



HAL
open science

Délimitation d'espèces et connectivité chez les coraux du genre *Pocillopora* dans l'Indo-Pacifique

Pauline Gélín

► To cite this version:

Pauline Gélín. Délimitation d'espèces et connectivité chez les coraux du genre *Pocillopora* dans l'Indo-Pacifique. Ecologie, Environnement. Université de la Réunion, 2016. Français. NNT : 2016LARE0041 . tel-01623609

HAL Id: tel-01623609

<https://theses.hal.science/tel-01623609>

Submitted on 25 Oct 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Université de La Réunion

École doctorale Science Technologie Santé E.D. n°542

THÈSE

Présentée pour l'obtention du grade de docteur de l'Université de La Réunion

Discipline : Écologie marine

Délimitation d'espèces et connectivité chez les coraux du genre *Pocillopora* dans l'Indo-Pacifique

Pauline GÉLIN

Thèse soutenue le 16 décembre 2016 devant un jury composé de :

Sophie ARNAUD-HAOND	<i>CR HDR, IFREMER Sète</i>	Rapporteur
Didier AURELLE	<i>MCF HDR, Université Aix-Marseille</i>	Rapporteur
François BONHOMME	<i>DR, CNRS</i>	Examineur
Claudie DOUMS	<i>Professeur, EPHE</i>	Examineur
Pascale CHABANET	<i>DR, IRD La Réunion</i>	Examineur
Henrich BRUGGEMANN	<i>Professeur, Université de La Réunion</i>	Examineur
Cécile FAUVELOT	<i>CR HDR, IRD Nouméa</i>	Co-directrice de thèse
Hélène MAGALON	<i>MCF HDR, Université de La Réunion</i>	Directrice de thèse

Résumé

Dans un monde changeant, la conservation et la gestion de la biodiversité sur le long terme nécessitent une connaissance de la répartition des taxons, mais également une estimation précise de leur diversité spécifique ainsi que des processus de spéciation ayant permis leur formation et des liens qui unissent différentes populations d'un même taxon (connectivité). En génétique des populations, ces liens sont les flux de gènes entre différentes populations. Il est crucial de connaître le degré de connectivité des populations, et ce pour différentes espèces marines, afin de déterminer les patrons de dispersion à différentes échelles spatiales et temporelles. La connaissance de ces patrons permet de dessiner des réseaux d'aires marines protégées de façon plus efficace afin de préserver et gérer la biodiversité.

Ce travail de thèse porte sur la connectivité des populations de coraux du genre *Pocillopora* dans le Sud-Ouest de l'océan Indien et l'océan Pacifique tropical. Ces coraux sont répartis sur toute la frange tropicale des océans Indien et Pacifique. Traditionnellement, les espèces étaient identifiées sur la base critères morphologies [17 espèces décrites dans Veron (2000)]. Différentes études utilisant des données génétiques ont révélé que la délimitation des espèces était parfois floue chez ces coraux. Ainsi, au cours de ce travail, l'utilisation de méthodes de délimitation d'espèces à partir d'ADN mitochondrial (ABGD, GMYC, PTP) et nucléaire (haplowebs) 16 hypothèses primaires d'espèces (PSH) ont été identifiées. Ces PSH ont ensuite été confrontées à des tests d'assignement à partir de marqueurs microsatellites, révélant un minimum de 18 hypothèses d'espèces secondaires (SSH).

Une fois que les hypothèses d'espèces sont définies, il est possible de réaliser des études de connectivité. Au cours de ce travail, deux hypothèses d'espèces présentant des écologies différentes ont été choisies pour mener ces analyses. La première, *Pocillopora damicornis* type β (SSH05) a été échantillonnée dans les lagons et la seconde, *Pocillopora eydouxi* (SSH09) a, quant à elle, été échantillonnée sur la pente externe. L'estimation de la structure génétique des populations a permis d'estimer les modes de reproduction (sexuée ou asexuée) chez ces deux hypothèses d'espèces et les analyses de connectivité ont révélé des patterns de structuration complexes pour chacune des SSHs.

Mots-clés : génétique des populations, hypothèse d'espèces, méthodes de délimitation d'espèces, tests d'assignement, scléactiniaires, Sud-Ouest de l'océan Indien, Sud-Ouest de l'océan Pacifique, microsatellites, séquences

Abstract

In a changing world, long-term conservation and management of biodiversity require knowledge of the taxa distribution, but also an accurate estimate of their specific diversity as well as the speciation processes that have allowed their formation and the links between different populations of the same taxon (connectivity). In population genetics, these links are represented by the gene flow between different populations. It is crucial to know the degree of connectivity among populations, and this for different marine species, in order to determine patterns of dispersal at different spatial and temporal scales. Understanding these patterns makes possible to design more accurately networks of marine protected areas in order to preserve and manage biodiversity.

This work focuses on the connectivity among populations of the coral genus *Pocillopora* in the Southwestern Indian Ocean and the Southwestern Pacific Ocean. These corals are widely distributed throughout the tropical fringe of the Indian and Pacific oceans. Traditionally, species were identified on the basis of morphological criteria [17 species described in Veron (2000)]. Different studies using genetic data revealed that the delimitation of species was sometimes blurred in these corals. Thus, in this work, the use of species delineation methods from mitochondrial (ABGD, GMYC, PTP) and nuclear (haplowebs) DNA, 16 primary species hypotheses (PSH) were identified. These PSHs were then confronted to assignment tests from microsatellite loci, revealing a minimum of 18 secondary species hypotheses (SSH).

Once the species hypotheses are defined, it is possible to conduct connectivity studies. In this work, two SSHs with different ecologies were chosen to carry out these analyses. The first, *Pocillopora damicornis* type β (SSH05) was sampled in the lagoons and the second, *Pocillopora eydouxi* (SSH09) was sampled on the outer slope. The estimation of the genetic structure of the populations made possible to estimate the reproductive modes (sexual or asexual) in these two SSHs and the connectivity analyzes revealed complex structuring patterns for each of the SSHs.

Key words: population genetics, species hypotheses, delimitation species methods, assignment tests, scleractinian, Western Indian Ocean, Tropical Southwestern Pacific, microsatellites, sequences

Remerciements

Après un accouchement douloureux, le voilà, le manuscrit, longuement couvé ces dernières années parfois dans la douleur. Merci à tous les copains et tous les collègues, qui m'ont aidée à tenir bon.

Je tiens tout d'abord à remercier les membres du jury, les rapporteurs, Sophie Arnaud-Haond et Didier Aurelle ainsi que les examinateurs, François Bonhomme, Claudie Doums, Pascale Chabanet et Henrich Bruggemann, qui ont tous accepté sans hésiter d'évaluer ce travail.

Je remercie également le LabEx CORAIL pour avoir permis de financer cette thèse et particulièrement Nathalie pour tes encouragements.

Merci à mes deux directrices de thèse, Cécile Fauvelot et Hélène Magalon. Un merci particulier à Hélène qui m'a suivi de (très) près pendant toutes ces années et dont le perfectionnisme débordant m'a poussée et me poussera toujours à tenter de faire mieux. Merci pour ton soutien sans failles (jusqu'au bout !), pour tes conseils, pour les longues soirées labo à essayer de comprendre les données, pour le dépannage d'urgence (finalement elle n'est pas si lourde que ça cette petite voiture !) et surtout merci pour ton amitié.

Merci à ceux qui m'ont accompagnée dans la dernière ligne droite de cette course folle, mes amis, camarades thésards : Agathe pour ton aide inconditionnelle, ton réconfort et ton amitié ; Bautisse pour ton aide, tes conseils et tes réponses à toutes mes questions malgré la distance ; Clément pour ta motivation ; Alina et Yoann pour vos encouragements et votre réconfort de dernière minute ; et tous les autres qui sont déjà passé par là, qui y arriveront bientôt ou qui y échapperont : Ti'Jem, Charlotte, Myriam, Cécile, Marc, Alexis, Simon, Florian, Alice et tous ceux que j'ai pu oublier...

Merci à tous les collègues du laboratoire et en particulier à Gaël pour ta macro super pratique et ton aide pour débayer R.

Merci aux différents stagiaires/esclaves, camarades rats de laboratoire (Agathe et Bautisse) ainsi qu'à Hélène pour m'avoir aidée au laboratoire, à extraire l'ADN de plus de 8000 colonies, soit plus de 85 plaques d'ADN, représentant plus 1 100 plaques de PCR soient plus de 106 000 réactions et plus de 320 000 pics lus (parce qu'il faut bien les lire plusieurs fois...), auxquels on peut ajouter les (presque) 2 000 séquences.. De quoi avoir mal aux yeux.

Merci à Géraldine, Élodie, Charles et toute l'équipe de la plateforme Gentyane de l'INRA de Clermont-Ferrand, pour avoir multiplexé et renvoyé les résultats des plus de 1 100 plaques que j'ai pu vous envoyer. Sans vous, rien n'aurait été possible.

Merci aux amis qui sont loin, mais sur qui j'ai toujours pu compter, Arthur, Morgane, Anne-Laure, Hélène, Boris, Marjorie et Gaëlle. Na r'trouve très vite !

Merci aux données et aux logiciels, toujours prêts à inventer n'importe quoi pour te faire rire (ou pas..), en inventant toujours des résultats plus farfelus les uns que les autres. Quelle imagination !

Merci à l'ASFH pour son soutien moral.

Et parce qu'il faut essayer de n'oublier personne, merci à Melchior et Lilith qui n'y comprendront jamais rien, mais dont les ronronnements ont toujours été réconfortants.

Merci à ma petite famille, ma maman, ma petite sœur et mes grands-parents, qui ont toujours été là, et pour qui il n'est pas nécessaire de dire quoi que ce soit, ça va de soi.

Enfin, merci à Alexis, qui m'a épaulée, soutenue et encouragée tous les jours. Je n'en dirai pas plus, les mots risqueraient d'abîmer le sentiment.

Pour tout, à tous, merci.

Préambule

Face aux risques liés aux changements globaux et à la découverte de l'incongruence entre les données génétiques et morphologiques, les études des diversités spécifique et génétique chez les scléractiniaires sont en plein essor. Cette thèse s'inscrit dans cette dynamique, en tentant d'estimer la diversité spécifique chez les coraux du genre *Pocillopora* et étudier la diversité génétique existante au sein de chaque espèce. Ainsi, la construction de ce travail suit un ordre naturel, partant de la définition d'hypothèses d'espèces (Chapitre 1) grâce à des méthodes de délimitation d'espèce basées sur des données génétiques. Par la suite, la diversité génétique de deux hypothèses d'espèces est estimée : *Pocillopora damicornis* type β et *Pocillopora eydouxi*. La première hypothèse d'espèce *P. damicornis* type β présente l'intérêt de se reproduire de façon sexuée et asexuée. Il s'agit dans un premier temps d'estimer la part relative de reproduction asexuée pour des populations échantillonnées sur quatre sites de la côte Ouest de La Réunion. Ces travaux (Chapitre 2) révèlent une part extrêmement importante de la propagation clonale chez les populations réunionnaises de *P. damicornis* type β . Ensuite, une étude menée à plus grande échelle, entre l'océan Indien (Mayotte, Juan de Nova, Madagascar, La Réunion et Rodrigues) et l'océan Pacifique (Nouvelle-Calédonie) tente d'estimer la structure génétique et la connectivité entre les populations étudiées (Chapitre 3). Par la suite, l'intérêt se porte sur la seconde hypothèse d'espèce, *P. eydouxi*, pour laquelle très peu de données sont disponibles. Ainsi, ce travail (Chapitre 4) permet d'affiner la délimitation des espèces, de spécifier le mode de reproduction et de déterminer les patrons de structuration génétique de ces populations dans l'océan Indien (Mayotte, Juan de Nova, Bassas da India, Europa, Madagascar, La Réunion, Rodrigues et Tromelin) et le Pacifique (plateau récifal des Chesterfield/Bampton/Bellona, Nouvelle-Calédonie et Polynésie Française). Ces différents chapitres ont été rédigés de manière à pouvoir être lus indépendamment, en anglais sous la forme d'articles scientifiques. Ils possèdent chacun leurs références bibliographiques, leurs annexes et leurs figures et tableaux dont la numérotation est unique à chaque chapitre. Les références bibliographiques de l'introduction et de la synthèse générale se situent à la fin du texte principal. La partie annexe qui se situe à la fin du document comprend des travaux publiés parallèlement à ce travail de thèse.

Table des matières

INTRODUCTION.....	13
1. Étudier la biodiversité	13
1.1. Biodiversité et changements globaux.....	13
1.2. Biodiversité et estimations	14
1.3. Biodiversité et écosystèmes coralliens	16
2. Décrire les espèces	16
2.1. La spéciation	17
2.2. Les critères de délimitation des espèces.....	19
2.3. Les méthodes d'identification et de délimitation d'espèces basées sur la génétique	20
2.3.1. Le barcoding.....	20
2.3.2. Automatic Barcoding Gap Discovery (ABGD)	21
2.3.3. L'approche « Generalized Mixed Yule-Coalescent » (GMYC)	22
2.3.4. Poisson Tree Processes.....	22
2.3.5. Haplowebs	23
2.3.6. Les tests d'assignement.....	24
3. Estimer la connectivité des populations	25
3.1. Le cadre conceptuel de la génétique des populations.....	25
3.2. La connectivité génétique des populations marines benthiques.....	26
3.2.1. Les modèles de structuration génétique	27
3.2.2. Les mesures de la différenciation génétique	28
3.3. La connectivité des populations clonales	29
4. Le modèle d'étude : les scléactiniaires du genre <i>Pocillopora</i>	31
4.1. Les scléactiniaires, reproduction et dispersion	32
4.1.1. La reproduction sexuée	32
4.1.2. La reproduction asexuée.....	35
4.2. Le cas des coraux du genre <i>Pocillopora</i>	36
4.2.1. <i>Pocillopora</i> : morphologie et phylogénie.....	36
4.2.2. Ce qui est connu de la reproduction chez <i>Pocillopora</i>	39
4.2.3. La connectivité des <i>Pocillopora</i> à travers le monde	40

5. Problématique.....	42
CHAPITRE 1. Réévaluation du nombre d'espèces, de la distribution et de l'endémisme chez les coraux du genre <i>Pocillopora</i> Lamarck, 1816 grâce à l'utilisation des méthodes de délimitation d'espèces et des microsattellites	45
CHAPITRE 2. Structure des populations à fine échelle, cas du corail <i>Pocillopora damicornis</i> à La Réunion	135
CHAPITRE 3. Etude de la structure génétique et de la connectivité à différentes échelles spatiales chez le corail <i>Pocillopora damicornis</i> type β dans le Sud-Ouest de l'océan Indien et le Pacifique Sud-Ouest	167
CHAPITRE 4. Étude de la structure génétique chez le corail <i>Pocillopora eydouxi</i> : un nouveau puzzle à résoudre	199
SYNTHESE GENERALE ET PERSPECTIVES	239
1. Synthèse des résultats et discussion	240
1.1. La délimitation des espèces chez <i>Pocillopora</i>	240
1.2. Spécificité des patterns modes de reproduction en fonction des hypothèses d'espèces	242
1.3. Spécificité des patterns de structuration génétique en fonction des hypothèses d'espèces	244
2. Conclusions et perspectives	247
ANNEXES	259
1. Isolation and characterization of 22 microsatellite loci from two coral species: <i>Acropora muricata</i> (Linnaeus, 1758) (Scleractinia, Acroporidae) and <i>Porites lutea</i> Milne-Edwards & Haime, 1851 (Scleractinia, Poritidae)	
2. The fine-scale genetic structure of the malaria vectors <i>Anopheles funestus</i> and <i>Anopheles gambiae</i> (Diptera: Culicidae) in the north-eastern part of Tanzania	
3. One species for one island? Unexpected diversity and weak connectivity in a widely distributed tropical hydrozoan	

INTRODUCTION

1. Étudier la biodiversité

1.1. Biodiversité et changements globaux

La biodiversité représente la diversité des formes de vie sur Terre, plusieurs niveaux peuvent être utilisés pour la décrire tels que la diversité des écosystèmes, la diversité des espèces, la diversité des gènes ou encore la diversité de traits fonctionnels (Cardinale et al., 2012).

Pour schématiser, la biodiversité peut être perçue comme un enchaînement de boîtes interdépendantes. Prenons la Terre, celle-ci est parsemée d'écosystèmes différents les uns des autres. Chacun d'entre eux présente des services écosystémiques différents tels que la production de l'oxygène de l'air ou l'épuration naturelle des eaux. Pour qu'un écosystème soit en bonne santé, il est nécessaire que chaque composante de l'écosystème, c'est-à-dire que les différentes formes de vie ou espèces qui le composent soient également en bonne santé. Le maintien d'une espèce passe souvent par son association avec les autres formes de vie qui l'entourent mais aussi par des paramètres abiotiques tels que la température ou la pluviométrie. L'adaptabilité des espèces à des conditions environnementales particulières est le résultat de l'évolution des gènes soumis, via les phénotypes des individus, aux pressions de sélection, exprimant ainsi des protéines particulières capables d'améliorer la survie des espèces dans des conditions environnementales particulières. Ainsi, la variété des écosystèmes est corrélée au nombre de formes de vie existantes, elles-mêmes étant la résultante de la variabilité des gènes. La préservation de la biodiversité dans son intégralité permet ainsi de maintenir une grande variabilité de services écosystémiques.

A l'inverse, la perte de biodiversité risque d'entraîner une transformation des habitats. L'appauvrissement des écosystèmes, les changements climatiques, les espèces exotiques envahissantes, la surexploitation des espèces et la pollution sont autant de facteurs qui entraînent l'appauvrissement de la diversité biologique et de la variabilité génétique des populations naturelles ou de l'extinction de certaines espèces (Vitousek et al., 1997). Aujourd'hui, l'humanité est entrée dans une nouvelle ère géologique : l'anthropocène, nommée ainsi car les stigmates de l'activité humaine sont déjà visibles (Crutzen, 2006; Smith and Zeder, 2013). Pour l'« *Anthropocene Working Group* », cette époque se distingue de toutes les autres

grâce à des marqueurs stratigraphiques spécifiques, décelables dans les roches sur toute la surface du globe, et légués par les activités humaines à partir de 1945. Ce terme a été popularisé à la fin du 20^e siècle par le météorologue et chimiste de l'atmosphère Paul Crutzen, pour désigner une nouvelle époque géologique, qui aurait débuté selon lui à la fin du 18^e siècle avec la révolution industrielle, et succéderait ainsi à l'Holocène.

Bien qu'ayant subi plusieurs épisodes d'extinctions massives, aujourd'hui la biodiversité connaît un sixième épisode d'extinctions massives d'une intensité et d'une rapidité sans précédent. C'est dans ce contexte que des hotspots de biodiversité, terrestres et marins ont été définis (Mittermeier et al., 2008; Myers et al., 2000; Roberts et al., 2002) car ils combinent à la fois une biodiversité extraordinaire (au moins 1 500 plantes terrestres endémiques) et des taux d'extinctions phénoménaux (perte de 70 % de leur surface). Ces zones sont très étudiées car leur préservation est primordiale.

1.2. Biodiversité et estimations

Pour comprendre un écosystème, il faut d'abord le décrire, connaître les espèces et les habitats qui le composent. L'accumulation de connaissances permet alors de définir des ensembles prioritaires pour les mesures de conservation.

Un des problèmes qui réside dans la description des écosystèmes est la description des espèces, unité de base de la conservation. Originellement, les espèces ont été décrites sur la base de critères morphologiques, mais la multiplication des études basées sur la génétique a révélé l'existence d'espèces cryptiques, c'est-à-dire d'espèces morphologiquement identiques, mais génétiquement différentes. Parfois, même des animaux emblématiques, longtemps considérés comme bien connus, peuvent s'avérer surprenants, comme la girafe par exemple : il était admis qu'il existait une seule espèce composée de neuf sous-espèces (Dagg and Foster, 1976) correspondant à des variations légères de la couleur de la robe et de leur répartition géographique, mais l'analyse génétique des neuf sous-espèces a révélé l'existence de quatre espèces cryptiques de girafes (Fennessy et al., 2016).

Lorsque les espèces sont définies, il faut ensuite procéder à l'étude de leurs traits de vie pour comprendre leur fonctionnement et rassembler des informations concernant notamment la reproduction (fréquence des événements de reproduction, sexuée et/ou asexuée, nombre de descendants, temps de générations, succès reproducteur, etc.) et les capacités de dispersion et la connectivité entre populations. La connectivité représente les flux d'individus existants entre différentes localités. De manière générale, deux populations sont considérées connectées

lorsque les propagules (gamètes, larves, juvéniles, adultes) ont rejoint une autre population, dans laquelle ils ont réussi à s'installer, survivre. Pour que des populations soient génétiquement connectées, il est nécessaire que les migrants se soient reproduits avec les individus résidents (*sensu* Pineda et al., 2007), on parle alors de dispersion efficace (Sale et al., 2005). En effet, il est important d'étudier la connectivité génétique car elle joue un rôle fondamental dans la distribution spatiale des individus, des allèles et des espèces. Elle permet de mesurer comment la diversité génétique de l'espèce se répartit entre populations dans l'espace. Outre son rôle central dans la dynamique et l'évolution des populations structurées spatialement, elle permet (1) la cohésion génétique de l'espèce, (2) sa persistance dans le temps et l'espace et (3) la colonisation d'habitats favorables dans un monde changeant.

Pour certains organismes, l'acquisition des connaissances sur tous ces traits d'histoire de vie peut se faire grâce à des observations *in situ*. Pour d'autres organismes, les suivis directs sont plus difficiles [grands migrateurs, bactéries (suivi des épidémies), organismes sessiles dont la dispersion est assurée par l'émission de propagules] et des méthodes indirectes devront être utilisées comme les suivis satellites, la méthode de capture-marquage-recapture, la modélisation des vecteurs de dispersion (courants marins : hydrochorie, vent : anémochorie, animaux : zoochorie), ou l'estimation des flux de gènes correspondant implicitement aux flux d'individus (les flux de gènes ne sont visibles que si les individus migrants se reproduisent avec les individus résidents dans la population d'arrivée) notamment grâce au cadre théorique de la génétique des populations. La génétique des populations permet, en étudiant la structure génétique des populations, de mesurer les degrés de différenciation génétique entre populations mais également d'estimer des indices de consanguinité, permettant ainsi d'obtenir des informations sur les modes de dispersion et les modes de reproduction des espèces et populations étudiées. Connaître les capacités de dispersion d'une espèce permettra d'estimer si des migrants seront capables de recoloniser des sites endommagés ou de trouver de nouveaux habitats favorables. Supposément, des individus ayant une capacité de dispersion importante pourront migrer d'une population à une autre, permettant de connecter les populations entre elles. Ainsi, il sera possible de prévoir les éventuels impacts d'une perturbation sur une espèce et estimer les capacités de résilience d'une population (capacité de retrouver son état d'origine), mais aussi proposer des mesures de protection pour mieux conserver la biodiversité.

