**Marine Biodiversity**
February 2022, Volume 52 Issue 1 Pages 5 (8p.)
https://doi.org/10.1007/s12526-021-01246-0
https://archimer.ifremer.fr/doc/00742/85391/



# Exploring the *Pocillopora* cryptic diversity: a new genetic lineage in the Western Indian Ocean or remnants from an ancient one?

Oury Nicolas [1, 2, *], Gélin Pauline [1], Rajaonarivelo Mamy [3], Magalon Hélène [1, 2]

[1] UMR ENTROPIE (Université de La Réunion, IRD, IFREMER, Université de Nouvelle-Calédonie, CNRS), Université de La Réunion, Saint-Denis, La Réunion, France
[2] Laboratoire Cogitamus, Paris, France
[3] CNRO (Centre National de Recherches Océanographiques), MESupReS, Hell-Ville, Nosy Be, Madagascar

* Corresponding author : Nicolas Oury, email address : nicolasoury@hotmail.fr

**Abstract :**

Cryptic species and lineages have been widely reported during the last decades, particularly in the marine realm. Misidentifications and ignoring species complexes imply many consequences, notably biasing biodiversity and connectivity assessments, which in turn mislead our understanding of ecosystems and impact the effective design and management of conservation plans. Focusing on the Indo-Pacific coral genus Pocillopora, playing key roles in reef ecosystems as one of the main bio-constructors, we report the first Pocillopora PSH16 (ORF53; sensu Gélin et al. 2017) colonies (N = 19) in the Western Indian Ocean (Nosy Tanikely, Madagascar), 6000 km further from its current distribution. Colonies were identified according to their mitochondrial open reading frame (ORF) haplotype and Bayesian assignment tests based on 13-microsatellite genotypes. Additionally, we performed genetic structure and diversity analyses with sympatric colonies from other Pocillopora species and Pocillopora PSH16 colonies from the tropical southwestern Pacific, revealing (1) a weak clonal richness, (2) a weak genetic diversity and (3) a relative isolation for the newly reported PSH16 colonies. These colonies thus represent either a new, distinct and uncommon, genetic lineage or isolated remnants of a wider one. In any case, unless specific management measures are implemented, their long-term maintenance seems compromised due to restricted gene flow within a restricted pool of genes.

**Keywords** : Genetic diversity, Microsatellite, Mitochondrial open reading frame, Scleractinian, Species distribution, Species hypothesis

# Introduction

Cryptic species are often defined as two or more distinct species that were classified as a single one due to their morphological similarity. Within the last decades, the development of sequencing technologies, coupled with the advances in bioinformatics, has led to the democratisation of genetic tools and analyses, and to the growing discovery of cryptic species and lineages (Pfenninger and Schwenk 2007). Thus, more and more species defined by morphological characteristics (i.e. morphospecies), and previously thought to be widely distributed, were demonstrated as complexes of different (cryptic) species, either each with distinct restricted distribution range or found in sympatry. As an illustration, many circumglobal or cosmopolitan red algae were in fact complexes of species with narrower distribution each, with the exception of human transported species (Díaz-Tapia et al. 2018; review in Hu et al. 2016). Such misidentifications have many consequences, the most obvious of which are incorrect biodiversity assessments and biased overviews of connectivity, which in turn impact the effective design and management of conservation plans.

This is particularly relevant for coral reefs, which are facing many threats (review in Burke et al. 2017), and whose conservation is a growing and pressing topic (see Abelson 2020). Despite the huge biodiversity and ecological services provided by coral reefs (Moberg and Folke 1999), many aspects of these ecosystems remain unknown, and many species, cryptic or not, remain undiscovered (Victor 2015).

The coral genus *Pocillopora* represents a key component of coral reefs 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, this genus remains a source of misunderstanding and a challenge for taxonomists. Indeed, recent taxonomic (Schmidt-Roach et al. 2014) and genetic (e.g. Pinzón et al. 2013; Gélin et al. 2017) studies identified several cryptic species and lineages within this genus. As an illustration, using species delimitation methods based on molecular markers, Gélin et al. (2017) defined within the *Pocillopora* genus 16 Primary Species Hypotheses (PSHs *sensu* Pante et al. 2015), and a few of these PSHs were partitioned into several Secondary Species Hypotheses (SSHs *sensu* Pante et al. 2015). Furthermore, some PSHs, and even more some SSHs, were found to be geographically restricted, while the corresponding morphospecies were thought to be widely distributed over the whole distribution range of the genus (Veron 2000). Thus, all *Pocillopora* species previously described with an Indo-Pacific distribution in Veron (2000) were complexes of cryptic species, each restricted to the ocean basin: *P. damicornis* (Schmidt-Roach et al. 2014;

Gélin et al. 2017), *P. eydouxi*/*meandrina* (Gélin et al. 2018a; Oury et al. 2021) and *P. verrucosa* (Oury et al. 2021).

Additionally, some PSHs were attributed to new or recently described species. Among them, *Pocillopora* PSH16 regroups the mitochondrial Open Reading Frame (ORF) haplotypes 53, 54 and 55 (*sensu* Gélin et al. 2017), corresponding to types 3d and 3f (*sensu* Pinzón et al. 2013) and to *P. damicornis* type $\chi$ (now *P. bairdi sensu* Schmidt-Roach et al. 2014), respectively. Up to now, PSH16 was only reported in the Pacific Ocean (more precisely in Taiwan, Palau, eastern Australia, New Caledonia, Tonga Islands and Moorea; Fig. 1), excepted one individual of type 3d (ORF53) in the Andaman Sea (Pinzón et al. 2013). Here, we report the first *Pocillopora* PSH16 (ORF53) in the western Indian Ocean (Madagascar), 6,000 km further from its current distribution, possibly representing a new *Pocillopora* genetic lineage, and possibly a new species.

