato*)]. Further investigations are nevertheless needed to state whether *P. damicornis* and *P. acuta* represent both species complexes. Indeed, *P. damicornis* was separated into two GSHs and SSHs (04a and 04b) not supported by other criteria, but which could be ecologically distinct as previously suggested (Oury et al. 2020a). Similarly, *P. acuta* was partitioned into several GSHs and SSHs/clusters, either sympatric or allopatric, and some associated to distinct Symbiodiniaceae. Multiple genetic entities were previously delimited in this species (Gélin et al. 2017a, 2018b; Torres et al. 2020), questioning its monophyly. Clades 3 and 4 are less congruent with current taxonomy. First, within Clade 3, all five GSHs are strongly supported by all other criteria, but we were not able to rely them with a currently accepted species (one of them, probably GSH13c, should correspond to *P. verrucosa*, but the others seem to be new species). Then, within Clade 4, four species seemed strongly supported. GSH09a, GSH09b and GSH13b each correspond to distinct species (GSH09b most probably corresponding to *P. meandrina* Dana, 1846,

while *P. villosa nomen nudem* was previously suggested for GSH13b; Gélin et al. 2017b). Only $GSH09c_{WIO}$ and $GSH09c_{TSP}$ could not be distinguished with certainty and are parsimoniously considered as different *P. grandis* lineages in allopatry for now, waiting further evidences.

In the light of these results, a new taxonomical revision of the *Pocillopora* genus, formally describing and naming these six new species (corresponding to GSH09a, GSH12, GSH13a, GSH13b, GSH14 and GSH15) becomes urgent. This will allow to throw off the multitude of current nomenclatures based on genetic lineages (Pinzón et al. 2013; Marti-Puig et al. 2014; Schmidt-Roach et al. 2014; Gélin et al. 2017b) which can be difficult to follow, even for specialists.

In conclusion, this study is the most extensive exploration to date of the taxonomy of the *Pocillopora* genus in terms of both genomic and geographic coverage. This genus represents a scleractinian taxon for which the definition of species limits has been challenging for decades. Several other criteria including morphology, biogeography or symbiosis ecology were also investigated to refine species limits and propose consensually and parsimoniously species hypotheses in the most integrative way possible. Some criteria appeared thus more informative than others, but all provided helpful insights for refining species limits. Here, we clearly delimited 21 genomic species hypotheses from 356 colonies sampled in three marine provinces (western Indian Ocean, tropical southwestern Pacific and south-east Polynesia), of which 13 species were strongly supported by all approaches and six appear to be new species. Importantly, we demonstrate once again the obsolescence of *corallum* macromorphology to identify most of the species. Conversely, micromorphological diagnostic characters and mtORF and PocHistone diagnostic haplotypes were highlighted for several species. Our recommendation is therefore to systematically identify *Pocillopora* species using these diagnostic criteria, prior to all types of studies involving the colonies (e.g., biodiversity, ecology, reproduction, adaptation, connectivity, exo- and endo-symbiosis…) in order to reduce misidentifications. Finally, our results give new insights into the

puzzle of defining *Pocillopora* species limits, supporting the existence of several new species. Next steps are to formally revise the taxonomy of the *Pocillopora* genus.

ACKNOWLEDGEMENTS

Coral sampling in New Caledonia was carried out during COBELO (http://dx.doi.org/10.17600/13100100), BIBELOT (http://dx.doi.org/10.17600/14003700), and CHEST (http://dx.doi.org/10.17600/15004500) oceanographic campaigns on board of RV Alis (IRD), and in the northeastern and northwestern of Madagascar during MAD (http://dx.doi.org/10.17600/16004700) oceanographic campaign on board of RV Antea (IRD). Sampling in Reunion Island was supported by program CONPOCINPA (LabEx CORAIL fund); in the south of Madagascar in collaboration with the Institut Halieutique des Sciences Marines (Tulear); and in Rodrigues Island with the collaboration of the Rodrigues Regional Assembly and the South-East Marine Protected Area supported by project Biodiversity (POCT FEDER fund); in Europa, Juan de Nova, and Glorioso Islands by program BIORECIE (financial supports from INEE, INSU, IRD, AAMP, FRB, TAAF, and the foundation Veolia Environnement); in Tromelin Island by program ORCIE (INEE), and in Mayotte by program SIREME (FED). HM thanks all the buddies who helped in photographs during diving (J. Butscher, S. Andréfouët, L. Bigot, and M. Pinault). We acknowledge the Plateforme Gentyane (microsatellite genotyping) of the Institut National de Recherche pour l'Agriculture, l'alimentation et l'Environnement (INRAE, Clermont-Ferrand, France), GenoScreen (Sanger sequencing; Lille, France), the Plateforme iGenSeq (library preparations and NGS sequencing) of the Institut du Cerveau et de la Moelle épinière (ICM, Paris, France), and G. Toutirais from the Plateau technique de Microscopie Électronique (scanning electron microscopy) of the Muséum National d'Histoire Naturelle (MNHN, Paris, France) for technical supports. We would also like to acknowledge S. Schmidt-Roach for his valuable advice on preparing SEM samples. Bioinformatics analyses were performed on the Genotoul bioinformatics platform Toulouse Occitanie (Bioinfo Genotoul,

https://doi.org/10.15454/1.5572369328961167E12). NO was financially supported by a PhD contract from the Doctoral School "Sciences, Technologies, Santé" of Reunion Island University.