A ce jour, 35 hotspots de biodiversité ont été décrits représentant des zones prioritaires pour la conservation de la biodiversité, dont font partie les récifs coralliens.

1.3. Biodiversité et écosystèmes coralliens

Les récifs coralliens regorgent d'une biodiversité exceptionnelle. En effet, plus de 5 % de la biodiversité mondiale décrite (en nombre d'espèce) y vit bien qu'ils ne représentent que 0,1 % de la surface du globe (Reaka-Kudla et al., 1997). Les scléactiniaires, ou coraux durs bio-constructeurs des récifs coralliens, sont la clé de voute de cet écosystème présent sur toute la surface du globe entre 30°N et 30°S. Bien que d'autres organismes participent à consolider les récifs (par exemple, les algues calcaires), leur structure tridimensionnelle offre un habitat pour une multitude d'organismes (poissons, algues, invertébrés, bactéries, etc.). Les scléactiniaires représentent un groupe fonctionnel ayant largement contribué à la formation des récifs coralliens depuis au moins 200 millions d'années (Veron, 1995).

De nombreuses populations humaines dépendent directement de la durabilité de cet écosystème, alors même que celui-ci risque de connaître un grand bouleversement lié aux nombreuses pressions anthropiques qu'il subit. La pêche ou la protection des côtes sont autant de services écosystémiques qui pourraient disparaître si cet écosystème déclinait. C'est dans ce contexte que les études se multiplient afin de mieux comprendre le fonctionnement des écosystèmes menacés, décrire les espèces et leur évolution dans le temps et l'espace afin de proposer des mesures pour limiter l'impact de l'anthropisation sur ces derniers.

Les coraux durs, ou scléactiniaires, sont des organismes benthiques sessiles. Ainsi, la phase dispersive est essentiellement assurée par l'émission des gamètes ou la production de larves. Etant donné la complexité de suivre directement les mouvements des larves ou des gamètes dans la colonne d'eau, la génétique des populations est très utilisée pour comprendre la structuration des populations marines benthiques et ainsi leur connectivité.

L'utilisation d'outils moléculaires a permis de révéler que chez de nombreux scléactiniaires, la génétique et la morphologie (historiquement utilisée pour décrire les espèces) manquaient de congruence dans de nombreux cas (par exemple, Flot et al., 2011; Pinzón et al., 2013; Schmidt-Roach et al., 2014b). Ainsi, la première étape pour l'étude de la connectivité entre populations est de décrire et délimiter les espèces afin de mener les études sur des groupes d'individus supposés appartenir à la même espèce et non pas des complexes d'espèces, ce qui pourrait biaiser les résultats, comme ce fut probablement le cas par le passé.

2. Décrire les espèces

Bien décrire les espèces *a priori* est essentiel pour comprendre leur fonctionnement. Cependant, leur description n'est pas toujours aisée, notamment à cause du caractère progressif de la spéciation.

2.1. La spéciation

La spéciation est un processus lent et progressif. Ainsi, suite à la divergence d'une espèce ancestrale en deux lignées divergentes, la spéciation est considérée comme totalement achevée lorsqu'il est possible d'observer et de discerner, sans ambiguïté, les deux nouvelles espèces formées. En revanche, si le processus de spéciation n'est pas terminé, on se trouve dans une zone grise (Figure 1). De chaque côté de cette zone, le nombre d'espèces sera communément admis, mais au sein de cette zone grise, la description des espèces peut devenir conflictuelle.

De manière générale, tout phénomène de spéciation comprend la perte de la capacité de reproduction de deux populations historiquement interfécondes. Cependant, les processus (ou modes de spéciation) menant à l'isolement reproducteur des populations peuvent être pluriels.

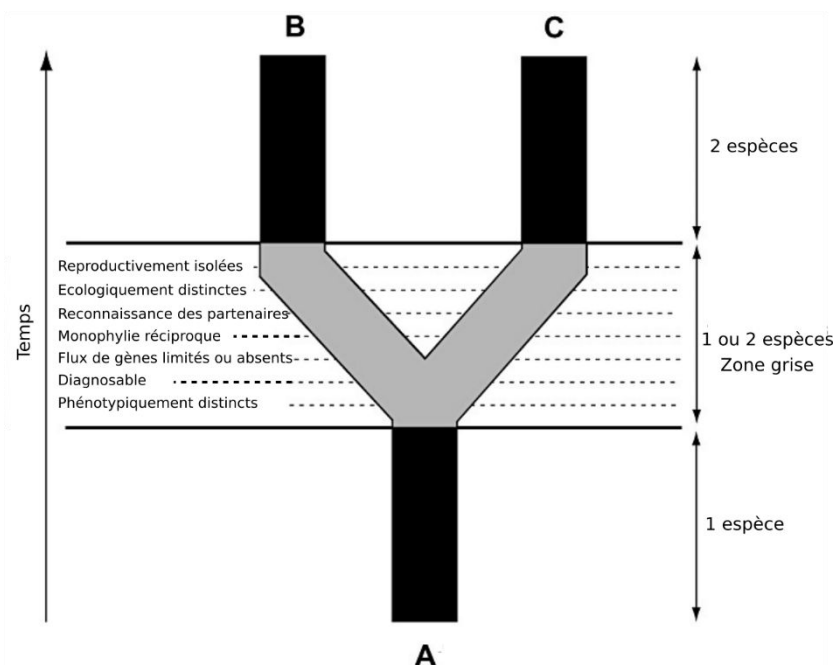


Figure 1. Représentation schématique du concept d'espèces génalogiques unifiées (adapté de (Braby et al., 2012)) montrant une lignée ancestrale (espèce A) qui diverge au cours du temps pour former deux lignées filles (espèces B et C). La zone grise entre les deux lignes noires représente le temps au cours duquel les lignées filles acquièrent différents caractères (représentés par les lignes pointillées), et qui servent de critères opérationnels pour identifier les limites entre les espèces sous différents concepts.

Le modèle le plus facile à comprendre est le mode de spéciation dit allopatrique. La spéciation allopatrique comprend plusieurs mode de spéciation : elle est dite (1) vicariante lorsqu'une barrière aux flux de gènes persiste dans le temps, ou (2) péripatrique (ou par effet fondateur) lorsqu'un nombre réduit d'individus colonise de nouveaux milieux en dehors de l'aire de répartition d'origine.

Dans la spéciation allopatrique, l'entrave au flux génétique entre les populations est la formation d'une barrière géographique physique (montagnes, océans, déserts, distance géographique, etc.). Le maintien de cette barrière, combiné aux pressions évolutives (dérive génétique aléatoire, sélection naturelle et accumulation de mutations) conduira au cours du temps à une différenciation génétique si grande entre les deux populations que celles-ci ne pourront plus se reproduire entre elles, même si elles se retrouvent de nouveau en contact.

Outre la spéciation allopatrique, on parle de spéciation parapatrique lorsque la barrière séparant les populations reste perméable aux flux de gènes. Dans ce cas, les populations possèdent une zone de contact étroite où il est possible d'observer des hybrides. Par ailleurs, l'isolement géographique n'est pas toujours le premier pas vers la spéciation. Parfois, deux lignées sympatriques (vivant au même endroit) divergent à partir d'un même ancêtre commun. Dans ce cas, c'est l'incompatibilité entre partenaires sexuels qui provoque l'isolement ; on parle alors de spéciation sympatrique. Ce dernier mode de spéciation a été mis en évidence chez des poissons cichlidés du lac du cratère Apoyo au Nicaragua, deux espèces étant apparues sans isolement géographique (Barluenga et al., 2006).

Au cours de la spéciation, la divergence des caractères varie quantitativement, les dernières étapes étant caractérisées par une différenciation marquée (caractères fixés) à plusieurs niveaux (Figure 1) : la monophylie réciproque de la majorité des généalogies de gènes et l'isolement reproducteur (Mallet, 2008). Lorsqu'ils ne sont pas soumis à la sélection, la fixation des caractères est stochastique, notamment à cause de processus neutres comme la dérive génétique. Dans des cas de spéciation directionnelle, il est attendu que les caractères sujets à la sélection se fixent au cours des premières étapes de la divergence (Nosil et al., 2009; Streelman and Danley, 2003). C'est particulièrement évident dans des cas de sympatrie ou parapatrie quand la sélection, en réponse à des perturbations, entraîne une fixation des caractères améliorant la valeur sélective (fitness) (Nosil et al., 2009; Rueffler et al., 2006). Dans la majorité des scénarios, l'origine de l'isolement reproducteur sera le résultat des effets d'épistasie en lien avec de nouvelles mutations [si un ou plusieurs gènes (dominants ou récessifs) masquent ou empêchent l'expression d'autres gènes] sous sélection directionnelle ou balancée (Schluter, 2009). En théorie, il est attendu que l'isolement reproducteur intervienne plus rapidement si

s'ajoute la sélection directionnelle, que ce soit en allopatrie (*e.g.* Fitzpatrick, 2002), en sympatrie ou en parapatricie (pour une revue, voir Nosil et al., 2009).

Comprendre comment se sont formées les espèces est un chemin pavé d'embûches. Répondre à la question : « Qu'est-ce qu'une espèce ? » est un domaine de recherche à part entière et l'émulation née a entraîné la définition de multiples concepts d'espèces.

2.2. Les critères de délimitation des espèces

La variabilité des concepts d'espèces réside dans les critères d'identification et chaque pan de la recherche a défini son concept.

Ainsi plusieurs concepts de base ont été définis à partir desquels d'autres ont été redéfinis : les espèces biologiques (Mayden, 2002; Mayr, 1942), évolutives (*e.g.* Samadi and Barberousse, 2006; Simpson, 1951), morphologiques (Michener, 1976), écologiques (Van Valen, 1976), phylogénétiques (Baum and Shaw, 1995; Donoghue, 1985; Meier and Willmann, 2000)), correspondant à un groupe génotypique (Mallet, 1995), ou cohésives (interchangeabilité des individus entre les populations (Templeton, 1998)] ; résumé dans (De Queiroz, 2007).

Cependant, chaque concept présente des avantages et des limites. A titre d'exemple, la définition d'espèces basées sur des critères morphologiques est commode, puisque les espèces peuvent être décrites directement, par une simple observation de caractères morphologiques. Cependant, elle peut être soumise à une grande subjectivité de la part de l'observateur et peut s'avérer difficile lorsque les organismes présentent une plasticité phénotypique importante. Par ailleurs, le concept d'espèce biologique qui considère un isolement reproducteur entre les espèces peut être facile à mettre en évidence (par exemple, le chat et le chien ne se reproduisent pas ensemble, ce sont donc des espèces différentes). Cependant, il ne permet pas de considérer les cas d'hybridation (par exemple, l'âne et le cheval peuvent se reproduire, appartiennent-ils chacun à une espèce différente ?). Le concept d'espèce phylogénétique retrace quant à lui l'évolution à partir de gènes, et il nécessite de choisir les gènes adéquats car tous ne présentent pas la même histoire évolutive et peuvent donc raconter des histoires évolutives différentes. En effet, certains allèles peuvent être maintenus dans différentes lignées (par la dérive, par exemple), on parle alors de « tri incomplet des lignées » (*incomplete lineage sorting*).

Chaque concept pris seul peut donner une vision tronquée de la réalité. De plus en plus d'études privilégient une approche intégrative pour délimiter les espèces (Dayrat, 2005; Padial et al., 2010; Schlick-Steiner et al., 2009). Cette approche suggère de combiner différents critères (par

exemple, morphologie, phylogénie, écologie, traits d'histoire de vie ou chimie) afin d'obtenir une vision plus proche de la réalité. Cependant, cette méthode peut être difficile à mettre en œuvre par l'importante quantité de données qu'il est nécessaire de collecter. Utiliser plusieurs types de données génétiques (ADN mitochondrial et nucléaire) peut représenter un premier pas vers la méthode intégrative.

2.3. Les méthodes d'identification et de délimitation d'espèces basées sur la génétique

Les méthodes de délimitation d'espèces permettent de proposer des hypothèses d'espèces qu'il sera nécessaire de vérifier en combinant d'autres critères.

2.3.1. Le barcoding

Né au début des années 2000, le *barcoding* ou *barcode* est devenu rapidement très populaire. Basé sur le séquençage d'un ou plusieurs gènes diagnostiques pour chaque règne [COI pour les animaux (Hebert et al., 2003a) ; rbcL et matK pour les plantes (Cbol Plant Working Group et al., 2009) ou ITS pour les champignons (Schoch et al., 2012)], il repose sur l'hypothèse que les distances génétiques au sein d'une seule espèce sont plus faibles que les distances génétiques entre espèces différentes.

L'utilisation du *barcode* est relativement simple, mais elle suggère de pouvoir comparer une séquence avec d'autres afin de savoir à quel organisme elle correspond. Pour cela, des banques de données génétiques comme GenBank ou Barcoding Of Life Data Systems [BOLD Systems, créée dans ce but, (Ratnasingham and Hebert, 2007)] permettent une analyse comparative d'une séquence avec toutes celles présentes dans la banque [Analyse *Basic Local Alignment Search Tool* ou BLAST ; (Altschul et al., 1990)].

Cependant, bien qu'universels, ces gènes *barcode* n'ont pas la même efficacité ou pouvoir discriminant dans toutes les lignées de chaque règne, comme par exemple, le gène COI ne peut pas être utilisé comme *barcode* chez les cnidaires car il est trop peu variable (Hebert et al., 2003b). L'utilisation du *barcode* à grande échelle soulève certains problèmes méthodologiques en ce qui concerne (1) les échantillonnages avec des effectifs permettant la détection d'une discontinuité entre les distances génétiques inter- et intra-spécifiques, (2) la fixation des spécimens permettant une bonne préservation de l'ADN, (3) l'extraction d'ADN sur des spécimens anciens conservés par des méthodes incompatibles avec la préservation de l'ADN

(par exemple, formol), (4) la gestion des grandes collections d'ADN, (5) la mise à disposition de ces données et leur accessibilité pour la communauté scientifique et (6) le problème d'identification des individus et de leurs séquences contenues dans ces collections. Il reste cependant indéniable que l'utilisation du *barcode* présente un intérêt considérable en écologie des communautés et en phylogéographie, permettant ainsi de décrire la composition spécifique des communautés et les variations spatio-temporelles de ces dernières.

La quantité de données issues du *barcoding* ayant rapidement augmenté, des méthodes d'analyses de ces données ont été développées pour délimiter les hypothèses d'espèces.

2.3.2. Automatic Barcoding Gap Discovery (ABGD)

L'*Automatic Barcoding Gap Discovery* (ABGD) repose sur le principe du *barcode*. Ainsi, pour identifier les limites entre espèces, des distances deux à deux sont calculées entre les séquences. Classiquement, les distances présentent une distribution bimodale permettant d'identifier le *barcode gap* supposé correspondre à la limite entre les espèces (Figure 2). En dessous de ce seuil, les séquences appartiennent à la même espèce ; au-dessus, à différentes espèces. Cette méthode est indépendante de la topologie de l'arbre construit à partir des données de séquences.

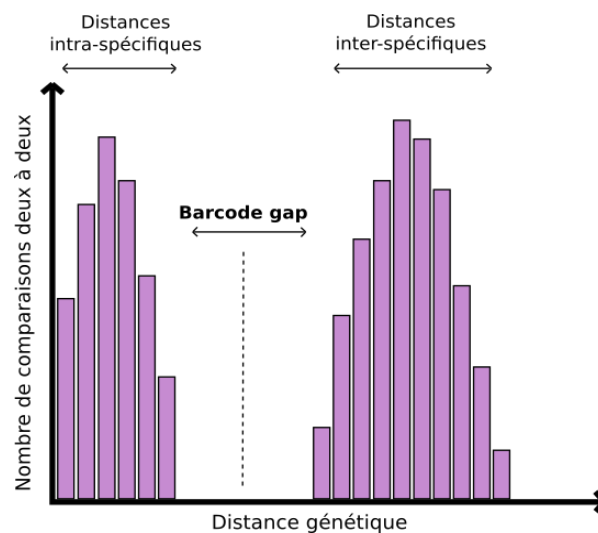


Figure 2. Représentation schématique de l'*Automatic Barcoding Gap Discovery* (ABGD). Le nombre de comparaisons deux à deux est représenté en fonction de la distance génétique. La distribution est bimodale. Le premier mode est supposé représenter la divergence intra-spécifique alors que le second mode, correspondant à des distances plus grandes représenterait les distances génétiques entre espèces différentes.

La méthode a été automatisée par Puillandre et al. (2012). Ainsi, à partir d'un alignement de séquences *barcode*, le *barcode gap* est détecté automatiquement et les séquences sont réparties en groupes correspondant aux différentes hypothèses d'espèces.

Cette méthode est basée sur les caractéristiques propres des séquences sans reconstruction phylogénétique. D'autres méthodes ont également été proposées, se basant sur la topologie des arbres.

2.3.3. L'approche « Generalized Mixed Yule-Coalescent » (GMYC)

L'approche GMYC (Generalized Mixed Yule-Coalescent; (Fontaneto et al., 2007; Fujisawa and Barraclough, 2013; Fujita et al., 2012; Pons et al., 2006) se base sur la topologie des arbres phylogénétiques reconstruits avec les marqueurs pour inférer des hypothèses d'espèces. En utilisant une fonction de vraisemblance qui modélise les processus évolutifs, ce modèle suppose que pour chaque nœud d'un arbre phylogénétique deux événements sont possibles : un événement de divergence entre deux espèces suivant un processus de spéciation de Yule (pas d'extinction, Yule, 1924) ou un événement de coalescence (construction rétrospective de la généalogie des gènes) entre lignées de la même espèce (Kingman, 1982). Le modèle de Yule suggère que chaque lignée, indépendamment des autres et indépendamment du passé, possède une probabilité donnée sur un pas de temps donné de se diviser en deux lignées (processus de spéciation ou cladogénèse). Par ailleurs, comme les événements de coalescence sont supposés survenir à une fréquence plus importante que les événements de spéciation, il est alors possible d'identifier une limite sur l'arbre phylogénétique entre la divergence inter- et intra-spécifique, délimitant ainsi des groupes de feuilles de l'arbre phylogénétique. Ces groupes représentent des métapopulations indépendantes: des lignées isolées et ayant évolué indépendamment les unes des autres, au sein desquelles opèrent mutation, sélection et dérive (Fujita et al., 2012), c'est-à-dire des hypothèses d'espèces.

Cette méthode requiert un arbre calibré dans le temps comme base pour l'analyse. Ainsi, il est nécessaire d'utiliser une horloge moléculaire pour calibrer les arbres. Comme le calibrage d'un arbre n'est pas toujours facile, d'autres méthodes ont été développées.

2.3.4. Poisson Tree Processes

Pour pallier le problème de calibration de l'arbre pour délimiter les espèces, la méthode PTP (Poisson Tree Processes ; Zhang et al., 2013) considère le nombre de substitutions entre les ramifications d'un arbre et les événements de spéciation, en modélisant les événements de spéciation en utilisant le nombre de substitutions plutôt que le temps, chaque événement de spéciation étant considéré comme un événement rare. Cette méthode fait l'hypothèse que chaque mutation présente une probabilité non nulle de former une nouvelle espèce et, par conséquent, que le nombre de substitutions entre espèces est significativement supérieur à celui qui est observé au sein d'une espèce. L'hypothèse sous-jacente est que chaque substitution, indépendamment des autres, possède une probabilité non nulle, bien que faible, de mener à un événement de spéciation, mais qu'au sein de chaque espèce le taux de substitution est grand. Le modèle cherche ainsi un point de transition où les schémas de ramifications de l'arbre changent entre les niveaux inter- et intra-spécifiques et suppose que les longueurs de branches sont le résultat de deux processus de Poisson différents. Le premier processus décrit la spéciation de telle sorte que le nombre moyen de substitutions jusqu'au prochain événement de spéciation suit une distribution exponentielle. Le second processus décrit les événements de ramification au sein d'une espèce de façon analogue aux événements de coalescence.

2.3.5. Haplowebs

Les méthodes décrites précédemment ont été dessinées pour être utilisées à partir de marqueurs *barcode*, souvent mitochondriaux notamment chez les animaux. Les marqueurs mitochondriaux permettent d'obtenir rapidement un grand nombre de caractères variables, en s'affranchissant des problèmes de recombinaison et de polymorphisme individuel propres à l'ADN nucléaire.

Une méthode a été proposée pour être utilisée à partir de marqueurs séquences nucléaires permettant de considérer les cas d'hétérozygotie chez les organismes diploïdes. Les haplowebs (Flot et al., 2010) sont des réseaux de séquences nucléaires sur lesquels sont ajoutées des connections entre les haplotypes qui sont trouvés chez un même individu hétérozygote. L'hypothèse est qu'une espèce possède son lot d'haplotypes, lesquels ne peuvent être retrouvés chez d'autres espèces. Ainsi, il est possible de définir des sl-FFR (single locus Fields For Recombination ; (Doyle, 1995)) définissant des « groupes d'haplotypes » représentant des hypothèses d'espèces.

2.3.6. Les tests d'assignement

À ces méthodes de délimitation d'espèces basées sur les séquences, il est aussi possible d'ajouter les données de génotypage issus de marqueurs microsatellites. Les données issues de marqueurs microsatellites peuvent être utiles pour redéfinir les limites entre les espèces. En effet, les microsatellites sont connus pour avoir un taux de mutation élevé (et supérieur à celui des séquences) induisant un polymorphisme élevé, les rendant utiles pour estimer la proximité génétique entre taxons proches (e.g. Guichoux et al., 2011). Cependant, leur taux de mutation important favorise l'homoplasie, ce qui peut rendre difficile l'interprétation de leur évolution (Slatkin, 1995) et rendant leur utilisation inappropriée au-delà d'espèces génétiquement proches (e.g. Mesak et al., 2014).