# Material and methods

## Sample collection and morphotype identification

Aiming at exploring the *Pocillopora* genus diversity, and *in fine* studying its population genetic connectivity, ~ 9,000 *Pocillopora* colonies were sampled (branch tip + photograph), independently of their *corallum* macromorphology (a non-discriminant character in this genus), from March 2001 to October 2016, in three marine provinces: the Western Indian Ocean (WIO), the Tropical Southwestern Pacific (TSP) and the South-East Polynesia (SEP), extended over six ecoregions (Spalding et al. 2007), 16 localities (Online Resource 1) and over hundred sampling sites. Different habitats (reef slope, fringing reef, flat reef or lagoon) were sampled, at various depths (from sea surface to 30 m depth), to maximise colonies genetic diversity. Samples were isolated into a numbered zip-lock bag on the field, then fixed in 90% ethanol at laboratory and stored at room temperature.

Additionally, examining the underwater photograph, each colony was attributed a morphotype (or several when morphology was unclear), defined 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)]. Morphotype identification was verified by sending a subset of photographs to three coral specialists (F. Benzoni, G. Faure and D. Obura).

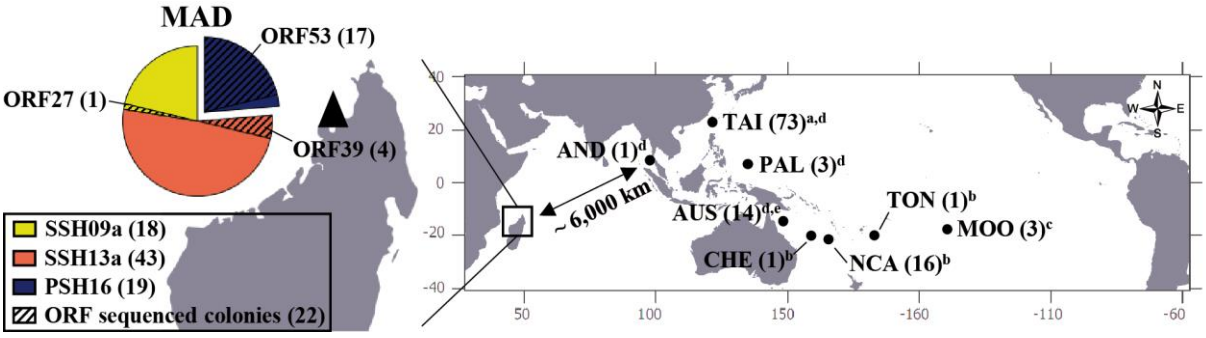


**Fig. 1** Records of *Pocillopora* PSH16. For each locality, the number of previously recorded colonies is indicated in parentheses, with corresponding references (a: De Palmas et al. 2018; b: Gélin et al. 2017; c: Johnston et al. 2021; d: Pinzón et al. 2013; e: Schmidt-Roach et al. 2014; right panel). Records from Nosy Tanikely (Madagascar; MAD; symbolised by the black triangle on the left panel; $N = 19$) were from this study. The *Pocillopora* species repartition in this site (over 80 sampled colonies) is detailed, with the number of colonies from each ORF haplotype sequenced (hatched parts).
AUS: Australia (Lizard Island); AND: Andaman Sea; CHE: Chesterfield Islands; MOO: Moorea; NCA: New Caledonia (Grande Terre and Loyalty Islands); PAL: Palau; TAI: Taiwan and TON: Tonga Islands.

**DNA extraction, genotyping, sequencing and species identification**

From the sampled colonies, total genomic DNA was extracted using DNeasy Blood & Tissue kit (Qiagen™), following the manufacturer protocol. All colonies were then genotyped with 13 microsatellite loci, as in Gélin et al. (2017). PCR products were analysed using an ABI 3730XL DNA Analyzer (Applied Biosystems) at the Plateforme Gentyane (INRAE, Clermont-Ferrand, France) and allelic sizes were determined with GENEMAPPER 4.0 (Applied Biosystems) using an internal size standard (Genescan LIZ-500, Applied Biosystems). Loci showing ambiguous peak profiles (e.g. faint peaks or more than two peaks) were processed again in simplex and, if remaining ambiguous, designated as missing data.

Colonies were then identified *a posteriori* of sampling and *a priori* of analyses using assignment tests performed with STRUCTURE 2.3.4 (Pritchard et al. 2000), by compiling all the genotypes with the 975 ones from Gélin et al. (2017), corresponding to colonies from various PSHs/SSHs already identified. Five iterations of STRUCTURE were run at $K = 12$ (this value was found to retrieve the main PSHs/SSHs in Gélin et al. 2017), with the same parameters as in Gélin et al. (2017).

Once the clusters retrieved, in order to identify them and maximise the ORF diversity explored, we sequenced a subset of the colonies within each cluster and each locality, presenting various general *corallum* macromorphologies (i.e. belonging to various morphotypes so that colonies both presenting similar and different morphotypes were selected). The FATP6.1 and RORF primers (Flot and Tillier 2007) were used, as in Gélin et al. (2017), and amplicons were

sent to GenoScreen (Lille, France) for sequencing in both directions on an ABI 3730XL DNA Analyzer (Applied Biosystems). Sequences were checked and edited using GENEIOUS 8.0 (Kearse et al. 2012). Then they were aligned with the sequences of the 55 reference ORF haplotypes from Gélin et al. (2017), all available in GenBank. Alignment was performed with MAFFT (Katoh et al. 2005), and sequences were trimmed to 842 bp (the length of the reference sequences).