FUNDING

This work was supported in part through grants from the LabEx CORAIL (AI PocillopoRAD).

DATA AVAILABILITY STATEMENTS

All data underlying this article are available online or upon reasonable request to the corresponding author. Raw sequencing reads (BioProject PRJNA831687; *Pocillopora* sequence capture: accession numbers SRR19052129-SRR19052500; Symbiodiniaceae ITS2 metabarcoding: accession numbers SRR19152377-SRR19152635) and new haplotype sequences (GenBank accession numbers ON155826-ON155833) were deposited on the NCBI. Microsatellite genotypes, morphometric data, reference sequences and SNP datasets were deposited on Dryad: https://doi.org/XXXXXXXXX.

AUTHOR CONTRIBUTION STATEMENTS

NO and HM designed the study. HM collected samples. NO and HM did lab steps. CN performed the bioinformatics for the ITS2 metabarcoding. NO performed all other bioinformatics and analysed the results with helpful guidance from SM and DA. NO wrote the original draft and all authors reviewed and edited the manuscript.

REFERENCES

Adams J., Maslin M., Thomas E. 1999. Sudden climate transitions during the Quaternary. Prog. Phys. Geogr. 23:1–36.

732 Anderson E.C., Thompson E.A. 2002. A model-based method for identifying species hybrids using
733 multilocus genetic data. Genetics. 160:1217–1229.
734 Armstrong E.J., Lê-Hoang J., Carradec Q., Aury J.-M., Noel B., Poulain J., Belser C., Da Silva C.,
735 Wincker P., Tara Pacific Consortium. 2021. Transcriptomic plasticity and symbiont
736 shuffling underpin *Pocillopora* acclimatization across heat-stress regimes in the Pacific
737 Ocean. BioRxiv.
738 Arrigoni R., Benzoni F., Terraneo T.I., Caragnano A., Berumen M.L. 2016. Recent origin and semi-
739 permeable species boundaries in the scleractinian coral genus *Stylophora* from the Red Sea.
740 Sci. Rep. 6:34612.
741 Arrigoni R., Berumen M.L., Mariappan K.G., Beck P.S.A., Hulver A.M., Montano S., Pichon M.,
742 Strona G., Terraneo T.I., Benzoni F. 2020. Towards a rigorous species delimitation
743 framework for scleractinian corals based on RAD sequencing: the case study of *Leptastrea*
744 from the Indo-Pacific. Coral Reefs. 39:1001–1025.
745 Arrigoni R., Huang D., Berumen M.L., Budd A.F., Montano S., Richards Z.T., Terraneo T.I.,
746 Benzoni F. 2021. Integrative systematics of the scleractinian coral genera *Caulastraea* ,
747 *Erythrastrea* and *Oulophyllia*. Zool. Scr.:zsc.12481.
748 Baird N.A., Etter P.D., Atwood T.S., Currey M.C., Shiver A.L., Lewis Z.A., Selker E.U., Cresko
749 W.A., Johnson E.A. 2008. Rapid SNP discovery and genetic mapping using sequenced
750 RAD markers. PLoS ONE. 3:e3376.
751 Baker A.C., McClanahan T.R., Starger C.J., Boonstra R.K. 2013. Long-term monitoring of algal
752 symbiont communities in corals reveals stability is taxon dependent and driven by site-
753 specific thermal regime. Mar. Ecol. Prog. Ser. 479:85–97.
754 Benzoni F., Bianchi C.N., Morri C. 2003. Coral communities of the northwestern Gulf of Aden
755 (Yemen): variation in framework building related to environmental factors and biotic
756 conditions. Coral Reefs. 22:475–484.

Benzoni F., Stefani F., Pichon M., Galli P. 2010. The name game: morpho-molecular species boundaries in the genus *Psammocora* (Cnidaria, Scleractinia). Zool. J. Linn. Soc. 160:421–456.

Benzoni F., Stefani F., Stolarski J., Pichon M., Mitta G., Galli P. 2007. Debating phylogenetic relationships of the scleractinian *Psammocora*: molecular and morphological evidences. Contrib. Zool. 76:35–54.

Bongaerts P., Cooke I.R., Ying H., Wels D., den Haan S., Hernandez-Agreda A., Brunner C.A., Dove S., Englebert N., Eyal G., Forêt S., Grinblat M., Hay K.B., Harii S., Hayward D.C., Lin Y., Mihaljević M., Moya A., Muir P., Sinniger F., Smallhorn-West P., Torda G., Ragan M.A., van Oppen M.J.H., Hoegh-Guldberg O. 2021. Morphological stasis masks ecologically divergent coral species on tropical reefs. Curr. Biol. 31:2286–2298.

Bongaerts P., Riginos C., Ridgway T., Sampayo E.M., van Oppen M.J.H., Englebert N., Vermeulen F., Hoegh-Guldberg O. 2010. Genetic divergence across habitats in the widespread coral *Seriatopora hystrix* and its associated *Symbiodinium*. PLoS ONE. 5:10871.