Cependant, il est possible d'utiliser les locus microsatellites communs aux différents organismes d'un même taxon, du fait que les microsatellites sont souvent conservés au niveau du genre. Grâce à des tests d'assignement par approche bayésienne ou par des méthodes statistiques multivariées, il est possible de distinguer un premier niveau de structuration créant des groupes génétiques correspondant aux hypothèses d'espèces putatives. La première de ces méthodes (STRUCTURE ; Pritchard et al., 2000) vise à créer des groupes génétiques qui suivent les hypothèses d'une population sous le modèle d'Hardy-Weinberg et qui minimise le déséquilibre de liaison. La deuxième méthode est l'analyse discriminante en composantes principales (DAPC). Elle ne fait aucune hypothèse sur les groupes génétiques auxquels les individus seront assignés, mais elle minimise la variance intra-groupes et maximise la variance inter-groupes (Jombart et al., 2010). Pour chacune de ces méthodes, des méthodes de détermination du nombre le plus probable de groupes génétiques ont été développées [ΔK , (Evanno et al., 2005) ; DIC, (Gao et al., 2011) ou TI, (Verity and Nichols, 2016)].

Les différentes méthodes décrites ci-dessus permettent chacune de définir des hypothèses d'espèces. De manière évidente, les méthodes donnent des résultats différents et il est nécessaire de les comparer et de trouver un critère pour définir un résultat consensus de l'ensemble des méthodes. À ce jour, il n'existe pas d'accord sur le critère à utiliser. Ainsi, certaines études ont utilisé le principe de parcimonie, c'est-à-dire que le nombre d'hypothèses d'espèces choisi est celui qui préserve au mieux l'intégrité du jeu de données (par exemple, Postaire et al., 2016). Cependant, ce critère est susceptible de sous-estimer le nombre réel d'espèces par son approche conservative. Par conséquent, une approche moins conservative serait d'observer les limites entre les hypothèses d'espèces données par les différentes méthodes (point de troncature) et de

définir les limites entre les hypothèses d'espèces dans les cas où le point de troncature apparaît le plus grand nombre de fois (sur l'ensemble des méthodes utilisées).

Une fois que les hypothèses d'espèces sont décrites, il est possible d'étudier chacune d'entre elles pour savoir comment leurs populations se structurent génétiquement, et donc estimer les capacités de dispersion et les modes de reproduction de chacune d'entre elles. La génétique des populations permet d'identifier de manière indirecte certains traits de vie des espèces, dont ceux précités.

3. Estimer la connectivité des populations

3.1. Le cadre conceptuel de la génétique des populations

La génétique des populations est née au début des années 1920 de la volonté de concilier la théorie darwinienne de l'évolution et les données acquises sur la transmission du matériel héréditaire. Elle repose largement sur la construction de modèles mathématiques dont l'objectif est de rendre compte de l'évolution en étudiant les fréquences des allèles et des génotypes, et les forces susceptibles de modifier ces fréquences au cours des générations successives. Certains de ces facteurs comme la sélection, les mutations, la dérive génétique et les migrations peuvent changer la fréquence des allèles et des génotypes.

Il existe quatre grandes forces qui exercent des pressions sélectives sur les allèles modulant ainsi la diversité génétique entre les populations : sélection, migration, mutation et dérive. La sélection naturelle représente la survie non aléatoire des génotypes. Elle peut être de plusieurs formes : directionnelle (si un génotype présente un avantage évolutif), balancée (s'il existe plusieurs optimums phénotypiques) ou stabilisatrice (éliminant les génotypes extrêmes et maintenant les génotypes intermédiaires). La dérive génétique est le résultat de l'évolution d'une population ou d'une espèce causée par des phénomènes aléatoires, impossibles à prévoir. Elle se produit par un échantillonnage aléatoire des allèles à chaque génération. Les fréquences alléliques fluctuent ainsi au cours du temps allant parfois jusqu'à causer la perte de certains allèles. L'impact de la dérive génétique dépend grandement de la taille efficace (nombre d'individus qui participent « efficacement » à la génération suivante en se reproduisant, plus la taille efficace est petite, plus le maintien de la variabilité génétique à long terme est compromise) des populations dans lesquelles elle opère. La migration, à condition que les migrants se reproduisent avec les individus des populations d'accueil, permet d'apporter des

nouveaux allèles dans la population, faisant ainsi évoluer les fréquences alléliques au cours des générations et tendant vers une homogénéisation des fréquences entre les populations liées par la migration. La mutation représente une modification de l'ADN pouvant mener à l'apparition de nouveaux allèles qui pourront se répandre dans les populations.

Toutes ces forces influencent la structure génétique des populations, notamment en faisant varier les fréquences alléliques. L'apparition de certains allèles et leur accumulation dans différentes populations peut conduire, en combinaison avec la dérive et la sélection, au processus de spéciation. A l'inverse, si la migration est très importante, les fréquences alléliques peuvent s'homogénéiser entre les populations. Ainsi, grâce à la génétique des populations, il est possible d'étudier la structure génétique des populations et estimer les degrés de connectivité génétique entre populations.

3.2. La connectivité génétique des populations marines benthiques

La connectivité en milieu marin, c'est-à-dire le processus qui lie les habitats et les populations par l'échange d'organismes, est un élément-clé de la dynamique des populations, de la structuration génétique et des processus de diversification des organismes marins (Bowen et al., 2013; Cowen and Foundation, 2003; Palumbi, 1992; Paulay and Meyer, 2002). Chez les scléroractiniaires qui sont des organismes benthiques, la connectivité dépend des capacités dispersives de leurs stades larvaires (processus qui inclut le transport et le recrutement, c'est-à-dire l'arrivée dans le nouvel habitat et la fixation sur le substrat) pour le maintien de leurs populations dans le temps et la colonisation de nouveaux habitats. Le transport des larves, leur déplacement dans la masse d'eau, est la résultante de deux phénomènes : le transport passif (advection et diffusion turbulente) et le transport actif (comportement natatoire des larves) (Ayata, 2010). En effet, entre le moment où la larve est formée et le moment où elle devient compétente (prête pour le recrutement), les courants marins peuvent déterminer où elle pourra s'établir, à condition qu'elle y trouve des habitats favorables. Cependant, la puissance des courants peut balayer les larves qui tenteraient de s'installer, les empêchant de se fixer dans un endroit précis (Jackson, 1986). En outre, une relation a été trouvée entre la durée de la phase planctonique et la structure génétique des populations (Bohonak, 1999). La dispersion sera donc plus ou moins grande en fonction des espèces et de leur mode de reproduction (fécondation interne ou externe; Amar et al., 2007; Nishikawa et al., 2003). Il a été montré que les espèces avec un fort pouvoir de dispersion présentent de faibles différenciations génétiques sur de grandes distances (Yu et al., 1999). En revanche, les espèces marines présentant de faibles

capacités de dispersion devraient présenter des aires de distribution restreintes et une forte différenciation génétique de leurs populations en lien avec la distance géographique, c'est-à-dire un patron d'isolement par la distance (Slatkin, 1993). Ce patron augmenterait le nombre d'opportunités de spéciation, principalement allopatrique, en favorisant les événements vicariants et la formation de lignées évolutives indépendantes au cours du temps (De Queiroz, 1998; Malay and Paulay, 2010; Paulay and Meyer, 2002).

Bien que le milieu marin soit parfois considéré comme un milieu complètement ouvert et sans entraves, les « open-waters », par exemple, peuvent représenter des barrières biogéographiques. L'importance relative des barrières environnementales (températures de surface, salinité, ...) s'opposant à la dispersion des organismes marins benthiques peut être estimée par le degré de connectivité entre les populations à différentes échelles géographiques.

Parmi les organismes marins, les scléactiniaires, de par leur rôle fondamental dans l'origine et la structure des récifs coralliens, présentent un intérêt tout particulier pour les études de conservation. Ainsi, décrire leur diversité spécifique et étudier la structure génétique de leurs populations est un élément clé dans la compréhension de ces écosystèmes.

3.2.1. Les modèles de structuration génétique

Par le passé, différents modèles de structuration des populations ont été décrits. Le modèle en îles (Wright, 1931) admet que, dans une population divisée en un nombre infini de groupes panmixtiques, les individus peuvent migrer d'un groupe à un autre avec la même probabilité, indépendamment de la distance qui les sépare. C'est un modèle très simple mais pas forcément réaliste ; en effet, il paraît peu probable que les individus aient la même probabilité de se disperser sur de longues distances que sur de petites distances. C'est pourquoi, en 1943, (Wright) introduit la notion de distance pour créer le modèle d'isolement par la distance (ou stepping-stone). Dans ce modèle, les migrants ont davantage de chance d'être issus de groupes proches. Ainsi, il est possible d'observer une plus forte différenciation entre des populations éloignées qu'entre des populations proches. Les modèles précités ne prennent pas en compte les problèmes de fragmentation d'habitats et considèrent chaque groupe d'individus isolés comme une population. Ainsi, le modèle en métapopulation a été proposé par (Levins, 1969, 1970) comme une *“population of populations that go extinct and recolonize”* (Levins, 1970). Ce modèle suggère que les individus sont organisés en populations plus ou moins isolées les unes des autres ayant chacune leur dynamique propre. Cependant, chaque population reste

dépendante des autres car quelques migrants permettent de connecter l'ensemble. Les populations qui viendraient à disparaître pourront être recolonisées par les migrants venant d'autres populations (Grimm et al., 2003; Hanski and Gilpin, 1991). Ainsi, une métapopulation peut se maintenir dans le temps bien que des extinctions locales surviennent.

La génétique des populations repose sur le modèle d'équilibre d'Hardy-Weinberg (Hardy-Weinberg Equilibrium, HWE) qui postule qu'au sein d'une population (idéale), les fréquences alléliques et génotypiques restent inchangées d'une génération à l'autre. Pour que cet équilibre existe, plusieurs hypothèses sont décrites concernant les populations : (1) elles ne sont pas soumises aux forces évolutives (sélection, migration, mutation, dérive), (2) elles sont de tailles infinies, (3) le croisement des individus se fait au hasard (pas de choix des partenaires, pas de structuration dans l'espace), (4) les individus sont diploïdes et la reproduction est sexuée et (5) les générations ne sont pas chevauchantes.

Afin d'étudier la structure génétique des populations, des estimateurs ont été décrits et permettent de mesurer la différenciation génétique entre populations.

3.2.2. Les mesures de la différenciation génétique

Au niveau inter-populationnel, la différenciation génétique peut être estimée grâce au F_{ST} ou indice de fixation (Weir and Cockerham, 1984) qui exprime la diminution de l'hétérozygotie d'une sous-population et permet de mesurer l'écart à l'équilibre de Hardy-Weinberg dû à la différenciation des sous-populations par rapport à une population totale panmictique. Cet indice varie entre 0 et 1 ; cela représente un continuum entre une absence de connectivité (populations fermées avec principalement de l'autorecrutement ; $F_{ST} = 1$) et une forte connectivité (populations ouvertes avec un fort taux d'allo-recrutement ; $F_{ST} = 0$). Ainsi, le cycle de vie et les stratégies reproductrices sont des caractéristiques spécifiques importantes qui vont moduler la différenciation des populations et, par conséquent, leur degré de connectivité. La différenciation des populations peut également être estimée par un test exact de Fisher dont l'hypothèse nulle est que la distribution de k génotypes parmi r populations est aléatoire (Raymond and Rousset, 1995). Il permet de tester si le déséquilibre à Hardy-Weinberg observé est lié à une variation des fréquences alléliques entre populations. L'analyse de variance moléculaire (AMOVA) est basée sur les analyses de variances des fréquences alléliques. Elle permet de définir le niveau d'organisation pour lequel la variabilité génétique est la plus forte (île, population ou individu).

Lorsque la différenciation génétique entre les populations est estimée, il est possible de la comparer avec la distance géographique pour tester l'isolement par la distance. Ainsi, la corrélation entre une matrice de distances géographiques [par exemple, avec une échelle logarithmique : $\log_{10}(\text{distances en km})$] et une matrice de distances génétiques [distance linéarisée de Slatkin : $(F_{ST}/(1 - F_{ST}))$] peut être estimée par un test de Mantel (Legendre, 2000; Mantel, 1967).

Par ailleurs, les tests d'assignement bayésiens (STRUCTURE ; Pritchard et al., 2000), comme vu précédemment (partie 2.3.6), permettent d'estimer la structuration génétique des populations en créant des groupes génétiques à l'équilibre d'Hardy-Weinberg. De la même façon, les analyses multivariées (DAPC ; (Jombart, 2008)) permettent de grouper les individus en fonction de leur proximité génétique sans faire d'hypothèses sur les marqueurs ni sur les populations, contrairement à STRUCTURE.

Au niveau intra-populationnel, la différenciation génétique est mesurée à partir du F_{IS} ou indice de consanguinité (Wright, 1931). Il varie entre -1 et 1 : une valeur négative exprime un excès d'hétérozygotes ; une valeur positive exprime un déficit en hétérozygotes et une valeur nulle, l'équilibre d'Hardy-Weinberg. Une consanguinité élevée chez une espèce à reproduction strictement sexuée, révélée par un déficit en hétérozygotes par rapport à une population conforme à HWE (panmixie), indique que le régime de reproduction est fermé, et que potentiellement la reproduction se fait préférentiellement entre individus apparentés, sans aucune action des autres forces évolutives.

Ainsi, la reproduction et les capacités dispersives, caractéristiques intrinsèques des organismes, peuvent représenter des facteurs structurant des populations. Certains cas particuliers, comme la reproduction asexuée voire même la reproduction clonale auront potentiellement un impact extrêmement important sur la structure des populations.

3.3. La connectivité des populations clonales

La reproduction asexuée présente l'intérêt de propager très rapidement des génotypes extrêmement bien adaptés à l'environnement et permet aux individus d'économiser leur énergie en évitant la production de gamètes ou de larves. Contrairement à la reproduction sexuée qui est théoriquement bénéfique bien que coûteuse (Crow, 1999), la reproduction asexuée n'offre pas l'opportunité de produire des génotypes recombinants qui pourraient être avantageux (Fisher, 1930; Muller, 1932), d'éliminer les mutations délétères (Kondrashov, 1988; Muller,

1964) et de s'adapter aux changements environnementaux en augmentant la diversité génétique des populations (Crow, 1994). Ainsi, pour persister à long terme, il est nécessaire que des lignées asexuées puissent occasionnellement se recombiner génétiquement, grâce à des processus de brassages intra- et inter-chromosomiques, afin de contrer les désavantages de ne pas se reproduire sexuellement. Sous cette considération, le meilleur modèle théorique correspondrait à une stratégie mixte. En effet, peu d'espèces sont exclusivement asexuées [par exemple, les rotifères bdelloïdes qui ont développé des mécanismes de réparation de l'ADN (Flot et al., 2013)] et plusieurs taxons présentent une alternance entre reproductions asexuée et sexuée (de Meeûs et al., 2007).

Cependant, dans le cas d'espèces présentant une reproduction en partie clonale ou une diversité génétique faible, les indices de mesure de la différenciation génétique seront biaisés par le fait que plusieurs individus, du fait de leur clonalité, présentent le même génotype multi-locus (MLG). Ainsi certains MLGs seront surreprésentés, ce qui biaisera les fréquences alléliques. C'est notamment le cas chez les coraux scléroustiniens, ceux-ci pouvant se reproduire par fragmentation ou par production de larves asexuées (*e.g.* Ayre and Hughes, 2000). En effet, la clonalité conduit à des corrélations fortes entre les allèles alors que la reproduction sexuée introduit de nouvelles combinaisons d'allèles. Néanmoins, par l'accumulation des mutations pour chaque allèle, la reproduction asexuée génère aussi des différences génétiques : l'effet Meselson (Birky, 1996; Welch and Meselson, 2000). Sur une échelle de temps courte, l'effet Meselson produit un excès d'hétérozygotes par rapport à ce qui est attendu sous l'hypothèse d'Hardy-Weinberg (Balloux et al., 2003; Stoeckel and Masson, 2014). L'étude des populations clonales soulève un autre problème : la production d'individus génétiquement identiques peut conduire à des populations ayant des diversités génétiques très faibles et l'utilisation des indices standards pour décrire les populations clonales en devient une expérience risquée. Ainsi, les statistiques traditionnelles utilisées en génétique des populations peuvent conduire à des interprétations erronées. La distribution de la fréquence des distances génétiques entre individus, définie comme le spectre de diversité génétique (*genetic diversity spectrum* GDS) a alors été beaucoup utilisée pour étudier les organismes clonaux (Arnaud-Haond et al., 2014; Moalic et al., 2011; Rozenfeld et al., 2007). Ainsi, ces analyses permettent la construction de réseaux présentant les relations génétiques entre les individus. Une des distances génétiques la plus communément utilisée est la distance des allèles partagés, qui suit implicitement le modèle infini d'allèles [*Infinite Allele Model*, IAM, (Lynch, 1990)]. Néanmoins, les marqueurs microsatellites sont plus susceptibles de varier d'un petit nombre de répétitions plutôt que d'un grand nombre à partir d'un allèle donné (Bruvo et al., 2004; Di Rienzo et al., 1994; Goldstein

et al., 1995; Rozenfeld et al., 2007). Le modèle d'évolution en pas japonais semble donc plus approprié [*Stepwise Mutation Model*, SMM, (Valdes et al., 1993)] car il prend en compte la différence du nombre de répétitions d'un motif microsatellite dans le calcul d'une distance entre deux allèles. Disposer d'un nombre de marqueurs important est donc nécessaire : plus le nombre de marqueurs disponibles sera élevé, plus il sera possible de déterminer si deux individus présentant le même MLG sont issus d'un même événement de reproduction et non d'un tirage aléatoire dans un pool d'allèles limité.

Parmi les scléactiniaires, les coraux du genre *Pocillopora* sont particulièrement intéressants car ils présentent une grande variété de modes de reproduction (sexuée, asexuée ou clonale). Ainsi, étudier le genre dans son ensemble pourrait permettre d'estimer les modes de reproduction de différentes espèces et savoir si celui-ci peut avoir un impact sur la structure des populations et la connectivité.

4. Le modèle d'étude : les scléactiniaires du genre *Pocillopora*

L'embranchement des cnidaires fait partie d'un des taxons présentant un panel extraordinaire de diversités d'espèces et d'écologies différentes et occupe une position basale dans l'évolution des métazoaires. Concernant les scléactiniaires (coraux durs), de nombreux fossiles ont été découverts et l'ensemble des découvertes couvrent une période de 240 millions d'années (Ma) d'histoire évolutive (Budd et al., 2010), datant leur apparition au cours du Trias pendant l'ère Mésozoïque. La phylogénie moléculaire confirme les hypothèses formulées à partir des observations de la structure du squelette calcaire quant à la polyphylie des organismes coralliens.

Une colonie corallienne est composée d'un ensemble d'individus, ou polypes, qui sont tous génétiquement identiques. Le polype est capable d'utiliser le carbonate de calcium (CaCO_3^{2-}) des océans pour synthétiser un squelette calcaire. Il est logé dans le squelette calcaire dans ce que l'on appelle un calice (Figure 3). Le calice est situé dans la partie extérieure du squelette, nommé exosquelette ou *corallum*. Chacune des loges individuelles s'étend en forme de tube sous le *corallum*, cette extension cylindrique est le corallite. Le polype corallien est zoophage, il se nourrit de zooplancton qu'il attrape la nuit à l'aide de cnidocystes. Cependant, le caractère oligotrophe des mers tropicales et, par conséquent, leur pauvreté en nutriments ainsi qu'en plancton ne permet pas aux polypes coralliens de couvrir tous leurs besoins énergétiques

(Muscatine and Porter, 1977). Ainsi pour assurer leur survie, les coraux ont développé une symbiose avec des dinoflagellés (algues unicellulaires) appartenant au genre *Symbiodinium*, appelées zooxanthelles (Muscatine and Porter, 1977; Rowan and Powers, 1991).

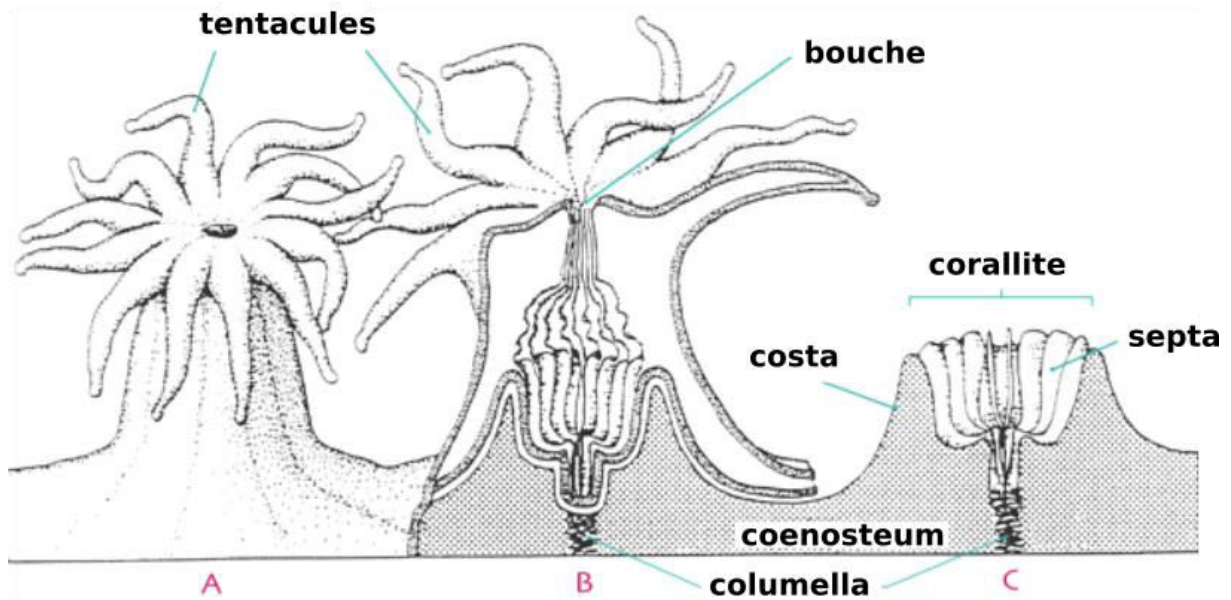


Figure 3 : Représentation schématique d'un squelette corallien. A Polype entier, B Coupe transversale du polype et de son squelette, C Squelette nu (parties vivantes retirées) – issue de (Mather and Bennett, 1993)

Les scléactiniaires, étant des organismes sessiles, ont développé des modes de reproduction et de dispersion variés.