Colonies belonging to PSH16 and having the ORF53 haplotype were found in Madagascar, 6,000 km further from its current distribution (see Results). To further explore the genetic structure and diversity of this PSH, we focused the rest of the analyses on it and its sympatric colonies from the WIO.

## Genetic analyses

*Allelic and clonal diversities*

From there, a population was considered as all sampled colonies from a single species sampled in the same site, at the same date (usually during a single dive, and thus in a limited depth-range). The occurrence of identical multi-locus genotypes (MLGs) within each population was assessed with GENCLONE 2.0 (Arnaud-Haond and Belkhir 2007). The probability of obtaining the same MLG twice or more from distinct random reproductive events was further estimated using $P_{SEX}(F_{IS})$ (Arnaud-Haond et al. 2007). The clonal richness $R$ (Dorken and Eckert 2001) was then calculated for each population, using the formula: $R = \frac{(N_{MLG}-1)}{(N-1)}$, with $N$, the number of colonies and $N_{MLG}$, the number of distinct MLGs.

Afterwards, only one representative for each MLG was kept. PSH16 colonies from the TSP ($N$ = 18 colonies from Gélin et al. 2017; Fig. 1), genotyped with the same 13 microsatellite loci as herein, were added to the dataset from this study, to assess the Indo-Pacific structure. Null allele frequencies and other potential technical biases were estimated with MICROCHECKER 2.2.3 (van Oosterhout et al. 2004), within each population (colonies from the Pacific were considered as a single population). Then, diversity indices [i.e. *Na* and *Np*, the mean numbers of alleles and private alleles per locus, *AR*, the allelic richness, *Ho* and *He*, the observed and expected heterozygosities, and $F_{IS}$, the inbreeding coefficient (Wright 1931)] and percentage of missing data (*%NA*) were estimated for each population and over the 13 loci, using FSTAT 2.9.3 (Goudet 2001). Linkage disequilibrium and departures from Hardy-Weinberg equilibrium were tested using GENEPOP 4.7.0 (Raymond and Rousset 1995; Rousset 2008).

*Genetic structure*

To assess the genetic structure among colonies, we used and compared the results of assignment tests [STRUCTURE (Pritchard et al. 2000) and DAPC (Jombart et al. 2010)], Minimum Spanning Trees (MST) and differentiation indices. STRUCTURE was run with the admixture model, assuming correlated allele frequencies. Five iterations of $10^6$ MCMC generations after an initial burn-in of $10^5$ generations were run for each *K*, varying from $K = 2$ to $K = 5$, and results were combined and visualised with CLUMPAK (Kopelman et al. 2015). The discriminant analysis of principal components (DAPC) was performed with the package '*adegenet'* (Jombart 2008) from the software R 3.1.1 (R Core Team 2021), the MST based on the shared allele distance between colonies was built with EDENETWORKS 2.18 (Kivelä et al. 2015), and differentiation indices [$F_{ST}$ (Weir and Cockerham 1984) and *Dest* (Jost 2008)] were estimated for each population pair with the R package *'diveRsity'* (Keenan et al. 2013). Finally, an unweighted pair group method with arithmetic mean (UPGMA) cluster dendrogram based on Nei (1972)'s genetic distance between populations was built with MEGA 7 (Kumar et al. 2016).

# Results and Discussion

## Species identification

Over all localities, ~ 9,000 colonies were sampled, genotyped and assigned to 12 clusters, as in Gélin et al. (2017). Based on these assignments, sampling localities and morphotypes of the colonies, a subset of 1,003 colonies were successfully sequenced for the ORF. Among them, no new haplotype was identified compared to those previously found in the phylogenetic study (Gélin et al. 2017; Online Resource 1). In particular, two of the three PSH16 ORF haplotypes (ORF53 and 54) were found in the TSP (previously reported in Gélin et al. 2017), and for the first time, one (ORF53; $N = 17$) was found in the WIO, more precisely in the National Marine Park of Nosy Tanikely, Nosy Be (northwestern Madagascar; 13.48702°S; 48.23548°E; 10-15 m depth; Fig. 1 & 2; Online Resource 2), the closest previous record being in the Andaman Sea (i.e. 6,000 km eastward; Pinzón et al. 2013; Fig. 1). We thus focused on colonies from this site for further analyses.

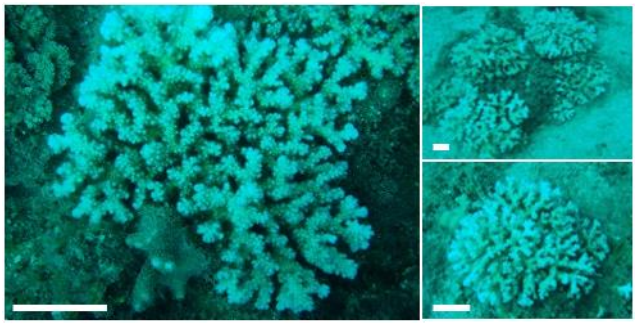

**Fig. 2** *Pocillopora* PSH16 colonies in Nosy Tanikely (scale: ~ 5 cm).

A total of 80 colonies were sampled in Nosy Tanikely, of which 18 were assigned to SSH09a corresponding to *P. eydouxi/meandrina* morphotype, 43 to SSH13a corresponding to *P. verrucosa* morphotype and 19 to PSH16 whose colonies morphotype is characterised by an horizontal growth with branches separated by large spaces and highly ramified (Table 1; Fig. 1; see Online Resource 2 for comparative photos of colonies from these three PSHs/SSHs). From them, a subset of 22 colonies were successfully sequenced for the ORF, and all colonies whose ORF has been sequenced were assigned to the PSH corresponding to the ORF haplotype. Thus, three ORF haplotypes were found: ORF27 (PSH09; $N = 1$), ORF39 (PSH13; $N = 4$) and ORF53 (PSH16; $N = 17$; Table 1; Fig. 1).