Bouckaert R., Vaughan T.G., Barido-Sottani J., Duchêne S., Fourment M., Gavryushkina A., Heled J., Jones G., Kühnert D., Maio N.D., Matschiner M., Mendes F.K., Müller N.F., Ogilvie H.A., Plessis L. du, 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. 15:1006650.

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

Bray J.R., Curtis J.T. 1957. An ordination of the upland forest communities of southern Wisconsin. Ecol. Monogr. 27:326–349.

Brener-Raffalli K., Clerissi C., Vidal-Dupiol J., Adjeroud M., Bonhomme F., Pratlong M., Aurelle D., Mitta G., Toulza E. 2018. Thermal regime and host clade, rather than geography, drive

Symbiodinium and bacterial assemblages in the scleractinian coral *Pocillopora damicornis sensu lato*. Microbiome. 6:39.

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., Fukami H., Smith N.D., Knowlton N. 2012. Taxonomic classification of the reef coral family Mussidae (Cnidaria: Anthozoa: Scleractinia). Zool. J. Linn. Soc. 166:465–529.

Budd A.F., Stolarski J. 2011. Corallite wall and septal microstructure in scleractinian reef corals: comparison of molecular clades within the family Faviidae. J. Morphol. 272:66–88.

Buitrago-López C., Mariappan K.G., Cardenas A., Gegner H.M., Voolstra C.R. 2020. The genome of the cauliflower coral *Pocillopora verrucosa*. Genome Biol. Evol. 12:1911–1917.

Chen C.A., Chang C.-C., Wei N.V., Chen C.-H., Lein Y.-T., Lin H.-E., Dai C.-F., Wallace C.C. 2004. Secondary structure and phylogenetic utility of the ribosomal internal transcribed spacer 2 (ITS2) in scleractinian corals. Zool Stud. 43:759–771.

Chevalier J.-P. 1971. Les scléractiniaires de la Mélanésie française (Nouvelle-Calédonie, Iles Chesterfield, Iles Loyauté, Nouvelles-Hébrides): 1ère partie. Fondation Singer-Polignac.

Combosch D.J., Guzman H.M., Schuhmacher H., Vollmer S.V. 2008. Interspecific hybridization and restricted trans-Pacific gene flow in the Tropical Eastern Pacific *Pocillopora*. Mol. Ecol. 17:1304–1312.

Combosch D.J., Vollmer S.V. 2015. Trans-Pacific RAD-Seq population genomics confirms introgressive hybridization in Eastern Pacific *Pocillopora* corals. Mol. Phylogenet. Evol. 88:154–162.

Cunha R.L., Forsman Z.H., Belderok R., Knapp I.S.S., Castilho R., Toonen R.J. 2019. Rare coral under the genomic microscope: timing and relationships among Hawaiian *Montipora*. BMC Evol. Biol. 19:153.

807 Cunning R., Bay R.A., Gillette P., Baker A.C., Traylor-Knowles N. 2018. Comparative analysis of
808 the *Pocillopora damicornis* genome highlights role of immune system in coral evolution.
809 Sci. Rep. 8:16134.
810 Cunning R., Gates R.D., Edmunds P.J. 2017. Using high-throughput sequencing of ITS2 to describe
811 *Symbiodinium* metacommunities in St. John, US Virgin Islands. PeerJ. 5:3472.
812 Cunning R., Glynn P.W., Baker A.C. 2013. Flexible associations between *Pocillopora* corals and
813 *Symbiodinium* limit utility of symbiosis ecology in defining species. Coral Reefs. 32:795–
814 801.
815 Davey J.W., Hohenlohe P.A., Etter P.D., Boone J.Q., Catchen J.M., Blaxter M.L. 2011. Genome-
816 wide genetic marker discovery and genotyping using next-generation sequencing. Nat. Rev.
817 Genet. 12:499–510.
818 De Queiroz K. 2005. A unified concept of species and its consequences for the future of taxonomy.
819 Proc. Calif. Acad. Sci. 56(Suppl 1):196–215.
820 De Queiroz K. 2007. Species concepts and species delimitation. Syst. Biol. 56:879–886.
821 Derkarabetian S., Starrett J., Hedin M. 2022. Using natural history to guide supervised machine
822 learning for cryptic species delimitation with genetic data. Front. Zool. 19:8.
823 Faircloth B.C., McCormack J.E., Crawford N.G., Harvey M.G., Brumfield R.T., Glenn T.C. 2012.
824 Ultraconserved elements anchor thousands of genetic markers spanning multiple
825 evolutionary timescales. Syst. Biol. 61:717–726.
826 Faircloth B.C., Sorenson L., Santini F., Alfaro M.E. 2013. A phylogenomic perspective on the
827 radiation of ray-finned fishes based upon targeted sequencing of ultraconserved elements
828 (UCEs). PLoS ONE. 8:65923.
829 Flot J.-F., Tillier S. 2007. The mitochondrial genome of *Pocillopora* (Cnidaria: Scleractinia)
830 contains two variable regions: the putative D-loop and a novel ORF of unknown function.
831 Gene. 401:80–87.