4.1. Les scléactiniaires, reproduction et dispersion

La reproduction des scléactiniaires peut se définir par son régime (sexué ou asexué), son mode (*broadcast-spawners* ou *brooders*) et son type (colonies gonochoriques ou hermaphrodites).

4.1.1. La reproduction sexuée

Étant donné le caractère sessile de ces organismes, ceux-ci n'ont pas de comportement actif de recherche de partenaire sexuel pour assurer leur reproduction. Ainsi, les gamètes sont émis dans la colonne d'eau. Il existe deux grands modes de reproduction des coraux : les *brooders* et les *broadcast-spawners*. Chez les coraux dits *brooders*, seuls les gamètes mâles

sont émis dans la colonne d'eau. Ceux-ci dérivent dans la colonne jusqu'à la rencontre avec un ovule (à l'intérieur des polypes de la colonie). En revanche, chez les coraux dits *broadcast-spawners*, les gamètes mâles et femelles sont émis en même temps et la fécondation se produit dans la colonne d'eau (Gleason and Hofmann, 2011; Harrison, 2011). Pour que la reproduction sexuée soit efficace chez les *broadcast-spawners*, toutes les colonies d'une même espèce doivent émettre leurs gamètes en même temps ; les chances de fécondation croisée sont ainsi augmentées. En effet, si d'autres espèces émettent leurs gamètes de manière synchrone (exemple *Goniopora* et *Pocillopora* ; Suzuki, 2012) et qu'il n'existe pas de barrières à la fécondation entre espèces différentes, la fécondation de gamètes d'espèces différentes devient possible.

Que les colonies soient *brooders* ou *broadcast-spawners*, la fécondation entraîne la formation d'un embryon qui se développera en larve planula. Chez les *brooders*, la planula est émise dans la colonne d'eau lorsqu'elle a achevé son développement alors que chez les *broadcast-spawners*, tout le développement jusqu'à la formation de la larve planula a lieu dans la colonne d'eau. La planula, une fois formée, va rejoindre le substrat pour s'y installer, on parle alors de recrutement. La mortalité des larves est très élevée pendant les quelques jours qui suivent leur émission (McCormick, 2003), mais le risque de prédation reste important tout au long de la vie larvaire. En effet, avant de pouvoir s'installer sur un récif, les larves rencontrent un "mur de bouches" (c'est-à-dire de nombreux prédateurs tels que les poissons ou les invertébrés filtreurs) qui entraîne une très forte mortalité (Fabricius and Metzner, 2004). Chez les coraux, un neuropeptide induit la métamorphose des larves et, par conséquent, leur installation sur le récif (Hirose et al., 2007). La métamorphose des larves est souvent induite par la présence d'algues corallines encroûtantes sur le substrat (Heyward and Negri, 1999). Si la planula parvient à se fixer, elle se métamorphose alors en polype primaire qui va se développer à son tour pour former une colonie juvénile (Figure 4).

Les capacités de dispersion dépendent, en partie, du mode de développement des larves. Intuitivement, il est admis que les larves dont le cycle de développement se déroule exclusivement dans la colonne d'eau (*broadcast-spawners*) passeront plus de temps dans celle-ci et auront donc davantage de temps pour leur dispersion. Au contraire, les larves ayant été incubées seront compétentes plus rapidement et pourront donc se fixer plus rapidement sur le substrat. Cependant, différentes études ont révélé des résultats et des théories contradictoires.

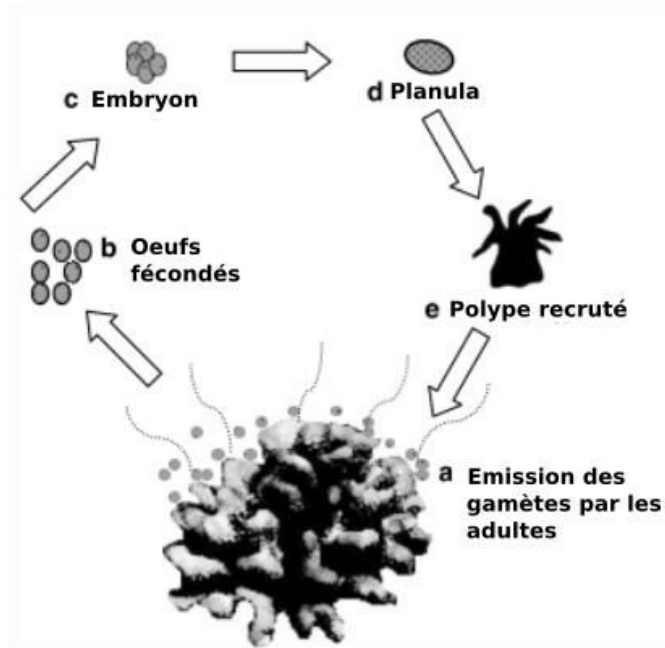


Figure 4. Représentation schématique du cycle de vie d'un corail. Les adultes émettent les gamètes dans la colonne d'eau (a). Les oocytes sont fécondés par les spermatozoïdes (b) ce qui conduit à la formation d'un embryon (c) qui va se développer en larve planula (d). Lorsqu'elle rencontre des conditions favorables, la planula se métamorphose en polype qui s'installe sur le récif (recrutement) pour former une nouvelle colonie (adapté de Apprill et al., 2009).

En effet, Fadlallah (1983) a montré que les larves incubées étaient plus grosses que les larves formées dans la colonne d'eau, suggérant que leur pouvoir de dispersion est plus important car leurs réserves leur permettraient de survivre plus longtemps dans la colonne d'eau. En outre, Isomura and Nishihira (2001) ont révélé que la taille des larves était variable en fonction des colonies et que plus une colonie est grosse, plus la larve est grosse et plus la durée de vie de la planula est longue et par conséquent plus ses capacités de dispersion sont grandes. Par ailleurs, concernant le recrutement des larves, plusieurs hypothèses ont été avancées. Ainsi, les larves incubées auraient tendance à recruter près de leurs parents d'après Ayre and Miller (2004) alors que Sherman et al. (2005) trouvent que le recrutement des larves incubées est rarement couronné de succès près des parents.

En plus de modes de reproduction différents, deux types de colonies peuvent également être trouvées. Les colonies peuvent être (1) de type gonochorique, c'est-à-dire que tous les polypes d'une colonie présentent le même type sexuel, soit mâle, soit femelle, ou (2) de type hermaphrodite, c'est-à-dire que pour une même colonie, on peut trouver des polypes mâles et des polypes femelles en même temps (hermaphrodisme simultané). L'hermaphrodisme peut également être séquentiel, c'est-à-dire que tous les polypes d'une colonie peuvent changer de

sexe au cours du temps [voir Baird et al., 2009; Gleason and Hofmann, 2011; Harrison, 2011; Richmond and Hunter, 1990, pour des revues sur la reproduction chez les coraux].

Parmi les combinaisons possibles de types et de modes de reproduction, les *broadcast-spawners* hermaphrodites sont les plus communs [58 %, recensés pour 444 espèces dont la reproduction a été décrite Harrison (2011)], suggérant que les phénomènes d'autofécondation pourraient être courants. Les *broadcast-spawners* gonochoriques correspondraient à 18 % des espèces et les *brooders*, qu'ils soient gonochoriques ou hermaphrodites ne représentent que 9 % des espèces. Des tests *in vitro* ont révélé que les coraux hermaphrodites simultanés sont, selon les espèces, complètement ou partiellement auto-fertiles, bien que l'autofécondation paraît peu fréquente (Carlson, 1999; Hatta et al., 1999; Heyward and Babcock, 1986; Knowlton et al., 1997; Miller and Mundy, 2005; Willis et al., 1997). Les phénomènes d'autofécondation ou de reproduction avec des individus apparentés auront des conséquences sur la structuration génétique des populations.

4.1.2. La reproduction asexuée

Outre la reproduction sexuée, les scléactiniaires, comme de nombreux autres cnidaires, sont également capables de se reproduire de façon asexuée.

Les scléactiniaires sont capables de se reproduire de façon végétative selon différentes méthodes : par fragmentation (Highsmith, 1982), par bourgeonnement ou expulsion d'un polype (Sammarco, 1982), ou par l'émission de larves produites de façon asexuée (larves parthénogénétiques). La production de larves parthénogénétiques a cependant été mise en évidence chez peu de scléactiniaires comme, par exemple, chez les coraux des genres *Pocillopora* (Miller and Ayre, 2004; Stoddart, 1983) et *Tubastrea* (Ayre and Resing, 1986). La production de larves parthénogénétiques a été décrite comme étant un mode de reproduction alternatif pour les coraux, notamment lorsque les taux de fécondation sont faibles à cause de la limitation en gamètes mâles (Yund, 2000).

Les individus issus de la reproduction asexuée sont potentiellement strictement identiques génétiquement (clones). Par conséquent, il est important d'identifier les clones ou genets (c'est-à-dire, l'ensemble des individus qui partagent le même génotype multi-locus) dans les populations échantillonnées. En effet, pour estimer ensuite la connectivité des populations, il est important d'estimer le plus justement possible les fréquences alléliques. La présence de clones biaise ces fréquences alléliques.

Les scléactiniaires font partie des groupes représentant un défi taxonomique. Historiquement, la morphologie était utilisée comme premier caractère pour les études systématiques et taxonomiques (combinant caractères morphologiques de la colonie et la microstructure du squelette calcaire et décrivant des morpho-espèces), alors que la morphologie peut être extrêmement variable et est probablement sous pression sélective. Parmi les scléactiniaires, les coraux du genre *Pocillopora* ont été largement étudiés et représentent un véritable défi taxonomique par leur extrême plasticité phénotypique (e.g. Pinzón and LaJeunesse, 2011; Veron and Pichon, 1976a).

4.2. Le cas des coraux du genre *Pocillopora*

Les coraux du genre *Pocillopora* sont des espèces branchues pionnières, c'est-à-dire qu'elles font partie des premières espèces à s'installer sur les récifs (Grigg, 1983). Ils sont très abondants dans tout l'Indo-Pacifique, répartis sur tous les récifs de la mer Rouge, et des océans Indien et Pacifique (Figure 5). Park et al. (2012) ont mis en évidence que l'ancêtre commun aux Pocilloporidae serait daté du Paléogène (environ 66 Ma) au sein desquels l'apparition du genre *Pocillopora* daterait de la fin du Néogène (environ 1.8 Ma). Ce genre serait plus jeune que tous les autres genres de cette famille (*Madracis*, *Seriatopora* et *Stylophora* ; Park et al., 2012).

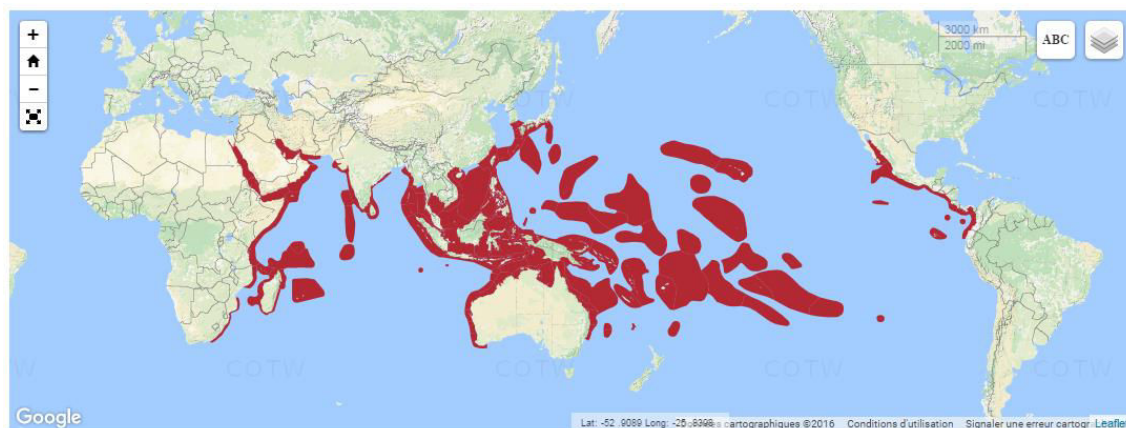


Figure 5 : Carte de répartition des coraux du genre *Pocillopora*. Carte extraite du site Corals of the world : <http://www.coralsoftheworld.org/>

4.2.1. *Pocillopora* : morphologie et phylogénie

La multiplication des études combinant phylogénie ont révélé que les chez les coraux du genre *Pocillopora*, la morphologie et la phylogénie manquent de congruence. Ce phénomène a

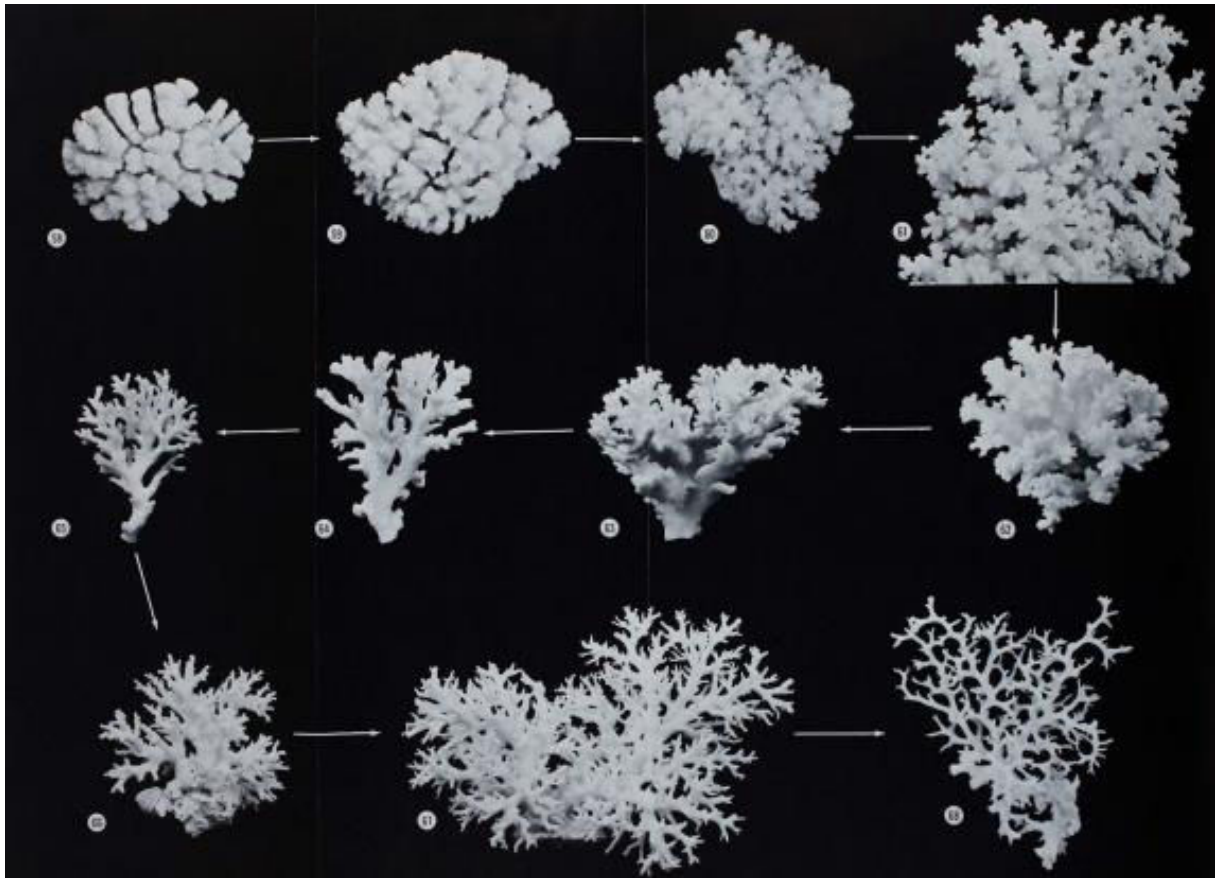


Figure 6. Écomorphes de *Pocillopora damicornis* (issue de Veron and Pichon, 1976a)

été pr esenti depuis plusieurs d ecennies puisqu'en 1976, Veron and Pichon (ont commenc e   regrouper diff erentes morphologies sous une m eme esp ece, d efinissant des  ecomorphes (variations morphologiques des colonies en lien avec l'environnement) (Figure 6). En 1995, (Veron)  met l'hypoth ese que les coraux sont en fait des « syngameons » ou « m eta-esp eces » repr esentant un regroupement d'organismes g en etiquement li es qui peuvent  tre morphologiquement similaires ou non et pouvant appartenir   des genres diff erents.

Cette derni ere d ecennie, des  tudes ont commenc e   explorer les limites d'esp eces du genre *Pocillopora*   l'aide de donn ees mol eculaires. En 2006, (Flot and Tillier) ont estim e les relations phylog en etiques de 37 colonies  chantillonn ees   Hawaii et appartenant   cinq morpho-esp eces (*P. meandrina*, *P. eydouxi*, *P. ligulata*, *P. damicornis* et *P. molokensis*) en s equenant l'ITS (Internal Transcribed Spacer). La phylog enie  tait congruente avec la monophylie de *P. ligulata* et *P. eydouxi*, mais *P. meandrina*, *P. damicornis* et *P. molokensis* repr esentaient des groupes polyphyl etiques. Par la suite, les m emes 37 colonies ont  t e  tudi ees

en ajoutant des marqueurs mitochondriaux et nucléaires (Flot et al., 2008). Alors que les reconstructions phylogénétiques identifient cinq lignées mitochondriales distinctes en accord avec la morphologie, *P. eydouxi* et *P. meandrina* sont résolus uniquement grâce au marqueur nucléaire ITS2. Cette observation a été attribuée à la nature du génome mitochondrial des coraux qui évolue lentement comparé à celui d'autres métazoaires (Hellberg, 2006; Shearer et al., 2002). Quelques années plus tard, (Schmidt-Roach et al., 2013) se sont penchés sur les différents écomorphes de *P. damicornis* décrits par Veron and Pichon (1976). Un réseau reconstruit à partir des séquences mitochondriales de l'Open Reading Frame (ORF) a révélé cinq lignées distinctes, appelées *P. damicornis* type α , β , γ , δ et ϵ (Schmidt-Roach et al., 2013). La même année, Pinzón et al. (2013) ont échantillonné 906 colonies représentant une grande variété de morphes de *Pocillopora*, et ce en couvrant l'aire de distribution du genre. En séquençant le marqueur mitochondrial ORF, ils ont identifié 21 haplotypes, qui peuvent être regroupés entre cinq et huit lignées génétiquement distinctes, le plus souvent sans lien avec la morphologie. Dans une autre étude, Marti-Puig et al. (2014) ont examiné les photographies *in situ*, les images de microscopie électronique et les données moléculaires (ORF, Dloop, ITS2) pour 59 colonies de *Pocillopora*, présentant une large gamme de morphologies atypiques supposées rares ou endémiques. Les marqueurs mitochondriaux (ORF et Dloop) ont permis de distinguer six clades distincts, alors que le marqueur nucléaire (ITS2) n'a pas permis de résoudre des groupes clairs en partie à cause des variations intra-génomiques de cette région. La forme générale de la colonie est extrêmement variable au sein de chaque clade (par exemple, le clade IIb est composé de colonies pouvant être attribuées aux morpho-espèces *P. damicornis*, *P. capitata*, *P. elegans*, *P. eydouxi*, *P. zelli* ou encore *P. molokensis*) alors que le niveau microscopique montre moins de variabilité. Ensuite, Schmidt-Roach et al. (2014b) ont étudié 78 colonies de *Pocillopora* de la côte Est de l'Australie en séquençant l'ORF et en analysant 10 caractères morphométriques de la forme générale de la colonie et de la microstructure des calices. Les clades de leur arbre phylogénétique sont congruents avec la forme générale des colonies, mais les variations de microstructure, en particulier la forme et le type de columelle, permettent de différencier les quatre clades identifiés. Le clade 1 est composé de *P. damicornis* type α (rebaptisé *P. damicornis*), *P. damicornis* type β (rebaptisé *P. acuta*) et *P. damicornis* type δ . Le clade 2 correspond à *P. verrucosa* et une nouvelle espèce que les auteurs ont appelée *P. bairdi*. Le clade 3 est composé de colonies de *P. meandrina* et *P. eydouxi* et le clade 4 de *P. damicornis* type ϵ (décrite sous le nom de *P. brevicornis*).

Toutes ces études montrent que (1) l'apport des données génétiques est crucial pour la taxonomie du genre *Pocillopora*, certaines morpho-espèces cachant une forte diversité

génétique, et (2) qu'il est nécessaire d'explorer cette diversité sur la totalité de l'aire de distribution, même aux marges, pour confirmer ou infirmer le statut des espèces et leur endémisme. Cependant, il est à noter que toutes ces études se sont basées uniquement sur des reconstructions phylogénétiques et sur la robustesse des nœuds pour déterminer les clades qui constituaient des espèces.