**Allelic and clonal diversities**

Three populations were thus considered in Nosy Tanikely, with respect to the three PSHs/SSHs identified (i.e. SSH09a, SSH13a and PSH16; Table 1; Fig. 1). Among them, two MLGs were found repeated more than once, both within PSH16 colonies. One was repeated twice, the second 14 times. For both MLGs, $P_{SEX}(F_{IS})$ was very low ($< 10^{-10}$), indicating that the colonies came from a single sexual reproduction event and belonged to the same genet. A clonal richness $R$ of 0.22 was thus calculated for PSH16 ($N_{MLG} = 5$; Table 1).

Missing data represented 5.2% of the Nosy Tanikely dataset, but 6.9% when considering the fourth population with the 18 Pacific PSH16 colonies from Gélin et al. (2017). No null allele nor other potential scoring error was detected. Moreover, for each population, no linkage disequilibrium was found. Among the four populations, the mean numbers of alleles and private alleles per locus (*Na* and *Np*; ± SE) were the lowest in PSH16 colonies from Madagascar, varying from 4.2 ± 0.3 to 7.2 ± 0.8, and from 0.7 ± 0.2 to 1.8 ± 0.4, respectively (Table 1). However, the allelic richness (*AR*; based on 10 alleles) and heterozygosis rates (*Ho* and *He*), varying from 3.61 ± 0.41 to 4.15 ± 0.34, from 0.386 ± 0.071 to 0.492 ± 0.057, and from 0.546 ± 0.077 to 0.719 ± 0.054, respectively, were the highest in this population (Table 1). All

four populations showed a significant heterozygote deficiency (0.242 ± 0.081* < $F_{IS}$ < 0.377 ± 0.076***; Table 1).

**Genetic structure**

At *K* =2, SSH09a and SSH13a colonies were each assigned to a specific cluster by STRUCTURE, while PSH16 colonies were almost all assigned to both clusters with similar probabilities. At *K* = 3, colonies assignment reflected the species. Only one PSH16 colony from the WIO was assigned to SSH13a cluster, and one SSH13a colony was admixed between SSH13a and PSH16 clusters (Fig. 3a). Yet, for these two colonies, the ORF corresponded to the PSH/SSH identification and not to the cluster they were assigned to (i.e. the first have the PSH16 ORF53 haplotype, and the second the SSH13a ORF39 haplotype). Then, at *K* = 4, PSH16 colonies from the WIO were assigned to a fourth cluster, with three PSH16 colonies from the TSP. Three PSH16 colonies from the TSP were admixed between both PSH16 clusters. At *K* = 5, all SSH13a colonies were assigned to two clusters with similar probabilities, suggesting no further genetic partitioning (Fig. 3a). DAPC grouped the colonies according to the three species and was thus congruent with STRUCTURE results at *K* = 3. PSH16 colonies from the WIO and the TSP were not separated (Fig. 3b). The MST clearly distinguished SSH09a and SSH13a colonies, as well as almost all PSH16 colonies from the TSP, but PSH16 colonies from the WIO were grouped either with SSH13a or with those of the TSP (Fig. 3c). Both differentiation indices (i.e. $F_{ST}$ and *Dest*) gave similar results, but *Dest* estimates were higher (0.072** < $F_{ST}$ < 0.282***; 0.137*** < *Dest* < 0.531***; Table 2). According to these indices (Table 2) and the UPGMA tree (Online Resource 3), SSH09a was the most differentiated population ($F_{ST}$ > 0.220***; *Dest* > 0.390***), then SSH13a, the latter being more differentiated from PSH16 (TSP) than from PSH16 (WIO).

252 **Table 1** *Pocillopora* populations genetic diversity

| Population/ Species | *N* | *ORF* | $N_{MLG}$ | *R* | *%NA* | *Na* | *Np* | *AR* (10) | *Ho* | *He* | $F_{IS}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *Nosy Tanikely (Madagascar; 13.48702°S, 48.23548°E)* | | | | | | | | | | | |
| SSH09a | 18 | ORF27 (1) | 18 | 1.00 | 11.5% | 5.5±0.6 | 1.4±0.4 | 3.61±0.41 | 0.386±0.071 | 0.546±0.077 | 0.242±0.081* |
| SSH13a | 43 | ORF39 (4) | 43 | 1.00 | 7.2% | 7.2±0.8 | 1.8±0.4 | 3.87±0.29 | 0.477±0.056 | 0.672±0.035 | 0.285±0.084** |
| PSH16 (WIO) | 19 | ORF53 (17) | 5 | 0.22 | 0.0% | 4.2±0.3 | 0.7±0.2 | 4.15±0.34 | 0.492±0.089 | 0.719±0.054 | 0.328±0.107* |
| *Tropical Southwestern Pacific (Gélin et al. 2017)* | | | | | | | | | | | |
| PSH16 (TSP) | 18 | ORF53 (14)<br>ORF54 (4) | 18 | 1.00 | 18.8% | 5.5±0.5 | 0.6±0.2 | 3.97±0.28 | 0.432±0.057 | 0.697±0.037 | 0.377±0.076*** |
| **TOTAL** | **98** | **-** | **84** | **0.86** | **6.9%** | **10.7±1.0** | **4.5±0.6** | **4.60±0.28** | **-** | **-** | **-** |