832 Forsman Z.H., Barshis D.J., Hunter C.L., Toonen R.J. 2009. Shape-shifting corals: molecular
833 markers show morphology is evolutionarily plastic in *Porites*. BMC Evol. Biol. 9:45.
834 Forsman Z.H., Concepcion G.T., Haverkort R.D., Shaw R.W., Maragos J.E., Toonen R.J. 2010.
835 Ecomorph or endangered coral? DNA and microstructure reveal Hawaiian species
836 complexes: *Montipora dilatata/flabellata/turgescens* & *M. patula/verrilli*. PLoS ONE.
837 5:15021.
838 Forsman Z.H., Knapp I.S.S., Tisthammer K., Eaton D.A.R., Belcaid M., Toonen R.J. 2017. Coral
839 hybridization or phenotypic variation? Genomic data reveal gene flow between *Porites*
840 *lobata* and *P. compressa*. Mol. Phylogenet. Evol. 111:132–148.
841 Forsman Z.H., Ritson-Williams R., Tisthammer K.H., Knapp I.S.S., Toonen R.J. 2020. Host-
842 symbiont coevolution, cryptic structure, and bleaching susceptibility, in a coral species
843 complex (Scleractinia; Poritidae). Sci. Rep. 10:16995.
844 Frichot E., Mathieu F., Trouillon T., Bouchard G., François O. 2014. Fast and efficient estimation
845 of individual ancestry coefficients. Genetics. 196:973–983.
846 Fukami H., Chen C.A., Budd A.F., Collins A., Wallace C., Chuang Y.-Y., Chen C., Dai C.-F., Iwao
847 K., Sheppard C., Knowlton N. 2008. Mitochondrial and nuclear genes suggest that stony
848 corals are monophyletic but most families of stony corals are not (Order Scleractinia, Class
849 Anthozoa, Phylum Cnidaria). PLoS ONE. 3:3222.
850 Gélin P., Fauvelot C., Bigot L., Baly J., Magalon H. 2018a. From population connectivity to the art
851 of striping Russian dolls: the lessons from *Pocillopora* corals. Ecol. Evol. 8:1411–1426.
852 Gélin P., Fauvelot C., Mehn V., Bureau S., Rouzé H., Magalon H. 2017a. Superclone expansion,
853 long-distance clonal dispersal and local genetic structuring in the coral *Pocillopora*
854 *damicornis* type *β* in Reunion Island, South Western Indian Ocean. PLoS ONE.
855 12:0169692.

Gélin P., Pirog A., Fauvelot C., Magalon H. 2018b. High genetic differentiation and low connectivity in the coral *Pocillopora damicornis* type β at different spatial scales in the Southwestern Indian Ocean and the Tropical Southwestern Pacific. Mar. Biol.:165–167.

Gélin P., Postaire B., Fauvelot C., Magalon H. 2017b. Reevaluating species number, distribution and endemism of the coral genus *Pocillopora* Lamarck, 1816 using species delimitation methods and microsatellites. Mol. Phylogenet. Evol. 109:430–446.

Glynn P.W., Ault J.S. 2000. A biogeographic analysis and review of the far eastern Pacific coral reef region. Coral Reefs. 19:1–23.

Gnirke A., Melnikov A., Maguire J., Rogov P., LeProust E.M., Brockman W., Fennell T., Giannoukos G., Fisher S., Russ C. 2009. Solution hybrid selection with ultra-long oligonucleotides for massively parallel targeted sequencing. Nat. Biotechnol. 27:182–189.

Grummer J.A., Bryson R.W. Jr., 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.

Harii S., Kayanne H., Takigawa H., Hayashibara T., Yamamoto M. 2002. Larval survivorship, competency periods and settlement of two brooding corals, *Heliopora coerulea* and *Pocillopora damicornis*. Mar. Biol. 141:39–46.

Harvey M.G., Smith B.T., Glenn T.C., Faircloth B.C., Brumfield R.T. 2016. Sequence capture versus restriction site associated DNA sequencing for shallow systematics. Syst. Biol. 65:910–924.

Hellberg M.E. 2006. No variation and low synonymous substitution rates in coral mtDNA despite high nuclear variation. BMC Evol. Biol. 6:24.

Hellberg M.E., Prada C., Tan M.H., Forsman Z.H., Baums I.B. 2016. Getting a grip at the edge: recolonization and introgression in eastern Pacific *Porites* corals. J. Biogeogr. 43:2147–2159.