4.2.2. Ce qui est connu de la reproduction chez *Pocillopora*

Les coraux du genre *Pocillopora* sont parmi les plus étudiés des scléactiniaires et en particulier, *P. damicornis*, le rat de laboratoire, a été étudié sous toutes les coutures. Cependant, à la lumière des résultats réévaluant le statut des espèces de *Pocillopora*, les résultats précédemment trouvés doivent être reconsidérés. Historiquement, le corail *P. damicornis* était connu pour posséder un large éventail de modes de reproduction. Aussi, à partir de leurs observations, Schmidt-Roach et al. (2013) ont pu résoudre une partie du puzzle et attribuer un mode de reproduction aux différentes lignées mitochondriales qu'ils ont identifiées et de réévaluer certaines études historiques. Ainsi, *P. damicornis* type α semble incuber ses larves (*brooder*) et les émettre après la pleine lune (type C et T Muir, 1984; type brown Richmond and Jokiel, 1984; Schmidt-Roach et al., 2013; Tanner, 1996) tandis que les types β et δ les libèrent après la nouvelle lune (type G ; Muir, 1984, type *yellow* ; Richmond and Jokiel, 1984 ; Schmidt-Roach et al., 2013). L'émission de larves issues de reproduction asexuée (larves parthénogénétiques) a été mise en évidence à la fois pour *P. damicornis* types α et type β . En outre, de nombreuses études rapportent des cas de reproduction clonale chez *P. damicornis* type β (par exemple, (Adjeroud et al., 2014; Souter et al., 2009; Torda et al., 2013a)). *Pocillopora damicornis* type γ est décrit comme proche de *P. verrucosa*, généralement décrit comme une espèce hermaphrodite simultanée émettrice de gamètes mâles et femelles avec fécondation externe (*broadcast-spawner*) [Afrique du Sud, Kruger and Schleyer (1998) ; Maldives, Sier and Olive (1994)]. En revanche, cette espèce a été décrite comme *brooder*, émettant ses larves après la nouvelle lune aux îles Marshall (Stimson, 1978) et aux Philippines (Villanueva et al., 2008), suggérant qu'il ait été confondu avec *P. damicornis* type β . Par ailleurs, à Hawaii, une émission de gamètes mâles a été observée pour *P. damicornis* type β et *P. meandrina* (Schmidt-Roach et al., 2014a). Cependant, comme aucun œuf n'a été observé chez *P. meandrina*, il n'est pas possible de savoir s'il est *brooder* ou *broadcast-spawner*. En revanche, *P. eydouxi* a été décrit comme un *broadcast-spawner* puisqu'il émet des gamètes

mâles et femelles (Hirose et al., 2001). A ce jour, aucune donnée n'est disponible sur le mode de reproduction du type ε (*P. brevicornis*).

Ces différentes études révèlent une grande variabilité dans les modes de reproduction, qui pourront influencer la connectivité et la structure génétique des populations des différentes espèces.

4.2.3. La connectivité des *Pocillopora* à travers le monde

Différentes études se sont intéressées à la structuration des populations et à la connectivité chez *Pocillopora* à travers le monde, cependant la plupart des études identifiaient les espèces sur la base de la morphologie des colonies. En prenant en compte le caractère imparfait de la morphologie pour décrire ces espèces, il est important de considérer les résultats de ces études avec précaution.

Concernant *P. damicornis*, plusieurs études ont estimé sa connectivité. Cependant, pour beaucoup d'entre elles, il n'est pas possible de savoir de quel type il s'agit. Ainsi, *P. damicornis sensu lato* a été étudié dans le Pacifique Tropical Est à l'aide de six locus microsatellites (Combosch and Vollmer, 2011). Cette étude a révélé que les flux de gènes étaient restreints entre les populations mais que la diversité génétique était élevée au sein de chaque population, suggérant un potentiel adaptatif important pour chacune d'entre elles (Combosch and Vollmer, 2011). En effet, aucune différenciation génétique n'a été trouvée entre les populations sur de grandes échelles, alors que, sur de faibles distances, la différenciation est forte entre les populations. A l'inverse, en étudiant des populations de *P. damicornis sensu lato* sur de petites distances (1,2 km) (Gorospe and Karl, 2013) n'ont pas trouvé de différenciation génétique. Ces derniers pourraient concorder avec un patron d'isolement par la distance, comme l'ont suggéré Paz-García et al. (2012) en trouvant une relation significative entre la différenciation génétique et la distance géographique.

Pour d'autres études en revanche, le type de *P. damicornis* a pu être identifié. Ainsi, en ce qui concerne *P. damicornis* type β , il a été montré que la différenciation génétique (en utilisant neuf microsatellites) était plus grande entre les populations de récifs différents qu'au sein d'un même récif (les deux sites les plus éloignés étant séparés de plus de 1 100 km) Torda et al. (2013a). Des résultats similaires ont été mis en évidence par Adjeroud et al. (2014) en Polynésie Française. En effet, des populations séparées de plus de 250 km sont génétiquement

différenciées alors que deux populations situées sur la même île et séparées de 13 km ne le sont pas. Étant donné que *P. damicornis* type β est capable de se reproduire par propagation clonale, ces derniers ont également estimé la différenciation génétique en ne gardant qu'un seul représentant de chaque génotype et ils trouvent que des populations séparées de plus de 280 km ne sont plus génétiquement différenciées alors que deux populations séparées de 18 km le deviennent. Ces résultats soulignent l'importance de considérer la clonalité dans les estimations de différenciation des populations. L'étude de (Pinzón et al., 2013) est la seule ayant essayé de couvrir l'ensemble de l'aire de distribution de cette espèce en utilisant cinq locus microsattellites. A partir de huit populations (dans le Pacifique central : Hawaii, Lizard Island and Heron Island ; dans la région Indo-Pacifique: Taïwan, Palau et le Golf de Thaïlande; et dans le Sud-Ouest de l'océan Indien : Maurice et Madagascar), ils ont pu identifier des niveaux élevés de différenciation génétique entre toutes les paires de populations (sauf entre Taiwan et Lizard Island) ce qui suggère un flux de gènes réduit sur de très grandes distances et semble confirmer un modèle d'isolement par la distance pour cette espèce.

Une seule étude a pu être attribuée à *P. damicornis* type α Torda et al. (2013a). Cette étude semble montrer une structuration différente pour ces espèces. En effet, pour cette espèce, de manière générale une différenciation génétique faible a été révélée sur l'ensemble de la zone étudiée sur la Grand Barrière de Corail en Australie (environ 1 000 km entre les sites les plus éloignés).

Concernant les autres morpho-espèces de *Pocillopora*, les études sont moins florissantes. Une étude menée en Polynésie Française, sur *P. meandrina* grâce à des marqueurs microsattellites a révélé une faible différenciation à grande échelle (2 000 km) alors que des populations plus proches (5-10 km) présentaient plus de différenciation (Magalon et al., 2005). En outre, Ridgway et al. (2001) ont utilisé six allozymes polymorphes pour déterminer la structure génétique des populations du corail *Pocillopora verrucosa* dans six récifs de l'est de l'Afrique du Sud et n'ont pas trouvé de différenciation génétique entre les sites. Une étude similaire, utilisant des locus microsattellites, a été menée sur la même morpho-espèce dans la même zone : une différenciation génétique a été trouvée entre le nord et le sud de la zone, indiquant une faible connectivité entre les populations (Ridgway et al., 2008). Sur la côte est de l'Afrique (en Tanzanie et au Kenya), une étude (attribuée à *P. verrucosa a posteriori*) montre que les populations éloignées (environ 700 km) sont moins différenciées que des populations proches (moins de 10 km) en utilisant six microsattellites (Souter et al., 2009). *Pocillopora verrucosa* (type 3 Pinzón et al., 2013) a été étudié sur de grandes distances (de Panama à Madagascar). Ainsi, en utilisant cinq locus microsattellites ils ont révélé une différenciation génétique

importante, sauf entre les populations de l'Indo-Pacifique Central (Taïwan, Palau et Lizard Island).

Ces résultats montrent que les schémas de structuration peuvent être différents en fonction des espèces et qu'il est fondamental de bien identifier les espèces *a priori* afin de pouvoir comparer les études entre elles.

5. Problématique

L'étude de la diversité des coraux scléactiniaires et de la connectivité génétique des récifs coralliens est un enjeu majeur pour comprendre leur fonctionnement et notamment, leur capacité de résilience après des perturbations anthropiques ou naturelles.

Les coraux de genre *Pocillopora* sont particulièrement utiles dans cette optique car ce sont des coraux pionniers et sont trouvés sur tous les récifs de la frange tropicale (sauf dans l'océan Atlantique). Cependant, étant donné les récentes études révélant les incongruences entre la morphologie et les données génétiques chez ces coraux (Flot et al., 2008; Flot and Tillier, 2006; Marti-Puig et al., 2014; Pinzón et al., 2013; Schmidt-Roach et al., 2013; Schmidt-Roach et al., 2014b), plusieurs questions se posent : Quelle est la diversité spécifique des *Pocillopora* ? Comment se reproduisent-ils ? Comment les populations se connectent-elles ?

Ainsi, le premier objectif de ce travail (présenté dans le Chapitre 1) sera de délimiter des hypothèses d'espèces robustes au sein du genre *Pocillopora*. Ainsi, en combinant des analyses de délimitation d'espèces sur données moléculaires (ABGD, GMYC, PTP, haplowebs) pour différents marqueurs (ORF, Dloop et ITS), tout en incluant des données déjà publiées et les données nouvellement récoltées au cours de ce travail provenant de zones encore sous étudiées (principalement le Sud-Ouest de l'océan Indien, le Pacifique Sud-Ouest et plus anecdotiquement la Polynésie Française), les hypothèses primaires d'espèces seront définies. Par la suite, ces données seront confrontées aux résultats de tests d'assignement (inférence bayésienne et analyse discriminante) pour définir les hypothèses d'espèces secondaires. Ce travail permettra de réévaluer le statut de certaines morpho-espèces et réévaluer les aires de distributions et les zones d'endémisme.

Une fois les hypothèses d'espèces secondaires définies, il sera possible d'étudier la structure des populations pour différentes hypothèses d'espèces, ce qui représente le deuxième objectif de ce travail de thèse, en s'intéressant particulièrement à la SSH *Pocillopora damicornis* type β . En effet, celle-ci ayant été décrite comme étant capable de se reproduire de façon asexuée,

cet aspect sera traité dans le Chapitre 2 en étudiant notamment sa propagation clonale et son impact sur la structure des populations autour de La Réunion.

Par la suite, la structure des populations de cette espèce sera estimée sur de grandes échelles (Chapitre 3) grâce à des populations échantillonnées de façon stratifiée dans le Sud-Ouest de l'océan Indien (Mayotte, Juan de Nova, Madagascar, Réunion et Rodrigues) et dans le Pacifique Sud-Ouest (Nouvelle-Calédonie). Ce chapitre apportera un éclairage sur les capacités de dispersion de cette espèce et sur les flux de gènes existants entre les populations échantillonnées.

Le dernier objectif portera son intérêt sur une autre hypothèse d'espèce, *P. eydouxi* (Chapitre 4). Cette dernière a été très peu étudiée par le passé et très peu de choses sont connues concernant sa biologie et son écologie. Ainsi, en étudiant la structure des populations à grande échelle, il sera possible de mieux comprendre cette espèce, notamment en étudiant sa reproduction (clonale ou non) et la connectivité en estimant les flux de gènes à différentes échelles à partir d'un échantillonnage stratifié mené dans le Sud-Ouest de l'océan Indien et dans le Pacifique Sud-Ouest.

CHAPITRE 1. Réévaluation du nombre d'espèces, de la distribution et de l'endémisme chez les coraux du genre *Pocillopora* Lamarck, 1816 grâce à l'utilisation des méthodes de délimitation d'espèces et des microsatellites

Résumé

Les méthodes de délimitation d'espèces basées sur les données génétiques permettent d'augmenter de façon considérable les taux de description de la biodiversité, mais elles peuvent également être utilisées pour aider à la clarification des cas de taxonomies complexes. Ainsi, au cours de cette étude, la diversité spécifique des scléactiniaires du genre *Pocillopora* a été explorée. Ces coraux présentent l'avantage d'être largement répartis sur toute la ceinture tropicale des océans Indien et Pacifique. Ainsi, à partir de 943 colonies de *Pocillopora* présentant une grande variété de morphes et ayant été échantillonnées dans trois provinces biogéographiques (le Sud-Ouest de l'océan Indien, le Sud-Ouest du Pacifique Tropical et le Sud-Est de la Polynésie), des hypothèses d'espèces primaires (PSH) ont été délimitées à partir de deux marqueurs mitochondriaux (ORF et Dloop) en utilisant le *Automatic Barcode Gap Discovery method* (ABGD), le *Poisson tree processes algorithm* (PTP) et le *Generalized mixed Yule-coalescent model* (GMYC) ainsi qu'en construisant un haploweb à partir du marqueur nucléaire ITS2. Par la suite, les PSHs ont été confrontées aux résultats des tests d'assignement (STRUCTURE and DAPC), obtenus à partir de 13 microsatellites, pour délimiter les hypothèses d'espèces secondaires (SSH). L'ajout de séquences de la littérature et la comparaison des différentes méthodes ont permis de délimiter au moins 18 hypothèses d'espèces secondaires au sein de 14 morphes, confirmant le caractère non diagnostique de la morphologie du corallum pour la définition des espèces chez le genre *Pocillopora* ainsi que la présence de lignées cryptiques. En effet, au cours de ce travail, il a été montré que plusieurs morphes peuvent correspondre à une SSH et *vice versa*. Par ailleurs, trois lignées génétiques non encore décrites ont pu être identifiées, supposant qu'il pourrait s'agir de trois nouvelles hypothèses d'espèces. En outre, la distribution géographique de certaines SSHs a pu être réévaluée grâce à ces nouvelles données génétiques ce qui pourrait avoir une implication directe dans les mesures de conservation.

Les résultats de ce chapitre ont permis la rédaction d'un article qui a été soumis au journal *Molecular Phylogenetics and Evolution*.

Reevaluating species number, distribution and endemism of the coral genus *Pocillopora* Lamarck, 1816 using species delimitation methods and microsatellites

Gélin P¹, Postaire B^{1,2}, Fauvelot C³, Magalon H¹

¹ UMR ENTROPIE (Université de La Réunion, IRD, CNRS), Laboratoire d'excellence-CORAIL; Faculté des Sciences et Technologies, 15 Bd René Cassin, CS 92003, 97744 St Denis Cedex 09, La Réunion, France

² present address: IMBE UMR 7263, Aix Marseille Université-CNRS-IRD-Avignon Université, Station marine d'Endoume, Chemin de la batterie des lions, 13007, Marseille, France.

³ UMR ENTROPIE (IRD, Université de La Réunion, CNRS), Laboratoire d'excellence-CORAIL; centre IRD de Nouméa, 101 Promenade Roger Laroque, BP A5, 98848 Nouméa cedex, New Caledonia

Corresponding author:

Hélène MAGALON, 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.

Email: helene.magalon@univ-reunion.fr

Key words: microsatellites, ORF, scleractinian, ABGD, GMYC, PTP

Abstract

Species delimitation methods based on genetic information, notably using single locus data, have been proposed as means of increasing the rate of biodiversity description, but can also be used to clarify complex taxonomies. In this study, we explore the species diversity within the cnidarian genus *Pocillopora*, widely distributed in the tropical belt of the Indo-Pacific Ocean. From 943 *Pocillopora* colonies sampled in the Western Indian Ocean, the Tropical Southwestern Pacific and Southeast Polynesia, representing a huge variety of morphs, we delineated Primary Species Hypotheses (PSH) applying the Automatic Barcode Gap Discovery method, the Poisson tree processes algorithm and the Generalized mixed Yule-coalescent model on two mitochondrial markers (ORF and Dloop) and reconstructing a haploweb using the nuclear marker ITS2. Then, we confronted identified PSHs to the results of clustering analyses (STRUCTURE and DAPC) using 13 microsatellites to determine Secondary Species Hypotheses (SSH). Based on the congruence of all methods used and adding sequences from the literature, we defined at least 18 Secondary Species Hypotheses among 14 morphs, confirming the lack of confidence in *corallum* macromorphology to define *Pocillopora* species and the presence of cryptic lineages. Indeed, several morphs corresponded to one SSH, and *vice versa*. We also identified three new genetic lineages never found to date, which could represent three new putative species. Moreover, the biogeographical ranges of several SSHs were re-assessed in the light of genetic data, which may have direct implications in conservation policies. Next generation sequencing, combined with other parameters (i.e. microstructure, zooxanthellae identification, ecology even at a micro-scale, resistance and resilience ability to bleaching) will be the next step towards an integrative framework of *Pocillopora* taxonomy, which will have profound implications for ecological studies, such as studying biodiversity, response to global warming and symbiosis.

Introduction

During the last decades, the use of genetic tools and breakthroughs in sequencing and faster phylogenetic algorithms have led to discover an increasing number of cryptic and/or endemic lineages (Pfenninger and Schwenk, 2007), most of which are awaiting to be named and even more are to be discovered (Appeltans et al., 2012; Mora et al., 2011). Species are fundamental units, routinely used for analysis of biogeography, ecology, macro-evolution and conservation biology (Sites and Marshall, 2004). Yet, converting genetic lineages into robustly and formally named taxonomic entities remains challenging. Indeed, beside the ongoing debate regarding species concepts (De Queiroz, 2007), species delineation resides in the issue of choosing a reliable criterion (or a set of) to identify distinct lineages. Historically, species identification was based on morphological criteria which represent the “visible diversity”, but it is now mostly accepted that several approaches of species delimitation are to be combined to improve our understanding of ecological and evolutionary patterns and processes (Gompert et al., 2006; Moritz, 2002; Wiens, 2007), in other words conducting integrative taxonomic studies (Padiál et al., 2010), even in emblematic animals that we thought to be well-known (for an example in giraffes, see Fennessy et al., 2016).

Within the marine realm, species description in cnidarians remains a difficult task due to the paucity of morphological characters and phenotypic plasticity (reviewed in Todd, 2008). Nevertheless, taxa in Scleractinia have been traditionally identified according to discontinuities of skeleton morphological characters (*corallum* macromorphology and microstructure of corallites) (Vaughan and Wells, 1943; Veron and Pichon, 1976; Wallace, 1999). However, since evolution of reproductive, ecological and physiological traits does not always present morphological outcomes (McFadden et al., 2014; Paz-García et al., 2015), these morphological characters may not necessarily be informative to elucidate phylogenetic relationships and define species boundaries. Relying on this assumption, several families of scleractinians have been re-evaluated during the few past years combining morphological and molecular data (e.g. Benzoni et al., 2010; Flot et al., 2008; 2011; Nakajima et al., 2012; Warner et al., 2015), with several studies revealing cryptic lineages (see for examples, Hayashibara and Shimoike, 2002; Keshavmurthy et al., 2013; Schmidt-Roach et al., 2014; Stefani et al., 2011)

The genus *Pocillopora* Lamarck, 1816 represents one of the most studied groups among scleractinians. It is widely distributed in the Indo-Pacific and the Tropical Eastern Pacific (Veron, 2000), but absent from the Atlantic Ocean. Historically, species boundaries within this

genus were investigated based on morphological characters, such as the shape and the organization of branches and verrucae, leading to 17 described morpho-species (Veron, 2000). Because of the extreme morphological plasticity in this group, as well as the paucity of diagnostic morphological characters, this last decade, molecular approaches have been added to investigate species boundaries. Hereafter we summarized the history of *Pocillopora* genetic work performed on the different morpho-species in view of species delimitation in this genus. First, in 2006, Flot and Tillier were the first to use molecular tools to investigate phylogenetic relationships among different morpho-species from Hawaii using ITS, revealing the monophyly of *P. ligulata* Dana, 1846 and *P. eydouxi* Milne Edwards, 1860 (accepted as *Pocillopora grandis* Dana, 1846) while the Hawaiian *P. damicornis* (Linnaeus, 1758), *P. meandrina* Dana, 1846 and *P. molokensis* Vaughan, 1907, appeared non-monophyletic. Later on, adding mitochondrial markers, Flot et al. (2008) identified five distinct mitochondrial lineages compatible with morphology, though *P. eydouxi* and *P. meandrina* were differentiated only by ITS2. A few years later, Schmidt-Roach et al. (2013), focusing on *P. damicornis* ecomorphs described in (Veron and Pichon, 1976) and sequencing mitochondrial ORF, revealed five distinct lineages further named *P. damicornis* types α , β , γ , δ and ϵ (note that type δ was previously named σ in Schmidt-Roach et al. (2012), such as the sequences available in GenBank). In addition, these five types were renamed as *P. damicornis*, *P. acuta* Lamarck, 1816, *P. verrucosa* (Ellis & Solander, 1786), *P. aliciae* Schmidt-Roach, Miller & Andreakis, 2013, and *P. brevicornis* Lamarck, 1816, respectively (Schmidt-Roach et al., 2014), presenting distinct features in their reproduction modes. Meanwhile, Pinzón et al. (2013) studied colonies from various *Pocillopora* morpho-species sampled across the distribution range of the genus. Using ORF and microsatellites, they identified between five and eight genetically distinct lineages (named type 1 to type 8) that were generally incongruent with colony morphology across the sampling range. In Marti-Puig et al. (2014), from *Pocillopora* colonies representing a range of atypical morphologies, thought to be rare or endemic, ORF and Dloop revealed six distinct clades, while the intragenomic variation of ITS2 led to unresolved groups. Moreover, *corallum* macromorphology was found highly variable within each mitochondrial clade, while less variation was found at the microscopic level. Lastly, Schmidt-Roach et al. (2014), sequencing the ORF of *Pocillopora* colonies from Australian waters, identified four clades and a fifth clade corresponding to *P. effusus*-like colonies from Pinzón et al. (2013). Moreover, while analyzing ten characters of the macromorphology and the calice size for a subset of colonies (from two for *P. eydouxi*-like to 31 for *P. damicornis* type α -like), mitochondrial molecular phylogenies were found to be more or less congruent with the groups based on

macromorphology, but fine-scale morphological variations (particularly the shape and type of columella) were useful for differentiation between the identified clades.

All these studies underline (1) the fact that genetic data are crucial to delineate *Pocillopora* species, as some morpho-species hide a high genetic diversity, and (2) the necessity to explore the whole *Pocillopora* distribution, even in margins, to confirm or infirm the species status and their distribution range. However, all these studies delineate species on the basis of phylogenetic trees (robustness of the nodes) or networks, without any criterion of where to place the limit between intra-specific and inter-specific diversities. In this context, we conducted an exhaustive survey of *Pocillopora* colonies presenting a large panel of morphological variants from the Southwestern Indian Ocean, New Caledonia and French Polynesia, some of the less studied regions within *Pocillopora* range. Considering De Queiroz's unified species concept (a separately evolving divergent lineage, De Queiroz, 2007), we used (1) different species delimitation methods (GMYC, PTP, ABGD and Haplowebs) based on several mitochondrial (ORF and Dloop) and nuclear markers (ITS2) to determine Primary Species Hypotheses (PSH, *sensu* Pante et al., 2015), and (2) assignment tests based on microsatellite data to determine the first level of differentiation and suggest Secondary Species Hypotheses (SSH, *sensu* Pante et al., 2015) in *Pocillopora*. Then, we reevaluated the distribution and endemism of the identified SSHs.