253 *N*: number of colonies; ORF: Open Reading Frame haplotype (occurrence in parentheses); $N_{MLG}$: number of multilocus genotypes (MLGs); *R*:
254 clonal richness (Dorken and Eckert 2001), *%NA*: percentage of missing data; *Na* and *Np*: mean numbers (± SE) of alleles and private alleles; *AR*:
255 mean (± SE) allelic richness based on 10 alleles; *Ho* and *He*: mean (± SE) observed and expected heterozygosities and $F_{IS}$: mean (± SE) inbreeding
256 coefficient (Wright 1931; *: $0.01 < P < 0.05$; **: $0.001 < P < 0.01$; ***: $P < 0.001$). WIO: Western Indian Ocean; TSP: Tropical Southwestern
257 Pacific

258

259 **Table 2** Genetic differentiation between *Pocillopora* populations

| | *N* | SSH09a | SSH13a | PSH16 (WIO) | PSH16 (TSP) |
|---|---|---|---|---|---|
| **SSH09a** | **18** | - | 0.456*** | 0.531*** | 0.397*** |
| **SSH13a** | **43** | 0.268*** | - | 0.147*** | 0.196*** |
| **PSH16 (WIO)** | **5** | 0.282*** | 0.104*** | - | 0.137*** |
| **PSH16 (TSP)** | **18** | 0.224*** | 0.130*** | 0.072** | - |

260 $F_{ST}$ (below diagonal; Weir and Cockerham 1984) and *Dest* (above diagonal; Jost 2008) estimates. **: $0.001 < P < 0.01$; ***: $P < 0.001$. *N*: number
261 of colonies (one representative per MLGs); WIO: Western Indian Ocean; TSP: Tropical Southwestern Pacific

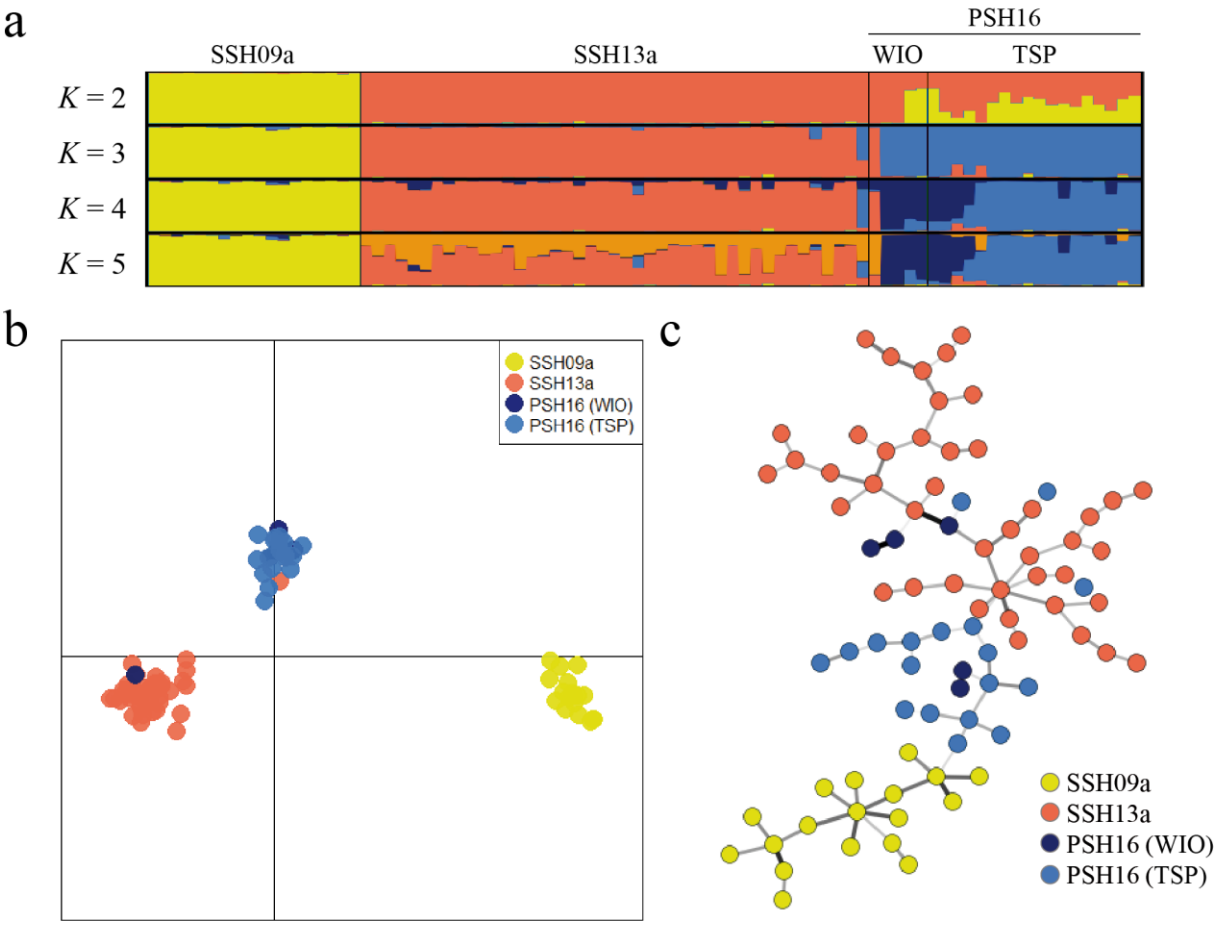


**Fig. 3** *Pocillopora* genetic structure. a: STRUCTURE plots from $K$ = 2 to $K$ = 5; b: DAPC assignments at $K$ = 3; c: Minimum Spanning Tree.
WIO: Western Indian Ocean; TSP: Tropical Southwestern Pacific