881 Heron S.F., van Hooidonk R., Maynard J., Anderson K., Day J.C., Geiger E., Hoegh-Guldberg O.,
882 Hughes T., Marshall P., Obura D. 2018. Impacts of climate change on World Heritage coral
883 reefs: update to the first global scientific assessment. Paris: UNESCO World Heritage
884 Centre.
885 Hervé M. 2021. RVAideMemoire: testing and plotting procedures for biostatistics. R package
886 version 0.9-64.
887 Hirose M., Kinzie III R.A., Hidaka M. 2001. Timing and process of entry of zooxanthellae into
888 oocytes of hermatypic corals. Coral Reefs. 20:273–280.
889 Hodges E., Xuan Z., Balija V., Kramer M., Molla M.N., Smith S.W., Middle C.M., Rodesch M.J.,
890 Albert T.J., Hannon G.J., McCombie W.R. 2007. Genome-wide in situ exon capture for
891 selective resequencing. Nat. Genet. 39:1522–1527.
892 Hoeksema B.W., Cairns S.D. 2022. World list of Scleractinia. *Pocillopora* Lamarck, 1816.
893 Available from http://www.marinespecies.org/aphia.php?p=taxdetails&id=206938.
894 Hughes T.P., Anderson K.D., Connolly S.R., Heron S.F., Kerry J.T., Lough J.M., Baird A.H., Baum
895 J.K., Berumen M.L., Bridge T.C. 2018. Spatial and temporal patterns of mass bleaching of
896 corals in the Anthropocene. Science. 359:80–83.
897 Hughes T.P., Barnes M.L., Bellwood D.R., Cinner J.E., Cumming G.S., Jackson J.B., Kleypas J.,
898 Van De Leemput I.A., Lough J.M., Morrison T.H. 2017. Coral reefs in the Anthropocene.
899 Nature. 546:82–90.
900 Hughes T.P., Kerry J.T., Baird A.H., Connolly S.R., Chase T.J., Dietzel A., Hill T., Hoey A.S.,
901 Hoogenboom M.O., Jacobson M. 2019. Global warming impairs stock–recruitment
902 dynamics of corals. Nature. 568:387–390.
903 Hume B.C.C., Smith E.G., Ziegler M., Warrington H.J.M., Burt J.A., LaJeunesse T.C.,
904 Wiedenmann J., Voolstra C.R. 2019. SymPortal: a novel analytical framework and platform
905 for coral algal symbiont next□generation sequencing ITS2 profiling. Mol. Ecol. Resour.
906 19:1063–1080.

907 Hundsdoerfer A.K., Lee K.M., Kitching I.J., Mutanen M. 2019. Genome-wide SNP data reveal an
908 overestimation of species diversity in a group of Hawkmoths. Genome Biol. Evol. 11:2136–
909 2150.
910 Huson D.H., Bryant D. 2006. Application of phylogenetic networks in evolutionary studies. Mol.
911 Biol. Evol. 23:254–267.
912 Johnston E.C., Cunning R., Burgess S.C. 2022. Cophylogeny and specificity between cryptic coral
913 species (*Pocillopora* spp.) at Mo'orea and their symbionts (Symbiodiniaceae). BioRxiv.
914 Johnston E.C., Forsman Z.H., Flot J.-F., Schmidt-Roach S., Pinzón J.H., Knapp I.S.S., Toonen R.J.
915 2017. A genomic glance through the fog of plasticity and diversification in *Pocillopora*. Sci.
916 Rep. 7:5991.
917 Johnston E.C., Forsman Z.H., Toonen R.J. 2018. A simple molecular technique for distinguishing
918 species reveals frequent misidentification of Hawaiian corals in the genus *Pocillopora*.
919 PeerJ. 6:4355.
920 Jokiel P.L., York Jr. Richard H. 1982. Solar ultraviolet photobiology of the reef coral *Pocillopora*
921 *damicornis* and symbiotic zooxanthellae. Bull. Mar. Sci. 32:301–315.
922 Jombart T., Devillard S., Balloux F. 2010. Discriminant analysis of principal components: a new
923 method for the analysis of genetically structured populations. BMC Genet. 11:94.
924 Kass R.E., Raftery A.E. 1995. Bayes factors. J. Am. Stat. Assoc. 90:773–795.
925 Katoh K., Standley D.M. 2013. MAFFT multiple sequence alignment software version 7:
926 improvements in performance and usability. Mol. Biol. Evol. 30:772–780.
927 Keshavmurthy S., Yang S.-Y., Alamaru A., Chuang Y.-Y., Pichon M., Obura D., Fontana S., De
928 Palmas S., Stefani F., Benzoni F., MacDonald A., Noreen A.M.E., Chen C., Wallace C.C.,
929 Pillay R.M., Denis V., Amri A.Y., Reimer J.D., Mezaki T., Sheppard C., Loya Y., Abelson
930 A., Mohammed M.S., Baker A.C., Mostafavi P.G., Suharsono B.A., Chen C.A. 2013. DNA
931 barcoding reveals the coral "laboratory-rat", *Stylophora pistillata* encompasses multiple
932 identities. Sci. Rep. 3:1520.