Material and methods

Sampling

Colonies of *Pocillopora* were sampled between 2001 and 2015, within three marine provinces (Spalding et al., 2007): the Western Indian Ocean (WIO), the Tropical Southwestern Pacific (TSP) and the Southeast Polynesia (SEP), representing 15 localities (Table 1, Fig. 1). In each locality, colonies presenting all possible morphologies were sampled (apex of branch) without *a priori* in SCUBA diving or snorkeling and photographed *in situ* (except colonies from French Polynesia in 2001-2003 that were not collected for this purpose and colonies from Tromelin Island due to field difficulties). Each fragment was conserved in 90° ethanol for molecular analyses. Examining the undersea photograph, each colony was attributed a morph based on its general morphological characteristics including branch shape and thickness, size and uniformity of verrucae, and overall growth form as described in Veron (2000) and Schmidt-Roach et al. (2014). In order to check our morph designations, a sub-sample of pictures was

examined by three coral specialists (F. Benzoni, G. Faure and D. Obura). When one colony's morphology was unclear, several morphs were assigned.

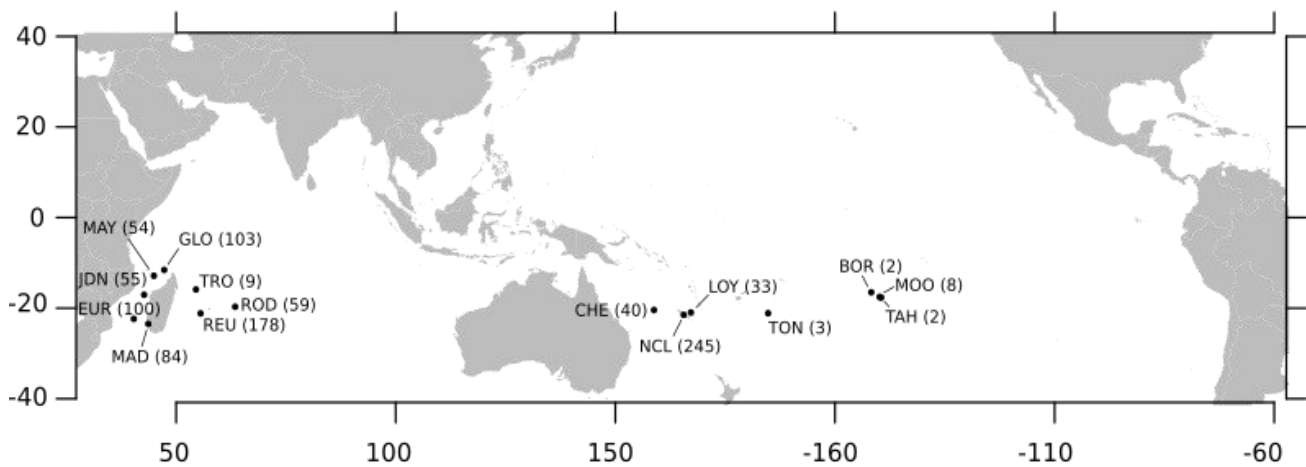


Fig 1. Sampling localities of *Pocillopora* colonies. The number of colonies sampled is indicated in parentheses near the locality code (from west to east and from north to south: GLO: Glorioso Islands; MAY: Mayotte; JDN: Juan de Nova Island; EUR: Europa Island; MAD: Madagascar (Tular); REU: Reunion Island; ROD: Rodrigues Island; TRO: Tromelin Island; CHE: Chesterfield Islands; NCL: New Caledonia; LOY: Loyalty Islands; TON: Tonga; BOR: Bora-Bora; MOO: Moorea; TAH: Tahiti).

DNA extraction and amplification

DNA was extracted using DNeasy Blood & Tissue Kit (Qiagen, Hilden, Germany) following the manufacturer's protocol. The mitochondrial open reading frame (ORF) was amplified with FATP6.1 and RORF primers (Flot et al., 2008) and the putative control region (Dloop) with FNAD5.2deg and RCOI primers (Flot et al., 2008). In addition, one nuclear marker (ITS2) was amplified with Scler5.8Sbforward and ITSrev primers (LaJeunesse and Pinzón, 2007). PCRs were performed in 25 μ L: 1X of MasterMix (Applied Biosystems), 0.3 μ M of forward and reverse primers and 1.6 ng/ μ L of genomic DNA. The thermocycling program was as follows: 94°C for 5 min + 40 \times (94°C for 60 s, 53°C for 60 s, 72°C for 60 s) + 72°C for 5 min for final elongation step. Amplicons were sent for sequencing to Genoscreen (Lille, France) on a capillary sequencer ABI 3730XL (Applied Biosystems).

Colonies were further genotyped using 13 microsatellite loci (Pd2-001, Pd3-004, Pd3-005, Pd2-006, Pd3-008, Pd3-009 (Starger et al., 2007), PV2, PV7 (Magalon et al., 2004), Poc40 (Pinzón and LaJeunesse, 2011), Pd4, Pd11, Pd13 (Torda et al., 2013) and Pd3-EF65 (Gorospe and Karl, 2013). Forward primers were indirectly fluorochrome labelled (6-FAM, VIC, NED, PET) by

adding a universal M13 tail at the 5'-end and were multiplexed post-PCR. Each amplification reaction was performed as in G elin et al. (unpublished results). PCR products were genotyped using an ABI 3730XL sequencer at the Plateforme Gentyane (INRA, Clermont-Ferrand, France). Allelic sizes were determined with GENEMAPPER 4.0 (Applied Biosystems) using an internal size standard (Genescan LIZ-500, Applied Biosystems).

Table 1. Sampling of *Pocillopora* colonies (N = 975). Are indicated for each locality: the marine province and the ecoregion according to Spalding et al. (2007), the locality code, the latitude and longitude, the total sample size (N) and the number of sequences obtained for ORF, Dloop and ITS2 markers (N_{ORF} , N_{Dloop} and N_{ITS2} , respectively).

Marine Province	Ecoregion	Locality	Locality code	Latitude	Longitude	N	N_{ORF}	N_{Dloop}	N_{ITS2}
Western Indian Ocean (WIO)	Western and Northern Madagascar	Madagascar (Tulear)	MAD	-23.47539	43.66148	84	80	22	59
		Europa Island	EUR	-22.36783	40.37185	100	100	13	74
		Juan de Nova Island	JDN	-17.04855	42.72176	55	53	33	36
		Mayotte	MAY	-12.83131	45.16044	54	47		
		Glorioso Islands	GLO	-11.56377	47.29394	103	91	36	70
	Cargados								
	Carajos/Tromelin Island	Tromelin Island	TRO	-15.88083	54.52714	9	9	1	4
	Mascarene Islands	Reunion Island	REU	-21.16115	55.57841	178	176	95	100
		Rodrigues Island	ROD	-19.69775	63.44172	59	59	21	25
	Tropical Southwestern Pacific (TSP)	New Caledonia	Chesterfield Islands	CHE	-20.41574	158.80233	40	39	38
		Grande Terre	NCL	-21.47567	165.57125	245	235	20	77
		Loyalty Islands	LOY	-20.96939	167.20426	33	29	8	17
	Tonga Island	Tonga	TON	-21.13061	-175.22125	3	3	2	3
Southeast Polynesia (SEP)		Bora-Bora	BOR	-16.50025	-151.73874	2	2	2	1
	Society Islands	Moorea	MOO	-17.52767	-149.83867	8	8	4	2
		Tahiti	TAH	-17.65834	-149.47704	2	2	2	2
Total						975	943	302	508

Sequence alignment and phylogenetic analyses

Sequences were checked and edited using GENEIOUS 8.0 (Kearse et al., 2012) and deposited in GenBank (Appendix S1 and S2). Additional ORF and Dloop sequences previously published (Appendix S1 and S2) were retrieved from GenBank. Sequences were aligned using MAFFT (Katoch et al., 2005) and trimmed to the shortest sequence length. For each marker, each haplotype was attributed a unique identifier (e.g. ORF01, Dloop01 and ITS2-001). The

nucleotide diversity (π) for each marker was estimated with DNASP 5.0 (Librado and Rozas, 2009). Then, JMODELTEST 2.5 (Darriba et al., 2012; Posada, 2008) was used to find the best substitution model based on AICc criterion of each fragment (ORF: HKY85+I+G; Dloop: HKY85+I+G; ITS2: K80+I+G). Moreover, as recent studies (Barley and Thomson, 2016; Zinger and Philippe, 2016) propose to use more complex sequence evolution models to better fit the data, we tested in addition for the ORF alignment a GTR+I+G model to reconstruct phylogenies and performed delineation methods.

Dloop sequences showed a 6 bp diagnostic gap between 190 and 195 bp. As no clear consensus about the treatment of gaps in alignments exists, analyses were both performed on the alignment with the gap as is and with the gap recoded as a single mutation. Concerning ITS2, some ambiguous profiles were detected for heterozygotes due to indels (one of 3 bp between positions 160 and 162 and a second of 4 bp repeated four times in maximum between positions 139 and 152) and double peaks. Ambiguities due to indels were resolved with reverse sequences and double peaks using PHASE 2.1 (Stephens and Donnelly, 2003; Stephens et al., 2001) implemented in DNASP 5.0 (Librado and Rozas, 2009). As for Dloop, two alignments were tested for ITS2: with gaps and with gaps recoded.

Analyses were conducted on each mitochondrial marker independently (ORF and Dloop) as well as on a concatenate dataset (ORF-Dloop). Additionally, we performed each analysis on several alignments: (1) containing all individual sequences (e.g. ORF_{all}), (2) containing a maximum of 10 sequences per haplotype (e.g. ORF₁₀) for counteracting the overrepresentation of certain haplotypes due to the imbalanced sampling, and (3) containing a single sequence per haplotype (e.g. ORF_{single}) as it is a prior for some analyses, giving each haplotype the same weight.

For GMYC and PTP species delimitation methods, we reconstructed phylogenies using Maximum Likelihood (ML) in PhyML (GENEIOUS plug-in; Guindon et al., 2010) and Bayesian inferences (BI) in MRBAYES 3.2 (3 independent runs of: 20×10^6 generations, 8 chains, temperature to 0.2, 10% burn-in length, sampling every 2×10^3 generations; GENEIOUS plug-in; Ronquist et al., 2012). Nodes were considered robust if their posterior probability (PP) was equal or higher to 0.95 for Bayesian reconstruction and when their bootstrap (BS) values were superior to 75% for ML reconstruction (Erixon et al., 2003). Members from other genera of the Pocilloporidae family [*Stylophora* Schweigger, 1820 and *Seriatopora* Lamarck, 1816 (accession numbers: KX618661- KX618678)] were used as outgroups for tree reconstructions but pruned for species delimitation analyses.

Delimiting PSHs

We analyzed sequence data using four quantitative methods of species delimitation: automatic barcode gap discovery (ABGD; Puillandre et al., 2012), generalized mixed Yule coalescent model (GMYC; Fontaneto et al., 2011; Pons et al., 2006), Poisson tree processes model (PTP; Zhang et al., 2013), and haplowebs (Flot et al., 2010). The three formers were applied on both mitochondrial markers (ORF and Dloop) and the latter on the nuclear one (ITS2). Each method uses different terminology to refer to the delimited taxa (ABGD = “groups” or “hypothetical species”; GMYC = “entities”; PTP = “phylogenetic species”; haplowebs = “field for recombination” or “FFR”), acknowledging that they may not represent true species. For clarity and consistency, whatever the method used, we will use the term “putative species”.

ABGD

We used the ABGD method developed by Puillandre et al. (2012) on the web-server: <http://wwwabi.snv.jussieu.fr/public/abgd/abgdweb.html>. This method employs a genetic distance-based approach to detect a “barcode gap” separating candidate species based on non-overlapping values of intra- and interspecific genetic distances and is independent of any tree topology (Hebert et al., 2003; Puillandre et al., 2012). After sequence alignment, we computed a matrix of pairwise distances using the K2P model (Kimura, 1980). A graphical representation of the pairwise distance distribution with different priors for each marker showed different barcoding gaps according to each marker (ORF: $P_{\min} = 0.001$, $P_{\max} = 0.08$ and $X = 1.0$; Dloop: $P_{\min} = 0.0005$, $P_{\max} = 0.001$; and $X = 0.5$; ORF-Dloop: $P_{\min} = 0.001$, $P_{\max} = 0.01$ and $X = 0.7$).

GMYC models

We used the GMYC method developed by Pons et al. (2006), implemented in R (R Development Core Team, 2016). This method infers species boundaries using the differences of the branching rates in an ultrametric phylogenetic tree (i.e. calibrated with a molecular clock) to discriminate between divergence between species following a strict Yule process [no extinction; (Yule, 1925)] and neutral coalescent events between lineages forming a species (Kingman, 1982). Ultrametric trees were constructed from the different input alignments of each marker using BEAST 1.8. (Drummond and Rambaut, 2007). We used a relaxed log-normal

clock with a coalescent tree prior as these have been identified as best priors for GMYC analyses (Esselstyn et al., 2012; Monaghan et al., 2009). Markov chains Monte Carlo (MCMC) were run for 50×10^6 generations, sampled every 5×10^3 generations. Chains convergence was assessed using TRACER 1.6 (Rambaut et al., 2014). The consensus trees (maximum clade credibility tree; 10% burn-in; tree not presented) were constructed with TREEANNOTATOR 1.7 (Rambaut and Drummond, 2013).

To account for uncertainty in species delimitation, we used three applications of the GMYC model: (1) the single-threshold species delimitation GMYC model, (2) the multiple-threshold species delimitation GMYC model using R (R Development Core Team, 2016) packages ‘ape’ (Paradis et al., 2004) and ‘splits’ (Ezard et al., 2014), and (3) Bayesian GMYC (bGMYC) model developed by Reid and Carstens (2012) in the package ‘bGMYC’ (Reid, 2014). Using the consensus tree, single- and multiple-threshold GMYC species delimitation models allow identifying respectively one or several thresholds, dividing coalescent and Yule processes on the tree (Monaghan et al., 2009). As the multiple-threshold model allows variation of evolution rates along branches and thus several shift points between Yule and coalescent processes across the phylogenetic tree, we compared the likelihood of both models’ outcomes. The Bayesian implementation of the GMYC model (Reid and Carstens, 2012) accounts for uncertainty in the phylogeny and in model parameters by sampling trees and conducting MCMC. This application gives marginal probabilities to putative species. As recommended by the authors, we conducted the bGMYC analysis by resampling the tree file generated by BEAST. Each of them was re-run for 5×10^4 generations, with 3×10^4 generations of burn-in and sampling every 100 steps, resulting in 200 new trees per initially sampled tree: *in fine*, 10^4 new trees were used. MCMC estimates from each tree were pooled to calculate the probabilities that two leaves in the phylogenies are conspecific. We set the probability of two leaves being conspecific at 0.60, 0.70, 0.80 and 0.90 and compared the partitions.

PTP analyses

PTP analyses incorporate the number of substitutions in the model of speciation and assume that the probability that a substitution gives rise to a speciation event follows a Poisson distribution. We ran PTP and the Bayesian implementation of PTP (bPTP) species delimitation analysis in the webserver (<http://species.h-its.org/ptp/>). Primarily, the PTP/bPTP analyses were performed with a ML tree as recommended by Zhang et al. (2013). As nodes from this tree were mostly unresolved (see Results), we also used ultrametric trees for a direct comparison

with GMYC results (identical input tree) and a Bayesian tree to compare species delimitation results with the Bayesian tree topology. We ran the bPTP analysis for 5×10^5 MCMC generations, with a thinning value of 500, a burn-in of 25%. We visually confirmed the convergence of the MCMC chain as recommended (Zhang et al., 2013).

Haploweb

The haplotype network for ITS2 was drawn with PopART (<http://popart.otago.ac.nz>) in Minimum Spanning and the haploweb (Flot et al., 2010) was constructed further by manually adding connectors between haplotypes co-occurring in at least one individual. Then, the haplotypes connected together were defined as single-locus fields for recombination (sl-FFR). These sl-FFRs reflect groups of individuals that share a common allele pool and thus can be considered as putative species following the criterion of mutual allelic exclusivity (Flot et al., 2010).

Finally, the combination of ABGD, GMYC, PTP and haploweb methods were used to identify the Primary Species Hypotheses (PSH). Each PSH was attributed a numerical identifier (e.g. PSH01). Indeed, since all these methods may present intrinsically different partitions, PSHs were defined by considering the most frequent partitions among the 16 different methods (see Fig. 2), rather than the most parsimonious ones. For example, ORF55 appeared as a singleton for only one method over 16 (ABGD analysis considering 9 putative species and the ORF_{single} input, see Fig. 2), so it was not considered as a robust PSH. Conversely, if one partition is observed across the majority of the methods, it will be considered as a valid PSH (e.g. the partition between ORF01 and ORF02 observed 16 times over 16, defining the limit between PSH01 and PSH02, or the partition between ORF23 and ORF24 observed 14 times over 16, delimiting the lower limit of PSH06 and the upper limit of PSH07).

Testing PSHs using microsatellite data to delimit SSH

We used two different methods to assess the genetic clustering of individuals within putative species. First, we performed a Discriminant Analysis of Principal Components (DAPC) in the R package ‘adegenet’ (Jombart, 2008; Jombart et al., 2010). DAPC is a non-model-based method that maximizes the differences between groups while minimizing variation within groups without prior information on individuals’ origin, without assuming

Hardy-Weinberg equilibrium (HWE) nor absence of linkage disequilibrium (LD). We assessed the optimal number of groups with the Bayesian information criterion (BIC) method (i.e. K with the lowest BIC value should reflect the optimal number of clusters). We tested values of $K = 1-20$, but BIC values may keep decreasing after the true value of K in case of genetic clines or/and hierarchical structure (Jombart et al., 2010). Furthermore, since retaining too many discriminant functions relative to the number of populations can lead to over-fitting of data, resulting in spurious discrimination of any set of clusters, the rate of decrease in BIC values was visually examined to identify values of K after which BIC values decreased only slightly (Jombart et al., 2010). Then, DAPC analysis was performed for different values of K , retaining axes of PCA sufficient to explain $\geq 80\%$ of the total variance. In addition, we conducted assignment analyses implemented in STRUCTURE 2.3.4 (Pritchard et al., 2000). We used the admixture model with default settings for inferring alpha and the correlated allele frequencies model, without any location or population priors using the following parameters (after validation of chains convergence): three iterations of 2×10^6 MCMC generations after an initial burn-in of 2×10^5 generations for each K , varying from $K = 1$ to $K = 20$. In addition to direct examination of STRUCTURE outputs, Evanno et al. (2005) proposed a method to choose the most likely K by analyzing the second-order rate of change of the posterior probability (PP) of the data between successive K values (ΔK), which was done using STRUCTURE HARVESTER 0.6.94 (Earl, 2012). Results files were combined using CLUMPP 1.1.2 (Jakobsson and Rosenberg, 2007) and visualized using DISTRUCT 1.1 (Rosenberg, 2004). Finally, we compared these assignment results to support or discard previously identified PSHs. Each SSH was attributed the identification number of the corresponding PSH (e.g. SSH01 for PSH01) and when one PSH was split into several SSH, a letter was added to each SSH (e.g. SSH05a and SSH05b).

Results

Sequences variability and trees construction

Over our 943 ORF sequences, 33 haplotypes were retrieved, to which we added 22 haplotypes from GenBank that we did not uncover, resulting in 55 haplotypes with 73 variable sites over 842 bp ($\pi = 0.015$). Two haplotypes from Forsman et al. (2013), not found among our sequences nor in other studies, were very short (217 bp) compared to all sequences from

the alignment and were therefore discarded (ORF56 and ORF57 in Appendix S1). We found that 15 of our haplotypes were identical to previous published sequences while 18 had never been described (in bold in Fig. 2 and GenBank accession numbers in Appendix S1). We numbered ORF haplotypes from ORF01 to ORF55, and added a reference sequence name, giving priority to the nomenclature found in Pinzón et al. (2013), since this study covers the largest area and the highest number of identified haplotypes. When haplotypes were not retrieved from this latter study, we named them following other studies, with given priority order: (1) Schmidt-Roach et al. (2013), (2) Marti-Puig et al. (2014) and (3) (Flot et al., 2008; Hsu et al., 2014; Mayfield et al., 2015) (see details in Appendix S1). Over the 18 newly identified ORF haplotypes, some were close to previously published haplotypes (e.g. ORF11 which is close to type α (Schmidt-Roach et al., 2013) and type 4 (a, b and c; Pinzón et al., 2013), or ORF22 close to type 8a (Pinzón et al., 2013); see Fig. 2 for more details), while some represented new lineages, such as ORF17 found in WIO, ORF24 in the SEP, ORF25 and ORF26 in WIO, and ORF51 and ORF52 in the TSP.

Among the 302 Dloop sequences, we identified 17 haplotypes (numbered from Dloop01 to Dloop17) to which 15 haplotypes retrieved from the literature were added (numbered from Dloop18 to Dloop32), resulting in 32 haplotypes (GenBank accession numbers in Appendix S2). Two of our 302 Dloop sequences, corresponding to two colonies of Juan de Nova Island and to Dloop17, presented large deletions and were removed from analysis. Thus, the 300 remaining sequences, representing 31 haplotypes, showed 32 variable sites over 881 bp ($\pi = 0.006$). Among the remaining 16 haplotypes we retrieved from our sampling, four have never been described (Dloop02, Dloop03, Dloop10 and Dloop16) but all were genetically close to previously published haplotypes and no new lineages were identified. Considering the difficulties in retrieving individuals with reference sequences for both markers in GenBank, the analysis on the concatenate alignment (ORF-Dloop) was conducted with our sequences only [i.e. 282 sequences, 35 haplotypes and 77 variable sites over 1723 bp ($\pi = 0.008$)]. Haplotype names resulted in the concatenation of ORF haplotype number and Dloop haplotype number, respectively (e.g. ORF01-Dloop09). Due to a weak number of haplotypes identified in the Dloop dataset and congruent results between Dloop and ORF datasets, we will only present the results from the ORF marker (results from the concatenate alignment are given in Appendix S3). Phylogenies constructed with Bayesian trees and those calibrated with a molecular clock (ultrametric trees) were congruent, but all trees constructed with the ML method resulted in few robust nodes and were not congruent with the previous ones.

