Systematic Biology – Supporting Information

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

Nicolas Oury, Cyril Noël, Stefano Mona, Didier Aurelle, and Hélène Magalon

Correspondence: N. 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. E-mail: nicolasoury@hotmail.fr

**Appendix 1** Sampling.

**Table of contents**

[*Sampling and Previous Identifications* 1](#Sampling)

[*References* 2](#References)

*Tables and Figures*

[Table S1 Sampling localities of *Pocillopora* colonies 3](#Table_S01)

[Table S2 Primers and PCR conditions 4](#Table_S02)

[Fig. S1 Sampling localities of *Pocillopora* colonies 6](#Fig_S01)

***Sampling and Previous Identifications***

The sampling was the same as in Gélin et al. (2017). Briefly, ca. 9,000 *Pocillopora* colonies were sampled [branch tip + photographs except for Tromelin Island (Scattered Islands) due to field difficulties and the Society Islands as colonies were not collected for this purpose], independently of their *corallum* macromorphology, from March 2001 to October 2016, within 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) and 18 islands/regions (Table S1; Fig. S1). In each locality, 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. Fragments were preserved in 90% ethanol at room temperature, and deposited at Reunion Island University (Saint-Denis, La Réunion).

All colonies were also previously genotyped with 13 microsatellites (Table S2) and for a subset, we also sequenced the mitochondrial ORF locus (mtORF; see Gélin et al. 2017b for more details; Table S2). 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, covering the totality of the localities and morphotypes sampled, as well as all SSHs and clusters, was considered to maximise the genetic diversity explored (Table S1; Fig. S1). Additionally, four *Seriatopora hystrix* and four *Stylophora pistillata* colonies were sampled in Glorioso Islands (WIO) and Grande Terre (New Caledonia; TSP) to serve as outgroups in the phylogenomic analyses [both species are Pocilloporidae, diverging from the *Pocillopora* genus in the middle-end Paleogene (42.7-28.4 Ma; Simpson et al. 2011)].

**References**

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.

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

Simpson C., Kiessling W., Mewis H., Baron‐Szabo R.C., Müller J. 2011. Evolutionary diversification of reef corals: a comparison of the molecular and fossil records. Evolution. 65:3274–3284.

Spalding M.D., Fox H.E., Allen G.R., Davidson N., Ferdaña Z.A., Finlayson M.A.X., Halpern B.S., Jorge M.A., Lombana A.L., Lourie S.A. 2007. Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. BioScience. 57:573–583.

**Table S1** Sampling localities of *Pocillopora* colonies.

*N*: total number of sampled *Pocillopora* colonies, *NmtORF* and *NHist*: numbers of sequences obtained for the mitochondrial open reading frame (mtORF) and the PocHistone loci, respectively, and *NITS2*: number of colonies high-throughput sequenced for the characterisation of Symbiodiniaceae communities.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Province** | **Ecoregion** | **Island/Region** | **Code** | **Latitude** | **Longitude** | ***N*** | ***NmtORF*** | ***NHist*** | ***NITS2*** |
| Western Indian Ocean (WIO) | Western and Northern Madagascar | Mayotte | MAY | -12.83131 | 45.16044 | 24 | 9 | 4 | 17 |
| Glorioso Islands | GLO | -11.56377 | 47.29394 | 10 | 10 | 2 | - |
| Juan de Nova Island | JDN | -17.04855 | 42.72176 | 20 | 10 | 4 | 17 |
| Europa Island | EUR | -22.36783 | 40.37185 | 13 | 8 | 3 | 13 |
| Northeastern Madagascar | MADne | -16.18321 | 49.94950 | 26 | 5 | 2 | 17 |
| Northwestern Madagascar | MADnw | -13.46366 | 48.25272 | 28 | 8 | 2 | 17 |
| Southwestern Madagascar | MADsw | -23.47539 | 43.66148 | 21 | 18 | 2 | 16 |
| Mascarene Islands | Reunion Island | REU | -21.16115 | 55.57841 | 30 | 15 | 6 | 22 |
| Rodrigues Island | ROD | -19.69775 | 63.44172 | 19 | 9 | 1 | 15 |
| Cargados Carajos/Tromelin Island | | TRO | -15.88083 | 54.52714 | 3 | 3 | - | 3 |
| Tropical Southwestern Pacific (TSP) | New Caledonia | Chesterfield Islands | CHE | -20.41574 | 158.80233 | 46 | 32 | 11 | 32 |
| Western Grande Terre | NCAw | -21.47567 | 165.57125 | 54 | 38 | 6 | 41 |
| Eastern Grande Terre | NCAe | 42 | 12 | - | 30 |
| Loyalty Islands | LOY | -20.96939 | 167.20426 | 7 | 5 | - | 7 |
| Tonga Islands | | TON | -21.13061 | -175.22125 | 3 | 1 | - | - |
| South-East Polynesia (SEP) | Society Islands | Bora-Bora | BOR | -16.50025 | -151.73874 | 2 | 2 | - | 1 |
| Moorea | MOO | -17.52767 | -149.83867 | 7 | 6 | - | 7 |
| Tahiti | TAH | -17.65834 | -149.47704 | 1 | 1 | - | 1 |
|  |  |  |  |  | **Total** | **356** | **192** | **43** | **256** |

**Table S2** Primers and PCR conditions used for the amplification of the mitochondrial open reading frame (mtORF), the PocHistone, the ribosomal RNA internal transcribed spacer 2 (ITS2) and the 13 microsatellite loci. *%NA*: percentage of missing data and *Na*: number of alleles (based on 356 *Pocillopora* colonies).