933 Kivelä M., Arnaud-Haond S., Saramäki J. 2015. EDENetworks: a user-friendly software to build
934 and analyse networks in biogeography, ecology and population genetics. Mol. Ecol. Resour.
935 15:117–122.
936 Kozlov A.M., Darriba D., Flouri T., Morel B., Stamatakis A. 2019. RAxML-NG: a fast, scalable
937 and user-friendly tool for maximum likelihood phylogenetic inference. Bioinformatics.
938 35:4453–4455.
939 LaJeunesse T.C. 2005. “Species” radiations of symbiotic Dinoflagellates in the Atlantic and Indo-
940 Pacific since the Miocene-Pliocene transition. Mol. Biol. Evol. 22:570–581.
941 Lamb P.D., Hunter E., Pinnegar J.K., Creer S., Davies R.G., Taylor M.I. 2019. How quantitative is
942 metabarcoding: a meta-analytical approach. Mol. Ecol. 28:420–430.
943 Lambeck K., Esat T.M., Potter E.-K. 2002. Links between climate and sea levels for the past three
944 million years. Nature. 419:199–206.
945 Lê S., Josse J., Husson F. 2008. FactoMineR□: an R package for multivariate analysis. J. Stat.
946 Softw. 25.
947 Leaché A.D., Fujita M.K., Minin V.N., Bouckaert R.R. 2014. Species delimitation using genome-
948 wide SNP data. Syst. Biol. 63:534–542.
949 Li J., Long L., Zou Y., Zhang S. 2021. Microbial community and transcriptional responses to
950 increased temperatures in coral *Pocillopora damicornis* holobiont. Environ. Microbiol.
951 23:826–843.
952 Magalon H., Flot J.-F., Baudry E. 2007. Molecular identification of symbiotic dinoflagellates in
953 Pacific corals in the genus *Pocillopora*. Coral Reefs. 26:551–558.
954 Marti-Puig P., Forsman Z.H., Haverkort-Yeh R.D., Knapp I.S., Maragos J.E., Toonen R.J. 2014.
955 Extreme phenotypic polymorphism in the coral genus *Pocillopora*; micro-morphology
956 corresponds to mitochondrial groups, while colony morphology does not. Bull. Mar. Sci.
957 90:211–231.

Mayden R.L. 1997. A hierarchy of species concepts: the denouement in the saga of the species problem. Species: the units of diversity. Chapman & Hall. p. 381–423.

McCormack J.E., Faircloth B.C., Crawford N.G., Gowaty P.A., Brumfield R.T., Glenn T.C. 2012. Ultraconserved elements are novel phylogenomic markers that resolve placental mammal phylogeny when combined with species-tree analysis. Genome Res. 22:746–754.

Metzker M.L. 2010. Sequencing technologies—the next generation. Nat. Rev. Genet. 11:31–46.

Nakajima Y., Nishikawa A., Iguchi A., Nagata T., Uyeno D., Sakai K., Mitarai S. 2017. Elucidating the multiple genetic lineages and population genetic structure of the brooding coral *Seriatopora* (Scleractinia: Pocilloporidae) in the Ryukyu Archipelago. Coral Reefs. 36:415–426.

Narum S.R., Buerkle C.A., Davey J.W., Miller M.R., Hohenlohe P.A. 2013. Genotyping-by-sequencing in ecological and conservation genomics. Mol. Ecol. 22:2841–2847.

Nei M. 1972. Genetic distance between populations. Am. Nat. 106:283–292.

O'Dea A., Lessios H.A., Coates A.G., Eytan R.I., Restrepo-Moreno S.A., Cione A.L., Collins L.S., De Queiroz A., Farris D.W., Norris R.D. 2016. Formation of the Isthmus of Panama. Sci. Adv. 2:1600883.

Oksanen J., Blanchet F.G., Friendly M., Kindt R., Legendre P., McGlinn D., Minchin P.R., O'Hara R.B., Simpson G.L., Solymos P., Stevens M.H.H., Szoecs E., Wagner H. 2020. vegan: community ecology package. R package verson 2.5-7.

Oury N., Gélin P., Magalon H. 2020a. Cryptic species and genetic connectivity among populations of the coral *Pocillopora damicornis* (Scleractinia) in the tropical southwestern Pacific. Mar. Biol. 167:142.

Oury N., Gélin P., Magalon H. 2020b. Together stronger: Intracolonial genetic variability occurrence in *Pocillopora* corals suggests potential benefits. Ecol. Evol. 10:5208–5218.

Oury N., Gélin P., Magalon H. 2021. High connectivity within restricted distribution range in *Pocillopora* corals. J. Biogeogr. 48:1679–1692.

Oury N., Gélin P., Rajaonarivelo M., Magalon H. 2022. Exploring the *Pocillopora* cryptic diversity: a new genetic lineage in the Western Indian Ocean or remnants from an ancient one? Mar. Biodivers. 52:5.

Pante E., Puillandre N., Viricel A., Arnaud-Haond S., Aurelle D., Castelin M., Chenuil A., Destombe C., Forcioli D., Valero M., Viard F., Samadi S. 2015. Species are hypotheses: avoid connectivity assessments based on pillars of sand. Mol. Ecol. 24:525–544.

Paz-García D.A., Aldana-Moreno A., Cabral-Tena R.A., García-De-León F.J., Hellberg M.E., Balart E.F. 2015a. Morphological variation and different branch modularity across contrasting flow conditions in dominant *Pocillopora* reef-building corals. Oecologia. 178:207–218.

Paz-García D.A., Hellberg M.E., García-de-León F.J., Balart E.F. 2015b. Switch between morphospecies of *Pocillopora* corals. Am. Nat. 186:434–440.

Pembleton L.W., Cogan N.O.I., Forster J.W. 2013. StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. Mol. Ecol. Resour. 13:946–952.

Philander S.G., Fedorov A.V. 2003. Role of tropics in changing the response to Milankovich forcing some three million years ago. Paleoceanography. 18:1045.

Pinzón J.H., LaJeunesse T.C. 2011. Species delimitation of common reef corals in the genus *Pocillopora* using nucleotide sequence phylogenies, population genetics and symbiosis ecology. Mol. Ecol. 20:311–325.