Among our 508 retrieved ITS2 sequences, 128 haplotypes were identified, with 81 variable sites over 481 bp ($\pi = 0.008$). Results obtained from the alignment with recoded gaps were congruent with those from the original alignment, but both resulted in non-robust phylogenies. Then, we decided to perform further analyses on ITS2 without recoding gaps.

Delimitation species method

First, when using the different ORF inputs (ORF_{all}, ORF₁₀, ORF_{single}), the ABGD method indicated a fuzzy barcoding gap between 0.004 and 0.012 divergence (Appendix S4), delimitating three to 34 groups for ORF₁₀ and seven to nine groups for ORF_{single} (Table 2, Fig. 2), both inputs giving identified putative species congruent with the phylogeny. Results for ORF_{all} were not congruent with the tree topology and are not shown.

Second, the GMYC models over-split the data with ORF₁₀ leading to a higher number of putative species than the actual number of haplotypes (Table 2), an obvious artefact. Therefore, the analyses were conducted keeping a single sequence per haplotype for GMYC models (e.g. Talavera et al., 2013). The single threshold GMYC model resulted in nine clusters and four singletons (i.e. putative species represented by only one haplotype) whatever the evolution model used (HKY or GTR), but could not be considered different from the null model (single species phylogeny with only coalescent processes; $P = 0.24$ and $P = 0.19$, respectively). Conversely, the multiple threshold GMYC model (phylogeny composed of several species with several coalescent time values) was preferred over the null model ($P = 0.029$ for the HKY model and $P = 0.007$ for the GTR model). Under the HKY model, four independent switches between speciation and coalescent processes were revealed, resulting in 16 putative species: 11 GMYC clusters and five singletons while under the GTR model, three independent switches between speciation and coalescent processes were detected, resulting in 14 putative species (Fig. 2).

Third, for ORF_{single}, the bGMYC analysis identified 15 putative species, among which 11 clusters presented a probability ≥ 0.70 of being conspecific [bGMYC(70)] under the HKY and GTR models (Fig. 2, Table 2). When considering a probability ≥ 0.80 of being conspecific [bGMYC(80)], we identified 19 putative species and 14 clusters under the HKY model, and 22 putative species and 14 clusters under the GTR model (Fig. 2, Table 2). For bGMYC(70), 11 clusters (73%) were phylogenetically robust for both models (Fig. 2) while for bGMYC(80), 13 clusters for HKY (68%) and for GTR (59%) were phylogenetically robust. Moreover, the bGMYC analysis probably over-split the data for $P > 0.90$ (39 and 45 putative species for HKY

and GTR models, respectively; Table 2) while considering $P < 0.60$ might be too parsimonious (12 and 13 putative species for HKY and GTR models, respectively; Table 2).

Fourth, for ORF_{single}, the PTP analysis gave a nearly identical result to the one given by the ultrametric tree built with HKY or GTR models, resulting in 17 putative species including 12 clusters and five singletons. Only a slight difference was observed in the group composed by ORF07-ORF17 (Fig. 2). The bPTP analysis resulted in different partitions depending on the evolution model (24 and 18 putative species with HKY and GTR, respectively; Fig. 2). Results obtained from the Bayesian and the ML trees were identical whatever the evolution model used but they over-split our data (45 putative species for bPTP with the Bayesian tree and 48 putative species for bPTP with the ML tree over 55 ORF haplotypes; Table 2). We did not present these latter results.

Finally, the haploweb built on ITS2 sequences (Appendix S5a) resulted in a complex network due to (1) the high number of haplotypes and (2) the presence of heterozygous individuals presenting highly divergent haplotypes. Applying the rule that one haplotype cannot be shared by more than one species, this ended up to identify four sl-FFRs.

Identification of PSH

The three methods of species delimitation (ABGD, GMYC, PTP) gave different results. ABGD using the ORF dataset with one single sequence per haplotype gave the most parsimonious result with 7 and 9 putative species. All the other methods, although slight differences appeared, gave a finer partition with common patterns. Among GMYC and PTP methods giving consistent results, the number of putative species ranged from 10 to 23 (Fig. 2). Based on our criterion considering the most frequent partitions, 16 PSHs were recovered, further named PSH01 to PSH16 (Fig. 2). Correspondences with types previously named in the literature are summarized in Table 3 and, in more details, in Appendix S1; the different morphs for each PSH are shown in Fig. 3 and, in more details, in Appendix S6.

PSH01 was composed of ten colonies that shared the same ORF haplotype (ORF01; Appendix S6a) identical to the ORF sequence type 2 from Pinzón et al. (2013) identified as *P. effusus* Veron, 2000 accepted as *P. effusa* (Veron, 2000). These colonies were sampled both in WIO (Reunion Island and Madagascar) and TSP (Chesterfield Islands). PSH02 was composed of haplotypes not uncovered from our data (ORF02-05), identified as Clade IIIb (Marti-Puig et al., 2014) and type 6 (Pinzón et al., 2013), which appeared restricted to the Hawaiian archipelago according to these studies.

Table 2: Results of the species delimitation methods. ABGD: number of putative species for the automatic barcoding gap discovery method; sGMYC and mGMYC: single and multiple-threshold generalized mixed Yule coalescent (GMYC) models, respectively [Nclust = number of clusters; Nsp = number of putative species (clusters + singletons); CI = confidence interval for the number of putative species; Ln-Null = likelihood of the null model (all sequences form a single cluster); Ln-GMYC = likelihood of GMYC model; LRT = p-value of the likelihood ratio test]; bGMYC: Bayesian implementation of the GMYC model; 60, 70, 80 and 90 correspond to probabilities (in percentage) of being conspecific, respectively. Ultrametric, Bayesian and Maximum Likelihood (ML) correspond to the construction method of the input tree used for PTP and bPTP analyses (Poisson Tree Process method). ORF_{all} and ORF-Dloop_{all}: all haplotypes included in the analysis; ORF₁₀: a maximum of 10 sequences kept for each haplotype; ORF_{single} and ORF-Dloop_{single}: one sequence per haplotype included

	ABGD	Evolution Model	sGMYC				mGMYC				bGMYC				ultrametric		bayesian		ML	
			Nclust	Nsp (CI)	Ln-Null	Ln-GMYC (LRT)	Nclust	Nsp	Ln-Null	Ln-GMYC (LRT)	60	70	80	90	PTP	bPTP	PTP	bPTP	PTP	bPTP
ORF _{all} (N = 943)	9 - 14 - 18																			
ORF ₁₀ (N = 189)	3 - 10 - 14 - 34	HKY+I+G	57	85 (21-94)	2055	2058 (0.040*)	45	66 (52-86)	2055	2065 (<0.001***)	15	19	26	44	21	56	24	149		
		GTR+I+G	47	64 (59-74)	2060	2064 (0.012*)	46	66 (60-80)	2060	2070 (<0.001***)	11	18	19	29	23	50	24	149		
ORF _{single} (N = 55)	7 - 9	HKY+I+G	9	13 (1-27)	431	433 (0.236)	11	16 (14-26)	431	435 (0.029*)	12	15	19	39	17	24	12	45	48	48
		GTR+I+G	9	13 (1-25)	432	433 (0.194)	8	14 (11-12)	432	437 (0.007**)	13	15	22	45	17	18	12	45	48	48
ORF-Dloop _{all} - (N = 282)	5 - 13 - 20	HKY+I+G	74	110 (36-118)	3387	3393 (0.002**)	66	113 (81-117)	3387	3405 (<0.001***)	8	8	10	175	18	128	66	79		
ORF-Dloop _{all} gap recoded (N = 282)	5 - 13 - 20	HKY+I+G	65	97 (87-101)	3432	3451 (<0.001***)	68	103 (60-103)	3432	3461 (<0.001***)	8	14	14	175	21	120	68	106		
ORF-Dloop _{single} - (N = 35)	4 - 5 - 11	HKY+I+G	8	10 (1-34)	239	241 (0.226)	8	13 (1-20)	239	241 (0.160)	14	15	19	31	9	11	21	21	30	31
ORF-Dloop _{single} gap recoded (N = 35)	4 - 5 - 11	HKY+I+G	8	10 (1-19)	240	242 (0.210)	8	12 (1-20)	240	242 (0.160)	15	16	19	31	9	11	21	21	21	30

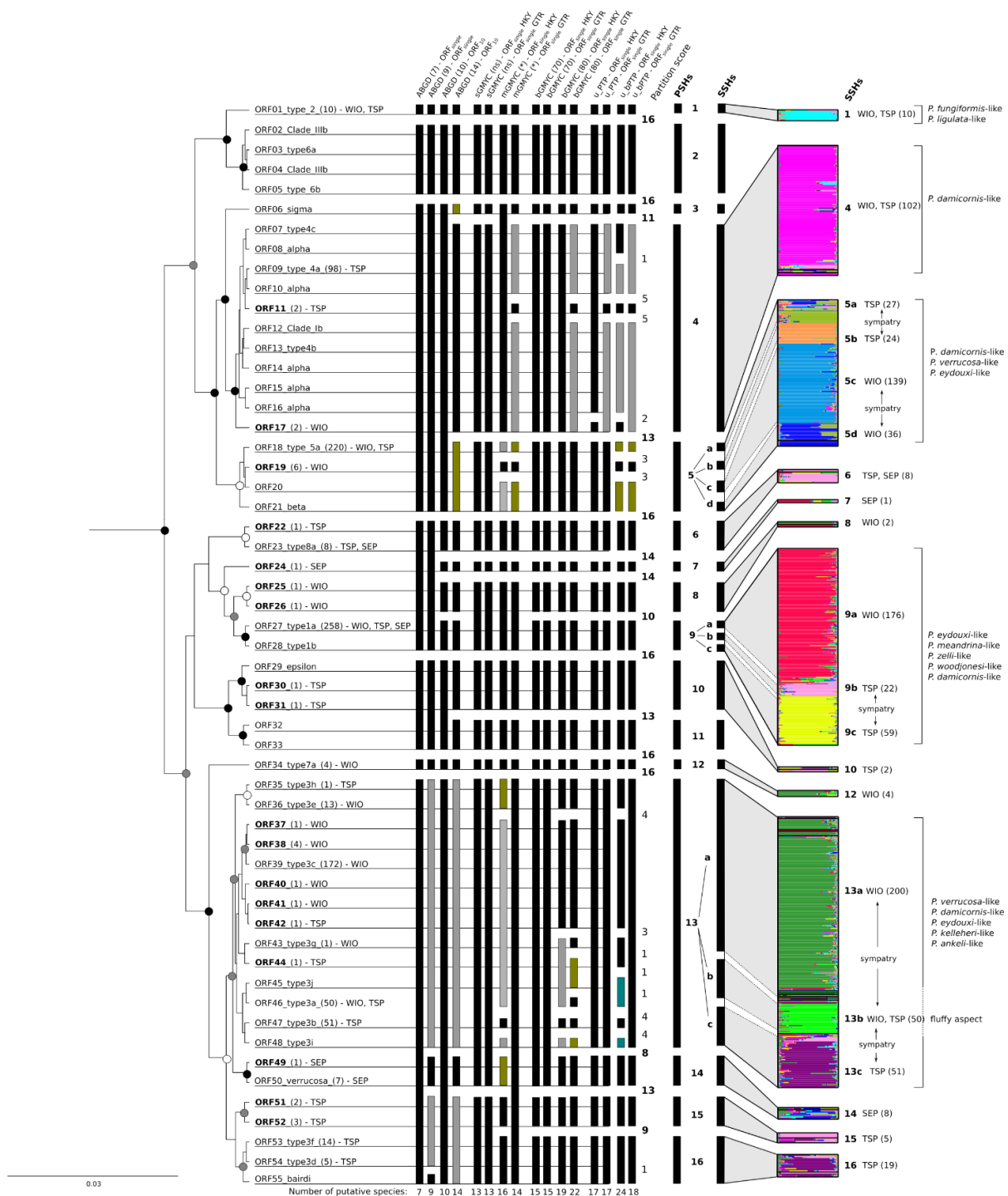


Fig. 2. Summary of the putative species limits as obtained through different methods. The phylogenetic tree is the ORF Bayesian tree; black nodes indicated posterior probability (PP) > 0.95 and bootstrap proportion (BS) > 0.70, gray nodes indicated 0.50 < PP < 0.95 and BS < 0.70 (PP/BS), and white nodes indicated PP > 0.95 and BS < 0.70. For nodes with PP < 0.50 and BS < 0.70, PP and BS values are not indicated. For each haplotype, are indicated the number of colonies in parentheses and the ecoregions of origin. In bold, are indicated haplotypes exclusively found in this study. ORF_{single} means that one sequence per haplotype was kept and ORF₁₀ means that a maximum of 10 sequences per haplotype were kept. ABGD (Automatic Barcoding Gap Discovery) results are presented for 7 and 9 putative species for ORF_{single} and for 10 and 14 putative species for ORF₁₀. GMYP (Generalized Mixed Yule Coalescent model) results are presented for ORF_{single} for single and multiple threshold models (sGMYP and mGMYP, respectively) and for the two evolution models tested (HKY and GTR). bGMYP results are presented for ORF_{single} at probabilities of 0.7 [bGMYP(70)] and 0.8 [bGMYP(80)] to be conspecific and for both evolution models (HKY and GTR). PTP/bPTP results are presented for the ultrametric tree built with the two evolution models (HKY and GTR). Among one repartition, the bars sharing the same color (except black) correspond to the same group. The number of putative species for each method is indicated below. The partition score indicated the number of partition found among all the methods. The Primary Species Hypotheses (PSH) are obtained according to the most frequent partition among all the delimitation methods. The STRUCTURE results are presented for K = 12, individuals are sorted by haplotype following the order of appearance in the tree. The Secondary Species Hypotheses (SSH) are the result of the consensus from the different delimitation methods based on ORF (PSH) and the assignment tests on microsatellite data.

Similarly, PSH03 was not observed over our sampling and corresponded to a single haplotype (ORF06) from Schmidt-Roach et al. (2013) named type δ . PSH04 was composed of 11 haplotypes (ORF07-17) covering type 4 (*sensu* Pinzón et al., 2013), type α (*sensu* Schmidt-Roach et al., 2014) and Clade Ib (*sensu* Marti-Puig et al., 2014) (Appendix S6b-d). Our colonies presenting these haplotypes (ORF09/ORF11/ORF17) were sampled both in the WIO (n = 2) and the TSP (n = 100). PSH05 was composed of four haplotypes (ORF18-21) which included type 5 (*sensu* Pinzón et al., 2013), Clade Ia (*sensu* Marti-Puig et al., 2014) and type β (*sensu* Schmidt-Roach et al., 2013). Our colonies presenting these haplotypes (Appendix S6e-f) were sampled both in the WIO (ORF18 and ORF19, n = 175) and the TSP (ORF18, n = 51). PSH06 (ORF22-23) included type 8 (*sensu* Pinzón et al., 2013) which was identified as *P. damicornis* by Pinzón et al. (2013), who also found that it matched with type ϵ (*sensu* Schmidt-Roach et al., 2013). Surprisingly, according to the Bayesian tree, ORF22 and ORF23 did not match with type ϵ in our study, but they appeared closer to type 1 (*sensu* Pinzón et al., 2013) than to type ϵ . Moreover, the nine colonies presenting these haplotypes were sampled both in the TSP (ORF22 and ORF23; n = 8) and the SEP (ORF23; n = 1). PSH07 (ORF24) was represented by a unique colony from Moorea (French Polynesia, SEP). This haplotype has never been found until now. PSH08 (ORF 25-26) was composed of two colonies from Juan de Nova (Mozambique Channel, WIO) and was not found in the literature (Appendix S6i). PSH09 (ORF27-28) corresponded to type 1 (*sensu* Pinzón et al., 2013) (Appendix S6j). All our colonies in this PSH shared the same haplotype (ORF27), whatever the geographic origin (Appendix S1). PSH10 (ORF29-31) included type ϵ (*sensu* Schmidt-Roach et al., 2013) and two new haplotypes from New Caledonia (ORF30 and ORF31; Appendix S6k). PSH11 (ORF32-33) was composed of two haplotypes from Taiwan (Hsu et al., 2014) and was not found in our sampling. PSH12 (ORF34) corresponded to type 7 (*sensu* Pinzón et al., 2013) and our colonies presenting this haplotype (n = 4) were only found in Tromelin Island. PSH13 (ORF35-48; Appendix S6l-t) was composed of 14 haplotypes, among which six newly sampled in both the WIO (ORF37, ORF38, ORF40 and ORF41) and the TSP (ORF42 and ORF44), and several already recorded types: 3a, 3b, 3c, 3e, 3g, 3h, 3i and 3j (*sensu* Pinzón et al., 2013), and γ (*sensu* Schmidt-Roach et al., 2013). PSH14 (ORF49-50) was composed of two haplotypes found in the SEP (French Polynesia and Cook Islands) and ORF50 matched with one haplotype previously found in (Mayfield et al., 2015), which the authors named *P. verrucosa*.

Table 3: Summary of the *Pocillopora* SSHs identified through species delimitation methods and microsatellite data. For each SSH, are indicated the ORF haplotypes, the name found in literature when possible, and finally, the associated *corallum* macromorphology found in our study and in the different studies cited in this table (b: *P. brevicornis*, c: *P. capitata*, d: *P. damicornis*, ef: *P. effusus*, ey: *P. eydouxi*, f: *P. fungiformis*, el: *P. elegans*, i: *P. inflata*, k: *P. kelleheri*, l: *P. ligulata*, me: *P. meandrina*, mo: *P. molokensis*, v: *P. verrucosa*, and w: *P. woodjonesi*). Data in bold are found in this study, data in bold and underlined are exclusively found in this study, other data are not found in this study, but in the literature.

PSH (this study)	1	2	3	4	5	6	7	8	9																				
SSH (this study)	1	2	3	4	5a, 5b, 5c, 5d	5d			9a, 9b, 9c																				
ORF haplotype (this study)	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	
Type (Pinzón et al 2013)	2		6a		6b		4c		4a			4c						5				8a					1a		
Type (Schmidt-Roach et al 2012, 2013, 2014)						δ aliciae											α damicornis		β acuta									e/m	
Clade (Marti-Puig et al 2014)	IIIa		IIIb					lb				lb							la									IIb	
Type (Souter et al 2010)																			F										
Type (Flot et al 2008)	B																											A	
Macromorphology (all studies)	f, l , ey, z, ef		l, d, mo, c				D												d, ey, v			d, v, el, me			d	v	ey, me , w , v, mo, d, z		

PSH (this study)	10	11	12	13	14	15	16																				
SSH (this study)	10	11	12	13a	13b	13c																					
ORF haplotype (this study)	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55
Type (Pinzón et al 2013)				7	3h	3e		3c			3g		3j	3a	3b	3i									3f	3d	
Type (Schmidt-Roach et al 2012, 2013, 2014)	ϵ brevicornis														γ verrucosa											γ verrucosa	bairdi
Clade (Marti-Puig et al 2014)																											
Type (Souter et al 2010)									NF																		
Type (Flot et al 2008)																											
Macromorphology (all studies)		d	b				v, d, me, ey , mo, k , z								v, d, ey, me , mo, k	v, k , d, me	v		v	y	d				v, d, k, l , me		

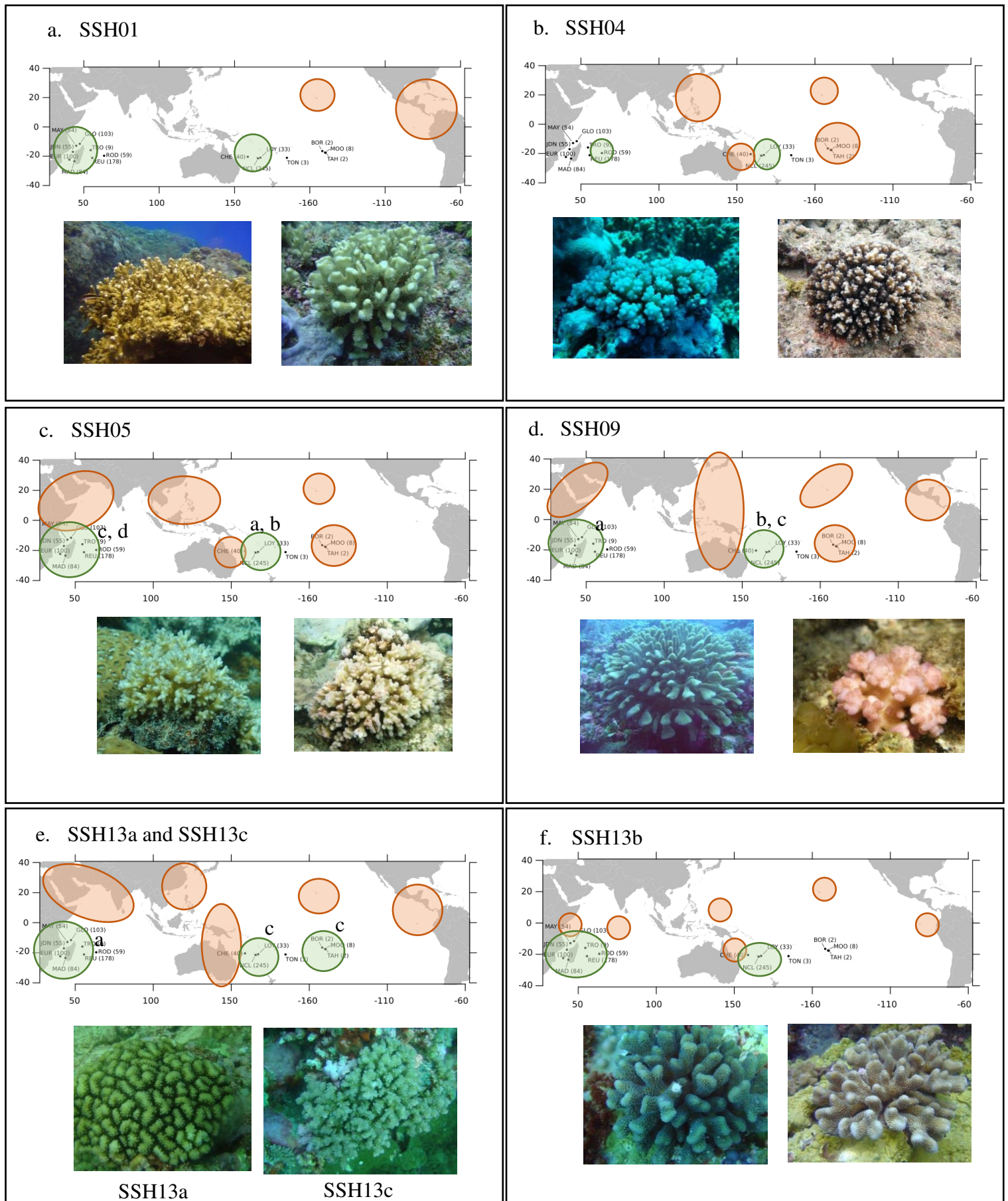


Fig. 3. Geographical distribution of six secondary species hypotheses (SSH): a. SSH01, b. SSH04, c. SSH05, d. SSH09, e. SSH13a and SSH13c, f. SSH13b. The red areas correspond to previous distribution identified from the literature (for which SSH could not be determined as microsatellite information was not available) and the green areas correspond to new areas from this study. For each SSH, two photographs are presented (more pictures for all SSHs are presented in Fig. A.4).