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Panel** | **Locus** | **Primers (5'-3')** | **Repeat motif** | **Dye** | **Size (bp)** | ***%NA*** | ***Na*** | **Reference** |
| N/A | mtORF | FATP6.1: TTTGGGSATTCGTTTAGCAG | N/A | N/A | N/A | N/A | N/A | [1] |
| RORF: SCCAATATGTTAAACASCATGTCA |
| N/A | PocHistone | F: ATTCAGTCTCACTCACTCACTCAC | N/A | N/A | N/A | N/A | N/A | [2] |
| R: TATCTTCGAACAGACCCACCAAAT |
| N/A | ITS2 | itsD: GTGAATTGCAGAACTCCGTG | N/A | N/A | N/A | N/A | N/A | [3] |
| its2rev2: CCTCCGCTTACTTATATGCTT | [4] |
| 1 | Pd3-004 | F: M13-ACCAGACAGAAACACGCACA | (ATG)8 | 6-FAM | 172-202 | 11.4% | 11 | [5] |
| R: GCAATGTGTAACAGAGGTGGAA |
| Poc40 | F: M13-GTTATTATATGGGTGTATGC | (CAA)X | 6-FAM | 302-344 | 34.4% | 13 | [6] |
| R: CTCAAAGTGCGATTAAAGCC |
| Pd3-005 | F: M13-AGAGTGTGGACAGCGAGGAT | (TGA)9 | NED | 206-264 | 7.2% | 21 | [5] |
| R: GTTCCTTCGCCTTCGATTTT |
| PV2 | F: M13-GCCAGGACCCATTTATACTCC | (GA)20 | VIC | 127-189 | 12.6% | 16 | [7] |
| R: TGCAGTGTTCTACTTGTCAGTGC |
| PV7 | F: M13-GGAGATGGATGGAGACTGC | (GT)5 (CT)2 GT (CT)3 | VIC | 238-266 | 6.7% | 15 | [7] |
| R: GGTATCTCTGTGCTCAGTTCTTTG |
| 2 | Pd2-001 | F: M13-CAGACTTGTCGGAATGAAAGC | (CA)11 | VIC | 206-230 | 6.7% | 15 | [5] |
| R: TTTTGTTTATAAGTCGATACAATGCA |
| Pd2-006 | F: M13-ATCTCCATGTGATCGGCATT | (CA)8 | NED | 202-237 | 6.6% | 13 | [5] |
| R: GTTCCCCCAGCTGAGAAGTT |
| Pd3-008 | F: M13-AGTTGAGGTTGTTGAAACATG | (CTG)7 | 6-FAM | 173-203 | 4.2% | 11 | [5] |
| R: TCCATGCAGAACCCC |
| Pd3-009 | F: M13-CCAATGCGTCCGTAGCTCTC | (CAA)7 (GAG)6 | 6-FAM | 334-369 | 25.8% | 12 | [5] |
| R: ATCACCTAAAAATTTCAGTCCCTTACC |
| 3 | Pd3-EF65 | F: M13-TGTGCAGGTGTTGTGACTGA | (GTT)5 (TGC)11 | PET | 194-242 | 12.5% | 17 | [8] |
| R: TGTCTTTTTCACTTTTGCTTCAA |
| Pd4 | F: M13-ACGCACACAAACCAACAAAC | (AAAC)5 | 6-FAM | 142-184 | 13.3% | 21 | [9] |
| R: TAATTCCATCAACTCAAAGGGG |
| Pd11 | F: M13-TCGTTTGAAGGGAAATGCTC | (CA)7 T (AC)13 | VIC | 144-184 | 28.5% | 20 | [9] |
| R: GGCATGCTATGTATGCGAGA |
| Pd13 | F: M13-TGTTCCTCTCTTTCTCTCTTCCA | (TCTT)5 | NED | 146-236 | 28.9% | 22 | [9] |
| R: CATTTATGTTCCTTTCACGGC |
|  |  |  |  |  | **Total** | **15.3%** | **15.9 ± 1.1** | |

**mtORF/PocHistone:**

PCR mix (total volume = 25 µL): 1X of MasterMix (Applied Biosystems) + 0.3 µM of F primer + 0.3 µM of R primer + 2 ng.µL-1 of genomic DNA.

Thermocycling program: 94°C/5 min + 40 × (94°C/60 s + 53°C/60 s + 72°C/60 s) + 72°C/5 min.

**ITS2:**

PCR mix (total volume = 20 µL): 1X of MasterMix (Applied Biosystems) + 0.3 µM of F primer + 0.3 µM of R primer + 2 ng.µL-1 of genomic DNA.