These results mainly suggest that the PSH16 colonies from the WIO and those from the TSP belonged to the same ancestral genetic group, probably isolated from an older ancestral one including PSH13. This is consistent with a previous phylogenetic study on the *Pocillopora* genus (Gélin et al. 2017), where PSH16 colonies from the TSP were found relatively genetically close to PSH13 colonies from the TSP. However, STRUCTURE results at $K$ = 4, $F_{ST}$ and *Dest* estimates, and Nei's distance between both PSH16 populations also indicate that PSH16 colonies from both ocean basins appear genetically different, and probably constitute different genetic lineages or different species. This observation is supported by the weak gene flow previously reported between both sides of the Indo-Pacific in this genus (Gélin et al. 2018a, b; Oury et al. 2021) questioning the existence of Indo-Pacific *Pocillopora* species.

Moreover, as the only PSH16 colonies reported within a radius of 6,000 km despite the thousands of *Pocillopora* colonies sampled in the WIO, this PSH16 lineage from the WIO seems uncommon and isolated, being mainly maintained through asexual reproduction, or

sexual reproduction from a very restricted pool of genes, as suggested by the weak clonal richness. However, as microhabitat data are lacking (the sampling initially aimed at studying population genetic connectivity), we cannot be sure that Nosy Tanikely does not represent a unique, only sampled once, habitat, specific to PSH16. In that case, regarding the hundred sites sampled, this specific habitat should be rare, and so PSH16 colonies.

Higher allelic richness and heterozygosis rates were found in this population compared to others, which does not reflect population isolation. Occasional hybridisations with genetically close lineages could explain these higher diversity indices. Interspecific hybridisations were already suggested within *Pocillopora* corals. Indeed, Combosch and Vollmer (2015) reported one-way introgressive hybridisation among tropical eastern Pacific *Pocillopora* morphospecies based on RAD-Seq. However, ITS2 heterozygous individuals, considered to be potential hybrids, were pooled in genomic libraries, thus possibly confusing hybridisation signals. More recently, trying to resolve phylogenetic relationships among seven *Pocillopora* species using RAD-Seq, Johnston et al. (2017) suggested hybridisation between the two recently derived sister species included in their study: *P. damicornis* and *P. acuta*. In our study, occasional reproductions with sympatric SSH13a colonies could be possible, given the relatively small genetic distance with PSH16 colonies reported by the MST, and $F_{ST}$, *Dest* and Nei's distances. In this case, the SSH13a colony assigned to the PSH16 cluster and the PSH16 (WIO) colony assigned to the SSH13a cluster by STRUCTURE and DAPC could represent colonies derived from hybridisation between both species.

Thus, in addition to the first report of PSH16 colonies in the WIO, 6,000 km further from the current PSH16 distribution, this study also reports a possibly new *Pocillopora* genetic lineage, or a remnant population from a wider lineage. Further investigations integrating morphological evidences (e.g., as in Stefani et al. 2011) are needed to clearly resolve the status of this lineage. Whatever these colonies are, they appear uncommon, isolated and processing genetic homogenisation, through hybridisation with closely related *Pocillopora* lineages. They thus process speciation, through a progressive differentiation from other lineages, and will probably end up as a new species, if they persist long enough. Indeed, the long-term maintenance of these PSH16 colonies appears compromised, especially if as rare as thought, and if no specific management measure is implemented. Colonies from Nosy Tanikely are already subject to conservation measures, as part of the National Marine Park of Nosy Tanikely. The report of this possibly new *Pocillopora* lineage should support the long-term implementation of the marine park protection measures and encourage managers to implement

adapted measures, especially in the context of Madagascar's willing to develop rare earth mining on Grande Terre.

**Acknowledgements** Coral sampling in Nosy Tanikely was performed during MAD oceanographic campaign (http://dx.doi.org/10.17600/16004700) on board of RV Antea (IRD). We gratefully thank GenoScreen (Lille, France) and the Plateforme Gentyane of the Institut National de Recherche pour l'Agriculture, l'Alimentation et l'Environnement (INRAE, Clermont-Ferrand, France) for sequencing and technical support. NO and PG were financially supported by PhD contracts from the Doctoral School of Reunion Island University and the LabEx CORAIL, respectively. We thank the reviewers for their comments to improve the manuscript.

## Declarations

**Funding:** None.

**Conflict of Interest:** The authors declare that they have no conflict of interest.

**Ethical approval:** All applicable international, national and/or institutional guidelines for animal testing, animal care and use of animals were followed by the authors.

**Sampling and field studies:** All necessary permits for sampling and observational field studies have been obtained by the authors from the competent authorities (authorization n°16/1040-AE/SG/DAJC/SAG/NAV/FRANCE).