Pinzón J.H., Sampayo E., Cox E., Chauka L.J., Chen C.A., Voolstra C.R., LaJeunesse T.C. 2013. Blind to morphology: genetics identifies several widespread ecologically common species and few endemics among Indo-Pacific cauliflower corals (*Pocillopora*, Scleractinia). J. Biogeogr. 40:1595–1608.

Price M.N., Dehal P.S., Arkin A.P. 2009. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. Mol. Biol. Evol. 26:1641–1650.

1010 Pritchard J.K., Stephens M., Donnelly P. 2000. Inference of population structure using multilocus
1011 genotype data. Genetics. 155:945–959.
1012 Quattrini A.M., Faircloth B.C., Dueñas L.F., Bridge T.C.L., Brugler M.R., Calixto-Botía I.F.,
1013 DeLeo D.M., Forêt S., Herrera S., Lee S.M.Y., Miller D.J., Prada C., Rádis-Baptista G.,
1014 Ramírez-Portilla C., Sánchez J.A., Rodríguez E., McFadden C.S. 2018. Universal target-
1015 enrichment baits for anthozoan (Cnidaria) phylogenomics: new approaches to long-standing
1016 problems. Mol. Ecol. Resour. 18:281–295.
1017 R Core Team. 2021. R: a language and environment for statistical computing. R Foundation for
1018 Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/: .
1019 Rabbani G., Huang D., Wainwright B.J. 2021. The mycobiome of *Pocillopora acuta* in Singapore.
1020 Coral Reefs. 40:1419–1427.
1021 Richards Z.T., van Oppen M.J.H., Wallace C.C., Willis B.L., Miller D.J. 2008. Some rare Indo-
1022 Pacific coral species are probable hybrids. PLoS ONE. 3:3240.
1023 Ros M., Suggett D.J., Edmondson J., Haydon T., Hughes D.J., Kim M., Guagliardo P., Bougoure J.,
1024 Pernice M., Raina J.-B., Camp E.F. 2021. Symbiont shuffling across environmental
1025 gradients aligns with changes in carbon uptake and translocation in the reef-building coral
1026 *Pocillopora acuta*. Coral Reefs. 40:595–607.
1027 Rueden C.T., Schindelin J., Hiner M.C., DeZonia B.E., Walter A.E., Arena E.T., Eliceiri K.W.
1028 2017. ImageJ2: ImageJ for the next generation of scientific image data. BMC
1029 Bioinformatics. 18:529.
1030 Schmidt-Roach S., Lundgren P., Miller K.J., Gerlach G., Noreen A.M.E., Andreakis N. 2012.
1031 Assessing hidden species diversity in the coral *Pocillopora damicornis* from Eastern
1032 Australia. Coral Reefs. 32:161–172.
1033 Schmidt-Roach S., Miller K.J., Lundgren P., Andreakis N. 2014. With eyes wide open: a revision of
1034 species within and closely related to the *Pocillopora damicornis* species complex
1035 (Scleractinia; Pocilloporidae) using morphology and genetics. Zool. J. Linn. Soc. 170:1–33.

1036 Shearer T.L., Van Oppen M.J.H., Romano S.L., Wörheide G. 2002. Slow mitochondrial DNA
1037 sequence evolution in the Anthozoa (Cnidaria). Mol. Ecol. 11:2475–2487.
1038 Shimpi G.G., Patel N.P., Haldar S. 2019. Molecular species delimitation of reef-building coral
1039 genera, *Porites* and *Turbinaria* (Anthozoa: Scleractinia), from the intertidal fringing reefs of
1040 Gulf of Kutch, India reveals unrecognized diversity. Syst. Biodivers. 17:541–557.
1041 Sier C.J.S., Olive P.J.W. 1994. Reproduction and reproductive variability in the coral *Pocillopora*
1042 *verrucosa* from the Republic of Maldives. Mar. Biol. 118:713–722.
1043 Simpson C., Kiessling W., Mewis H., Baron□Szabo R.C., Müller J. 2011. Evolutionary
1044 diversification of reef corals: a comparison of the molecular and fossil records. Evolution.
1045 65:3274–3284.
1046 Stat M., Gates R.D. 2010. Clade D *Symbiodinium* in scleractinian corals: a “nugget” of hope, a
1047 selfish opportunist, an ominous sign, or all of the above? J. Mar. Biol. 2011:730715.
1048 Stefani F., Benzoni F., Yang S.-Y., Pichon M., Galli P., Chen C.A. 2011. Comparison of
1049 morphological and genetic analyses reveals cryptic divergence and morphological plasticity
1050 in *Stylophora* (Cnidaria, Scleractinia). Coral Reefs. 30:1033–1049.
1051 Stoltz M., Baeumer B., Bouckaert R., Fox C., Hiscott G., Bryant D. 2021. Bayesian inference of
1052 species trees using diffusion models. Syst. Biol. 70:145–161.
1053 Terraneo T.I., Benzoni F., Arrigoni R., Baird A.H., Mariappan K.G., Forsman Z.H., Wooster M.K.,
1054 Bouwmeester J., Marshell A., Berumen M.L. 2021. Phylogenomics of *Porites* from the
1055 Arabian Peninsula. Mol. Phylogenet. Evol. 161:107173.
1056 Terraneo T.I., Benzoni F., Arrigoni R., Berumen M.L. 2016. Species delimitation in the coral genus
1057 *Goniopora* (Scleractinia, Poritidae) from the Saudi Arabian Red Sea. Mol. Phylogenet.
1058 Evol. 102:278–294.
1059 Todd P.A. 2008. Morphological plasticity in scleractinian corals. Biol. Rev. 83:315–337.