Thermocycling program: 94°C/5 min + 35 × (94°C/45 s + 54°C/30 s + 72°C/60 s) + 72°C/5 min.

**Microsatellites:**

PCR mix (total volume = 10 µL): 1X of MasterMix (Applied Biosystems) + 0.025 µM of F primer (M13 tailed) + 0.25 µM of R primer + 0.25 µM of fluorescent dyed M13 tail + 2 ng.µL-1 of genomic DNA.

Thermocycling program: 94°C/5 min + 7 × (94°C/30 s + 62°C [-1°C at each cycle]/30 s + 72°C/30 s) + 30 × (94°C/30 s + 55°C/30 s + 72°C/30 s) + 8 × (94°C/30 s + 56°C/30 s + 72°C/30 s) + 72°C/5 min

**References**

[1] Flot J-F, Magalon H, Cruaud C, Couloux A, Tillier S (2008) Patterns of genetic structure among Hawaiian corals of the genus *Pocillopora* yield clusters of individuals that are compatible with morphology. Comptes Rendus Biologies 331:239–247. doi: 10.1016/j.crvi.2007.12.003.

[2] Johnston EC, Forsman ZH, Toonen RJ (2018) A simple molecular technique for distinguishing species reveals frequent misidentification of Hawaiian corals in the genus *Pocillopora*. PeerJ 6:e4355. doi: 10.7717/peerj.4355

[3] Pochon X, Pawlowski J, Zaninetti L, Rowan R (2001) High genetic diversity and relative specificity among *Symbiodinium*-like endosymbiotic dinoflagellates in soritid foraminiferans. Mar Biol 139:1069–1078. doi: 10.1007/s002270100674

[4] Stat M, Pochon X, Cowie R, Gates R (2009) Specificity in communities of *Symbiodinium* in corals from Johnston Atoll. Mar Ecol Prog Ser 386:83–96. doi: 10.3354/meps08080

[5] Starger CJ, Yeoh SS, Dai CF, Baker AC, Desalle ROB (2008) Ten polymorphic STR loci in the cosmopolitan reef coral, *Pocillopora damicornis*. Mol Ecol Res 8:619–621. doi: 10.1111/j.1471-8286.2007.02017.x

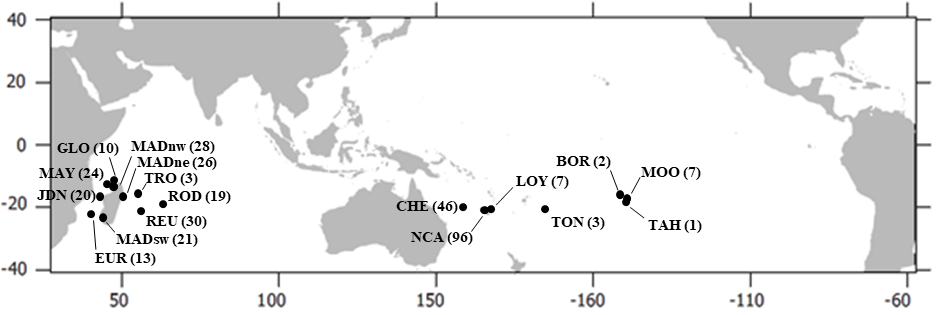
[6] Pinzón JH, LaJeunesse TC (2011) Species delimitation of common reef corals in the genus *Pocillopora* using nucleotide sequence phylogenies, population genetics and symbiosis ecology. Mol Ecol 20:311–325. doi: 10.1111/j.1365-294X.2010.04939.x

[7] Magalon H, Samadi S, Richard M, Adjeroud M, Veuille M (2004) Development of coral and zooxanthella-specific microsatellites in three species of *Pocillopora* (Cnidaria, Scleractinia) from French Polynesia. Mol Ecol Notes 4:206–208. doi: 10.1111/j.1471-8286.2004.00618.x

[8] Gorospe KD, Karl SA (2013) Genetic relatedness does not retain spatial pattern across multiple spatial scales: dispersal and colonization in the coral, *Pocillopora damicornis*. Mol Ecol 22:3721–3736. doi: 10.1111/mec.12335

[9] Torda G, Schmidt-Roach S, Peplow LM, Lundgren P, van Oppen MJH (2013) A rapid genetic assay for the identification of the most common *Pocillopora damicornis* genetic lineages on the great barrier reef. PLoS ONE 8:e584. doi: 10.1371/journal.pone.0058447

**Fig. S1** Sampling localities of *Pocillopora* colonies (number of sampled colonies in parentheses).

****MAY: Mayotte, GLO: Glorioso Islands, JDN: Juan de Nova Island, EUR: Europa Island, MADne: northeastern Madagascar, MADnw: northwestern Madagascar, MADsw: southwestern Madagascar, REU: Reunion Island, ROD: Rodrigues Island, TRO: Tromelin Island, CHE: Chesterfield Islands, NCA: Grande Terre (New Caledonia), LOY: Loyalty Islands (New Caledonia), TON: Tonga Islands, BOR: Bora-Bora, MOO: Moorea and TAH: Tahiti.