**Data availability:** ORF haplotype sequences were already available on GenBank. Microsatellite genotypes were deposited on Zenodo: http://doi.org/XXXXXX

**Author Contribution Statement:** MR and HM collected samples. NO, PG and HM did lab steps and analysed results. NO wrote the original draft and NO and HM reviewed and edited the manuscript.

## References

Abelson A (2020) Are we sacrificing the future of coral reefs on the altar of the "climate change" narrative? ICES J Mar Sci 77:40–45. https://doi.org/10.1093/icesjms/fsz226

Arnaud-Haond S, Belkhir K (2007) GENCLONE: a computer program to analyse genotypic data, test for clonality and describe spatial clonal organization. Mol Ecol Notes 7:15–17. https://doi.org/10.1111/j.1471-8286.2006.01522.x

Arnaud-Haond S, Duarte CM, Alberto F, Serrão EA (2007) Standardizing methods to address clonality in population studies. Mol Ecol 16:5115–5139. https://doi.org/10.1111/j.1365-294X.2007.03535.x

Benzoni F, Bianchi CN, Morri C (2003) Coral communities of the northwestern Gulf of Aden (Yemen): variation in framework building related to environmental factors and biotic conditions. Coral Reefs 22:475–484. https://doi.org/10.1007/s00338-003-0342-1

Burke LM, Reytar K, Spalding M, Perry A (2017) Reefs at risk revisited: World Resources Institute

Combosch DJ, Vollmer SV (2015) Trans-Pacific RAD-Seq population genomics confirms introgressive hybridization in Eastern Pacific *Pocillopora* corals. Mol Phylogenet Evol 88:154–162. https://doi.org/10.1016/j.ympev.2015.03.022

De Palmas S, Soto D, Denis V, et al (2018) Molecular assessment of *Pocillopora verrucosa* (Scleractinia; Pocilloporidae) distribution along a depth gradient in Ludao, Taiwan. PeerJ 6:e5797. https://doi.org/10.7717/peerj.5797

Díaz-Tapia P, Maggs CA, Macaya EC, Verbruggen H (2018) Widely distributed red algae often represent hidden introductions, complexes of cryptic species or species with strong phylogeographic structure. J Phycol 54:829–839. https://doi.org/10.1111/jpy.12778

Dorken ME, Eckert CG (2001) Severely reduced sexual reproduction in northern populations of a clonal plant, *Decodon verticillatus* (Lythraceae). J Ecol 89:339–350. https://doi.org/10.1046/j.1365-2745.2001.00558.x

Flot J-F, Tillier S (2007) The mitochondrial genome of *Pocillopora* (Cnidaria: Scleractinia) contains two variable regions: The putative D-loop and a novel ORF of unknown function. Gene 401:80–87. https://doi.org/10.1016/j.gene.2007.07.006

Gélin P, Fauvelot C, Bigot L, et al (2018a) From population connectivity to the art of striping Russian dolls: the lessons from *Pocillopora* corals. Ecol Evol 8:1411–1426. https://doi.org/10.1002/ece3.3747

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. https://doi.org/10.1007/s00227-018-3428-6

Gélin P, Postaire B, Fauvelot C, Magalon H (2017) 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. https://doi.org/10.1016/j.ympev.2017.01.018

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices version 2.9.3.2, updated from Goudet 1995. Version 2.9.3.2

Hu Z-M, Duan D-L, Lopez-Bautista J (2016) Seaweed phylogeography from 1994 to 2014: An overview. In: Hu Z-M, Fraser C (eds) Seaweed Phylogeography: Adaptation and Evolution of Seaweeds under Environmental Change. Springer Netherlands, Dordrecht, pp 3–22

Johnston EC, Forsman ZH, Flot J-F, et al (2017) A genomic glance through the fog of plasticity and diversification in *Pocillopora*. Sci Rep 7:5991. https://doi.org/10.1038/s41598-017-06085-3

Johnston EC, Wyatt AS, Leichter JJ, Burgess SC (2021) Niche differences in co-occurring cryptic coral species (*Pocillopora* spp.). Coral Reefs. https://doi.org/10.1007/s00338-021-02107-9

Jombart T (2008) adegenet: a R package for the multivariate analysis of genetic markers. Bioinformatics 24:1403–1405

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

Jost LOU (2008) $G_{ST}$ and its relatives do not measure differentiation. Mol Ecol 17:4015–4026. https://doi.org/10.1111/j.1365-294X.2008.03887.x

Katoh K, Kuma K, Toh H, Miyata T (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. Nucleic Acids Res 33:511–518

Kearse M, Moir R, Wilson A, et al (2012) Geneious Basic: An integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28:1647–1649. https://doi.org/10.1093/bioinformatics/bts199

Keenan K, McGinnity P, Cross TF, et al (2013) diveRsity : an R package for the estimation and exploration of population genetics parameters and their associated errors. Methods Ecol Evol 4:782–788. https://doi.org/10.1111/2041-210X.12067