1060 Toonen R.J., Puritz J.B., Forsman Z.H., Whitney J.L., Fernandez-Silva I., Andrews K.R., Bird C.E.
1061 2013. ezRAD: a simplified method for genomic genotyping in non-model organisms. PeerJ.
1062 1:203.
1063 Torda G., Schmidt-Roach S., Peplow L.M., Lundgren P., van Oppen M.J.H. 2013. A rapid genetic
1064 assay for the identification of the most common *Pocillopora damicornis* genetic lineages on
1065 the great barrier reef. PLoS ONE. 8:e58447.
1066 Torres A.F., Forsman Z.H., Ravago-Gotanco R. 2020. Shifts in coral clonality along a gradient of
1067 disturbance: insights on reproduction and dispersal of *Pocillopora acuta*. Mar. Biol.
1068 167:161.
1069 van Oppen M.J.H., McDonald B.J., Willis B., Miller D.J. 2001. The evolutionary history of the
1070 coral genus *Acropora* (Scleractinia, Cnidaria) based on a mitochondrial and a nuclear
1071 marker: reticulation, incomplete lineage sorting, or morphological convergence? Mol. Biol.
1072 Evol. 18:1315–1329.
1073 van Oppen M.J.H., Willis B.L., Miller D.J. 1999. Atypically low rate of cytochrome b evolution in
1074 the scleractinian coral genus *Acropora*. Proc. R. Soc. B Biol. Sci. 266:179–183.
1075 van Oppen M.J.H., Willis B.L., Vugt H.V., Miller D.J. 2000. Examination of species boundaries in
1076 the *Acropora cervicornis* group (Scleractinia, Cnidaria) using nuclear DNA sequence
1077 analyses. Mol. Ecol. 9:1363–1373.
1078 Vaughan T.W., Wells J.W. 1943. Revision of the suborders, families, and genera of the
1079 Scleractinia. Boulder: Geological Society of America.
1080 Veron J.E.N. 2000. Corals of the world. Australia: Australian Institute of Marine Science.
1081 Vidal-Dupiol J., Chaparro C., Pratlong M., Pontaroti P., Grunau C., Mitta G. 2019. Sequencing, *de*
1082 *novo* assembly and annotation of the genome of the scleractinian coral, *Pocillopora acuta*.
1083 BioRxiv.
1084 Vollmer S.V., Palumbi S.R. 2004. Testing the utility of internally transcribed spacer sequences in
1085 coral phylogenetics. Mol. Ecol. 13:2763–2772.

1086 Warner P.A., van Oppen M.J.H., Willis B.L. 2015. Unexpected cryptic species diversity in the
1087 widespread coral *Seriatopora hystrix* masks spatial-genetic patterns of connectivity. Mol.
1088 Ecol. 24:2993–3008.
1089 Warnes G.R., Bolker B., Bonebakker L., Gentleman R., Huber W., Liaw A., Lumley T., Maechler
1090 M., Magnusson A., Moeller S., Schwartz M., Venables B. 2020. gplots: various R
1091 programming tools for plotting data. R package version 3.1.1.
1092 Weir B.S., Cockerham C.C. 1984. Estimating F-statistics for the analysis of population structure.
1093 Evolution. 38:1358–1370.
1094 Wells J.W. 1956. Scleractinia. Treatise on Invertebrate Paleontology, Part F. Boulder: Geological
1095 Society of America. p. F328–F444.
1096 Wepfer P.H., Nakajima Y., Sutthacheep M., Radice V.Z., Richards Z., Ang P., Terraneo T., Sudek
1097 M., Fujimura A., Toonen R.J., Mikheyev A.S., Economo E.P., Mitarai S. 2020.
1098 Evolutionary biogeography of the reef-building coral genus *Galaxea* across the Indo-Pacific
1099 ocean. Mol. Phylogenet. Evol. 151:106905.
1100 Wham D.C., Ning G., LaJeunesse T.C. 2017. *Symbiodinium glynnii* sp. nov., a species of stress-
1101 tolerant symbiotic dinoflagellates from pocilloporid and montiporid corals in the Pacific
1102 Ocean. Phycologia. 56:396–409.
1103 Willis B.L., van Oppen M.J.H., Miller D.J., Vollmer S.V., Ayre D.J. 2006. The role of
1104 hybridization in the evolution of reef corals. Annu. Rev. Ecol. Evol. Syst. 37:489–517.
1105

1106 **APPENDICES**

1107 **Appendix 1** Sampling.

1108 **Appendix 2** Molecular Analyses.

1109 **Appendix 3** Morphological Analyses.

1110 **Appendix 4** Symbiodiniaceae Analyses.

1111 **Appendix 5** Illustration of the *Pocillopora* Genomic Species Hypotheses.