Kivelä M, Arnaud-Haond S, Saramäki J (2015) EDENetworks: a user-friendly software to build and analyse networks in biogeography, ecology and population genetics. Mol Ecol Resour 15:117–122. https://doi.org/10.1111/1755-0998.12290
Kopelman NM, Mayzel J, Jakobsson M, et al (2015) CLUMPAK: a program for identifying clustering modes and packaging population structure inferences across *K*. Mol Ecol Resour 15:1179–1191. https://doi.org/10.1111/1755-0998.12387
Kumar S, Stecher G, Tamura K (2016) MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol 33:1870–1874. https://doi.org/10.1093/molbev/msw054
Moberg F, Folke C (1999) Ecological goods and services of coral reef ecosystems. Ecol Econ 29:215–233. https://doi.org/10.1016/S0921-8009(99)00009-9
Nei M (1972) Genetic distance between populations. Am Nat 106:283–292. https://doi.org/10.1086/282771
Oury N, Gélin P, Magalon H (2021) High connectivity within restricted distribution range in *Pocillopora* corals. J Biogeogr. https://doi.org/10.1111/jbi.14104
Pante E, Puillandre N, Viricel A, et al (2015) Species are hypotheses: avoid connectivity assessments based on pillars of sand. Mol Ecol 24:525–544. https://doi.org/10.1111/mec.13048
Pfenninger M, Schwenk K (2007) Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. BMC Evol Biol 7:121. https://doi.org/10.1186/1471-2148-7-121
Pinzón JH, Sampayo E, Cox E, et al (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. https://doi.org/10.1111/jbi.12110
Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. Genetics 155:945–959
R Core Team (2021) R: a language and environment for statistical computing. Version 4.0.4. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/
Raymond M, Rousset F (1995) GENEPOP: Population genetics software for exact tests and ecumenism. J Hered 86:248–249. https://doi.org/10.1111/j.1558-5646.1995.tb04456.x
Rousset F (2008) Genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Mol Ecol Resour 8:103–106. https://doi.org/10.1111/j.1471-8286.2007.01931.x
Schmidt-Roach S, Miller KJ, Lundgren P, Andreakis N (2014) With eyes wide open: a revision of species within and closely related to the *Pocillopora damicornis* species complex (Scleractinia; Pocilloporidae) using morphology and genetics. Zool J Linn Soc 170:1–33. https://doi.org/10.1111/zoj.12092
Spalding MD, Fox HE, Allen GR, et al (2007) Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. BioScience 57:573–583. https://doi.org/10.1641/b570707
Stefani F, Benzoni F, Yang S-Y, et al (2011) Comparison of morphological and genetic analyses reveals cryptic divergence and morphological plasticity in *Stylophora* (Cnidaria, Scleractinia). Coral Reefs 30:1033–1049. https://doi.org/10.1007/s00338-011-0797-4
van Oosterhout C, Hutchinson WF, Wills DP, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Mol Ecol Notes 4:535–538. https://doi.org/10.1111/j.1471-8286.2004.00684.x
Veron JEN (2000) Corals of the world. Australian Institute of Marine Science, Australia
Victor BC (2015) How many coral reef fish species are there? Cryptic diversity and the new molecular taxonomy. In: Ecology of fishes on coral reefs. Cambridge University Press, Cambridge, pp 76–88
Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. Evolution 38:1358–1370. https://doi.org/10.1111/j.1558-5646.1984.tb05657.x
Wright S (1931) Evolution in Mendelian populations. Genetics 16:97–159

## Electronic Supplementary Material

**Online Resource 1** Number of *Pocillopora* colonies per ORF haplotype and locality.

**Online Resource 2** *Pocillopora* colonies in Nosy Tanikely.

**Online Resource 3** UPGMA cluster dendrogram.

## Tables and Figures

**Table 1** *Pocillopora* populations genetic diversity
*N*: number of colonies; ORF: Open Reading Frame haplotype (occurrence in parentheses); $N_{MLG}$: number of multilocus genotypes (MLGs); *R*: clonal richness (Dorken and Eckert 2001), *%NA*: percentage of missing data; *Na* and *Np*: mean numbers (± SE) of alleles and private alleles; *AR*: mean (± SE) allelic richness based on 10 alleles; *Ho* and *He*: mean (± SE) observed and expected heterozygosities and $F_{IS}$: mean (± SE) inbreeding coefficient (Wright 1931; *: $0.01 < P < 0.05$; **: $0.001 < P < 0.01$; ***: $P < 0.001$). WIO: Western Indian Ocean; TSP: Tropical Southwestern Pacific

**Table 2** Genetic differentiation between *Pocillopora* populations
$F_{ST}$ (below diagonal; Weir and Cockerham 1984) and *Dest* (above diagonal; Jost 2008) estimates. **: $0.001 < P < 0.01$; ***: $P < 0.001$. *N*: number of colonies (one representative per MLGs); WIO: Western Indian Ocean; TSP: Tropical Southwestern Pacific

**Fig. 1** Records of *Pocillopora* PSH16. For each locality, the number of previously recorded colonies is indicated in parentheses, with corresponding references (a: De Palmas et al. 2018; b: Gélin et al. 2017; c: Johnston et al. 2021; d: Pinzón et al. 2013; e: Schmidt-Roach et al. 2014; right panel). Records from Nosy Tanikely (Madagascar; MAD; symbolised by the black triangle on the left panel; $N = 19$) were from this study. The *Pocillopora* species repartition in this site (over 80 sampled colonies) is detailed, with the number of colonies from each ORF haplotype sequenced (hatched parts).
AUS: Australia (Lizard Island); AND: Andaman Sea; CHE: Chesterfield Islands; MOO: Moorea; NCA: New Caledonia (Grande Terre and Loyalty Islands); PAL: Palau; TAI: Taiwan and TON: Tonga Islands.

**Fig. 2** *Pocillopora* PSH16 colonies in Nosy Tanikely (scale: ~ 5 cm).

**Fig. 3** *Pocillopora* genetic structure. a: STRUCTURE plots from $K = 2$ to $K = 5$; b: DAPC assignments at $K = 3$; c: Minimum Spanning Tree.
WIO: Western Indian Ocean; TSP: Tropical Southwestern Pacific