We retrieved colonies only in French Polynesia. PSH15 (ORF51-52) corresponded to two new haplotypes found in New Caledonia. The five sampled colonies presenting these two haplotypes showed a *P. verrucosa*-like macromorphology (Appendix S6v). PSH16 (ORF53-55) covered types 3d and 3f (*sensu* Pinzón et al., 2013) and *P. bairdi* (ORF55), species newly described in Schmidt-Roach et al. (2014). Our colonies harbored only ORF53 and ORF54 (type 3f and type 3d, respectively), and were found in the TSP (Chesterfield Islands, New Caledonia and Tonga Archipelago; Appendix S6w).

Comparing PSHs with ITS2 haploweb, only two PSHs appeared as sl-FFRs: PSH05 and PSH14. Taken by pairs, all other PSHs shared at least one ITS2 haplotype; *vice versa*, some ITS2 haplotypes were shared by at least two PSHs to a maximum of four PSHs (8 haplotypes over 128; Appendix S5b). Overall, this analysis using ITS2 did not lead to a conclusive result and will not be considered further.

PSH testing and SSH identification

The sequenced colonies were also genotyped in the aim to compare partitions obtained with the mtDNA (i.e. PSHs) to partitions obtained with clustering analyses based on microsatellite multi-locus genotypes, to define SSHs. PSHs represented by a small number of colonies (PSH06, PSH07, PSH08, PSH10, PSH12, PSH14, PSH15; $n \leq 8$) and those for which we did not get samples (PSH02, PSH03, PSH11) were not considered for comparison with clustering results obtained from the microsatellite data. Therefore, for further SSH identification, we considered colonies belonging to six ORF-based PSHs: PSH01, PSH04, PSH05, PSH09, PSH13 and PSH16, representing a total of 913 colonies. Nevertheless, all the sequenced colonies for the ORF marker ($n = 943$) were included in the clustering analyses. From Bayesian assignment (STRUCTURE), the Evanno's method indicated several values of K as valid partitions ($K = 4, 5, 8, 10$ or 12 ; Appendix S7a and S7b). As it is now admitted that there is no "true K " (Jombart and Collins, 2015; Raj et al., 2014; Verity and Nichols, 2016) and that the retained K is "merely a question of personal taste" (Jombart and Collins, 2015), we chose to keep $K = 12$. Indeed, the PSHs with few colonies continued to appear till this K (and not over-splitting the dataset). Therefore, all considered PSHs were recovered by the Bayesian clustering analysis, allowing us to designate them as SSHs (keeping the same identification number as PSH) (Schlick-Steiner et al., 2009). PSH05, PSH09 and PSH13, which presented the highest number of colonies, were subdivided in several genetic groups. Actually, colonies identified as PSH05 were further separated into four distinct genetic groups, named SSH05a,

SSH05b, SSH05c and SSH05d. Colonies belonging to SSH05c and SSH05d were sampled in the WIO: SSH05c was almost exclusively composed of colonies from Reunion Island (REU) and Rodrigues Island (ROD) ($n = 138$, 99%) while SSH05d was mainly composed of colonies from Juan de Nova Island (JDN), Madagascar (MAD) and Mayotte (MAY) ($n = 35$, 97%), with few colonies of both SSHs found in sympatry in MAD and REU ($n = 2$). Likewise, colonies belonging to SSH05a and SSH05b came from New Caledonia (TSP): SSH05a was composed of colonies sampled around Grande Terre (Nouméa, Bélep and Poindimié) while SSH05b was composed of colonies from the southwest of Grande Terre (Nouméa), where six colonies from SSH05a were also found. In a similar way, colonies belonging to PSH09 were assigned in three genetic groups, therefore corresponding to three SSHs (SSH09a, SSH09b and SSH09c). SSH09a colonies came from the WIO (all sampled localities), those of SSH09b were sampled in the TSP [Chesterfield Islands ($n = 2$) and New Caledonia ($n = 12$)] and in the SEP [French Polynesia ($n = 1$)] and those of SSH09c came from the TSP [Chesterfield Islands ($n = 18$) and New Caledonia ($n = 37$)]. SSH09b and SSH09c were found in sympatry at the same sampling sites in Chesterfield Islands and New Caledonia. PSH13 was also divided in three genetic groups (SSH13a, SSH13b, SSH13c): SSH13a was found in the WIO (all sampled localities), SSH13c in the TSP and the SEP [Chesterfield Islands ($n = 2$), New Caledonia ($n = 41$) and French Polynesia ($n = 1$)] while all colonies from SSH13b shared the exact same ORF haplotype (ORF46) and were found in nearly all sampling locations of the WIO and the TSP. Conversely, colonies from PSH16 failed to be assigned to a distinct genetic group based on their microsatellite data. However, it is worth noting that nine species delimitation methods distinguished them as a putative species (Fig. 2), so that they may be considered as members of a distinct SSH (SSH16).

Regarding the results from the DAPC analysis, the plot of BIC as a function of the number of clusters K (ranging from 1 to 50) does not present a real minimum. Nevertheless, we were able to detect slight slope decreases in the BIC plot for $K = 4, 7, 11$ and 13 (Appendix S7c). All the graphs obtained from STRUCTURE and DAPC for the different values of K are presented in Appendix S8. Even if the optimal values of K were not equal between methods (STRUCTURE or DAPC), the assignment results were highly similar for a given K , except that from $K \geq 8$ since DAPC revealed two genetic groups in sympatry within SSH09a (ORF27, *P. eydouxi/meandrina*) not retrieved using STRUCTURE. Therefore, DAPC analysis suggested that SSH09 may be split into four distinct putative species instead of three with the Bayesian assignment (two putative species for SSH09a, one for SSH09b and one for SSH09c).

Finally, reconsidering PSHs that were not represented by a sufficient number of colonies in view of the clustering analyses, we could extrapolate that all these PSHs are likely to represent distinct SSHs, as PSHs for which we sampled a sufficient number of colonies, were all confirmed as at least one SSH. So they should represent at least 7 SSHs, keeping the same identification number as their respective PSHs.

When examining the *in situ* macromorphology (determined just from undersea photographs) in regard of the PSH/SSH revealed by the different analyses, it is clear that morphs are not exclusive of one PSH/SSH (Table 3): one PSH/SSH could be composed of different morphs; reciprocally, several PSH/SSHs could share the same morph. For example, some colonies belonging to PSH01 in our sampling could be attributed to *P. fungiformis*-like morphs (tall and upright branches like candles, encrusting at the basis) and others to *P. ligulata*-like morphs (irregular branches with flattened ends and truncated tips) (Fig. 3a, Appendix S6a). Moreover, Pinzón et al. (2013) found also some colonies belonging to PSH01 and named them as *P. effusus* Veron 2000 (prostrate flattened branches which fuse). In addition, the colonies belonging to SSH13a and SSH09, either from our study or from the literature, were attributed various morphs in common such as *damicornis*, *eydouxi*, *meandrina*, *molokensis*, *verrucosa*, *zelli* (for details, see Table 3). Likewise, the *damicornis sensu lato* morph was found in 12 SSHs.

Discussion

Among the 943 colonies sampled in three biogeographic provinces and analyzed (Western Indian Ocean, Tropical Southwestern Pacific and Southeast Polynesia), we were able to identify 33 ORF and 16 Dloop haplotypes among which 18 and 4 were brand new, respectively. For the first time in *Pocillopora* genus, we used species delimitation methods (ABGD, GMYC and PTP) on both markers and we revealed 16 Primary Species Hypotheses (PSH). Confronting them to assignment tests using multi-locus genotypes from 13 microsatellites, we identified at least 18 Secondary Species Hypotheses (SSH) that could represent new putative species.

Species delimitation methods efficiency for *Pocillopora*

All species delimitation methods used in this study are designed for barcode markers while the two mitochondrial markers we used (ORF and Dloop) are not designed for this purpose initially in corals. Nevertheless, they both have been described as hypervariable and are widely used in studies focusing on *Pocillopora*, as they allow identifying several robust lineages (e.g. Marti-Puig et al., 2014; Pinzón et al., 2013; Schmidt-Roach et al., 2014). However, mitochondrial markers, which redraw only maternal lineages evolution, might not be appropriate to study speciation in cnidarians due to its slow evolution (Shearer et al., 2002). Park et al. (2012) evidenced that one of the youngest most recent common ancestor for Pocilloporidae is dated at the beginning of the Paleogene period (approximately 66 Myr) and *Pocillopora* appearance is dated at the end of the Neogene period (approximately 1.8 Myr), this genus appearing younger than the other genera from the Pocilloporidae family (*Madracis*, *Seriatopora* and *Stylophora*; Park et al., 2012). Furthermore, Veron (2000) highlighted the fact that *Stylophora* and *Pocillopora* were the most affected genera by the glaciation events of the Plio-Pleistocene. Thus the weak mtDNA diversity found in the *Pocillopora* genus compared to other models [e.g. 42 haplotypes for *Dascyllus quadrimaculatus* in French Polynesia (Fauvelot et al., 2003), or 33 haplotypes for *Chelonia mydas* in the Western Indian Ocean (Taquet, 2007)] might be the combined result of an important bottleneck that occurred between the Neogene and the Quaternary periods during the glaciation events, and the slow evolution of mtDNA (in average) in Anthozoans. However, in *Pocillopora* genus, the ORF marker “may be variable enough to be used as molecular markers to shed light on the taxonomy and phylogeography of *Pocillopora*”, as said Flot and Tillier (2007), and seems appropriate to discriminate putative species and might be considered as a potential barcode marker. Furthermore, Dloop, which is longer than ORF and thus harder to amplify, was less variable than ORF, and led to the same interpretations than ORF. So it seems redundant with this latter, making it unnecessary for future studies. On the other hand, the nuclear ITS2 marker, which corresponds to a multigene family showing high level of heterozygosity within colony (Flot et al., 2008) with possible incomplete lineage sorting (Márquez et al., 2002; van Oppen et al., 2001), did not lead to informative results in our study. Indeed, this extraordinary number of haplotypes shared by different PSHs in our study, could be the sign of introgressive hybridization. However, as ITS homogenization may proceed extremely slowly when divergent haplotypes are combined in a single genome (Modrich and Lahue, 1996), rDNA data cannot readily distinguish between ancient or recent hybridization events. Therefore, it appeared to be difficult to interpret for *Pocillopora* species delimitation when focusing on a large number of colonies and a vast geographic scale.

Regarding species delimitation methods, as results may vary following the input, it is important to test several parameters and evolution models to choose the most appropriate ones for each analysis. ABGD appeared to be the most parsimonious method when keeping a single sequence per haplotype, congruent with the tree topology. However, restricting drastically the dataset might impede the identification of several putative species, some being revealed when ten sequences of each haplotype were kept, corresponding to those obtained with the other methods. ABGD lumping tendency when compared to other methods has recently been described (Pentinsaari et al., 2016), highlighting the necessity to compare results from several methods. The highly parsimonious results from ABGD might represent ancestral lineages (corresponding to older nodes in the phylogeny) and this method may not be able to detect recent speciation events (Reid and Carstens, 2012).

Tree-based methods appear to be more accurate but rely on a batch of hypotheses (Barley and Thomson, 2016). More precisely, PTP has been shown to outperform the GMYC method, particularly when evolutionary distances between species are small (Zhang et al., 2013). We evidenced a small nucleotide diversity among haplotypes, suggesting that PTP would be the more appropriate method for our data. Nevertheless, our data did not fit with the best prior for the PTP analysis: a ML tree as input. In our case, the ML tree gave no robust results and the PTP analyses performed with this tree largely over-split the data and did not seem to outperform GMYC analyses. Conversely, a recent study using species delimitation methods (ABGD, GMYC, PTP and haplowebs) on *Goniopora* corals from Saudi Arabian Red Sea (Terraneo et al., 2016), found that all the methods gave congruent results except the GMYC method which over-split the data. Furthermore, to better fit the data, it had been suggested by Barley and Thomson (2016) to use more complex evolution models than the one identified with JMODELTEST. Nevertheless, our results did not show a better partition with the HKY+I+G nor the GTR+I+G models. As a conclusion, the species delimitation methods performed differently according to the biological model.

Unsurprisingly, all species delimitation methods used in this study gave different partitions. No consensus exists about the choice criterion: it had been suggested to consider the most parsimonious partition (e.g. Postaire et al., 2016), but it would probably underestimate the true number of putative species. Here, it would correspond to the ABGD partitions while the other methods (GMYC and PTP) gave less parsimonious results, these latter differing slightly from each other but remaining congruent. So considering the most frequent truncation points among all methods should give a good picture of putative species boundaries and seems more relevant than considering only one (e.g. Tang et al., 2014). Anyhow, assignment tests using

microsatellite data might help to refine these boundaries. Indeed, microsatellite markers are known to have a high mutation rate inducing high polymorphism that make them powerful for assessing genetic similarity between individuals and close taxa (*e.g.* Guichoux et al., 2011). Nevertheless, their high mutation rate promotes homoplasy, confusing the interpretations of their evolution (Slatkin, 1995) and making them usually unsuitable for phylogenetic analyses beyond closely related species (*e.g.* Mesak et al., 2014). Still, the results we obtained from microsatellite data were congruent with species delimitation methods for *Pocillopora*. Though, the two clustering analyses (STRUCTURE and DAPC) revealed differences in the optimal number of identified clusters: the number of clusters (K) may vary in structured populations and must be seen as a flexible parameter (Verity and Nichols, 2016). That is why we used several criteria (ΔK and BIC) to help estimating the most appropriate K .

Revision of *Pocillopora* species and their geographical distribution

The combination of species delimitation methods based on mitochondrial markers allowed us to identify 16 PSHs. Our protocol using assignment tests based on microsatellite markers led us to delineate several SSHs among these PSHs and many question the validity of the current taxonomy of *Pocillopora* genus.

Interestingly, among the 18 new haplotypes found in this study, five of them, representing eight colonies, were well-supported divergent lineages and the species delimitation methods classified them in three distinct SSHs: SSH07, SSH08, and SSH15. These latter were uncommon (less than six colonies each) and were found in restricted areas [Moorea (French Polynesia), Juan de Nova Island (Scattered Islands), Grande Terre and Loyalty Islands (New Caledonia), respectively]. Thus they could represent endemics but additional sampling is needed to confirm their status.

Likewise, SSH06, SSH12, and SSH14 were represented by very few colonies (each ≤ 8), their ORF haplotypes having already been found by other studies. These SSHs, although rare, are present in the WIO and the TSP and some of them require a revision of their geographic distribution. Indeed, SSH12 (ORF34) corresponds to type 7 from Pinzón et al. (2013), who found that ORF types 3 and 7 consisted of a single genetically homogeneous population in the Red Sea and Arabian Gulf while the microsatellite data suggested them that type 7 was a divergent variant. Conversely, we found that type 7 (*sensu* Pinzón et al., 2013) is well-defined by all the species delimitation methods we tested while our clustering analysis using microsatellite data did not, potentially because we found only four colonies harboring this

haplotype in Tromelin Island (Scattered Islands), at the North-East of Madagascar. Combining both results, we conclude that SSH12 seems to be a true species according to our criterion (De Queiroz, 2007); its geographic distribution initially restricted to the north part of the Indian Ocean (Pinzón et al., 2013) could be extended to the south till Tromelin Island. In the same way, SSH06 previously found in Taiwan (type 8; Pinzón et al., 2013) has also been found in the TSP and SEP, enlarging its geographic distribution to the south-east. For SSH14 that we found in SEP, no extension of its range is noticed, despite the addition of one new haplotype. Indeed, colonies belonging to this SSH have already been found in French Polynesia (Forsman et al., 2013), in the Cook Islands and Austral Islands (Mayfield et al., 2015) and might be endemic to SEP.

Concerning SSHs presenting sufficient sampling ($n \geq 10$), SSH04 corresponding to *P. damicornis* type α (Schmidt-Roach et al., 2013) and type 4 (Pinzón et al., 2013) formed a well-defined homogeneous genetic lineage considering all methods, making it a possibly true species. Moreover, the distribution area for this species, which was only described, till now, in the Pacific Ocean, is extended in the WIO (two colonies found in Rodrigues Island), but further investigations are needed to precise its range in the WIO. So the name *Pocillopora damicornis* given by Schmidt-Roach et al. (2014) is meaningful, under the hypothesis that it is taxonomically valid.

Whatever the methods used, SSH01 seems to represent a true species, even if some identifications based on morphology diverge between authors. Indeed, colonies were named *P. effusus* in the TEP (type 2, Pinzón et al., 2013), considered as endemic to this zone by Veron (2000). Conversely we attributed *P. fungiformis* morph to our colonies in the WIO and the TSP [considered as endemic to the South of Madagascar by Veron (2000)], and also *P. ligulata* morph for some colonies in the TSP. Although Pinzón et al. (2013) found that ORF type 2 (SSH01) and type 6 (SSH02) were grouped together, all our delimitation species methods separated them into two distinct SSHs with type 6 possibly endemic from Hawaii. Thus, SSH01 was previously found in the TEP (Pinzón et al., 2013), Hawaii (Flot et al., 2010; Marti-Puig et al., 2014) and Phoenix Islands (Marti-Puig et al., 2014). Now, we can enlarge its geographic distribution to the TSP (Chesterfield Islands) and to the WIO (Reunion Island and South Madagascar). Thus the names *P. fungiformis* and *P. effusus* may refer to a morph (more or less encrusting, with fusing branches at the basis) rather than to real species.

For PSH05, PSH09 and PSH13, each presenting a large sampling, the pattern is more complex as each might represent more than one SSH. Indeed, concerning PSH05 that corresponds to colonies from *P. damicornis* type β , also named *P. acuta* (Schmidt-Roach et al., 2013; Schmidt-

Roach et al., 2014), type 5 (Pinzón et al., 2013), type F (Souter, 2010) and type b (Flot et al., 2008), four distinct lineages were revealed, two in the WIO and two in the TSP. As they were not found in real sympatry at the reef level, the differentiation between SSH05a and SSH05b or between SSH05c and SSH05d may result from geographic structuring. Thus, we did not get enough proofs to conclude either they represent one or several (maximum of four) distinct species.

Concerning PSH09, corresponding to the *P. eydouxi/meandrina* complex (Schmidt-Roach et al., 2014) and types 1 and 9 (Pinzón et al., 2013), results are also difficult to interpret. It is noteworthy that colonies attributed to the morphs *P. eydouxi* and *P. meandrina* shared the exact same ORF haplotype but are morphologically different (Schmidt-Roach et al., 2014). Nevertheless, this latter study focused on morphometric characters of only four and seven colonies from *P. eydouxi* and *P. meandrina*, respectively. Considering the morphological plasticity in the *Pocillopora* genus (Todd, 2008), these results are to be taken with cautious and morphology should be further investigated. Anyway, from our analyses, SSH09b and SSH09c are found in sympatry at the locality level in the TSP and thus could be considered as two distinct species. Performing assignments considering only the multi-locus genotypes harboring ORF type 1, Pinzón et al. (2013) also found two clusters (called type 1 and type 9), which were found in sympatry at least in Hawaii and Lizard Island in the Pacific. SSH09b and SSH09c might refer to these two clusters. However, they found type 9 also in the WIO, while our colonies from the WIO were assigned in a unique cluster, SSH09a. So SSH09a represents either the geographic structuring between ocean basins and it is associated either with SSH09b or SSH09c, or it could be considered as a third distinct species among PSH09. So the complex *P. eydouxi/meandrina* could represent a complex of at least two species. The geographical range of each species should be reevaluated in new insight of microsatellite data.

Concerning PSH13, corresponding to *P. damicornis* type γ or *P. verrucosa* (Schmidt-Roach et al., 2013; Schmidt-Roach et al., 2014) and type 3 (Pinzón et al., 2013), SSH13a and SSH13c were not found in sympatry. Thus we could not conclude whether they represent two distinct species, as they might be hidden by the geographic structuring between ocean basins. However, SSH13b corresponds to a unique ORF haplotype [ORF46 or type 3a (Pinzón et al., 2013)] found in sympatry with both SSH13a and SSH13c. So even if it was not detected by species delimitation methods based on mitochondrial sequences, it should be considered as a distinct species. Furthermore, other criteria point in the direction of its species status. Indeed, all the colonies belonging to SSH13b showed a stout morph with a fluffy aspect [tiny and tight verrucae with polyps highly extended; Fig. 3f, Appendix S6u and photographs 3 and 4 in

P. elegans monograph in Veron (2000)] making them highly recognizable, even in the field. Additionally, when we performed blind tests including pictures of this latter colonies mixed with pictures of other morphs, the three taxonomists and we always recognized them as a distinct morph, whatever the name each one gave. From an ecological point of view, in April-May 2016 in Mayotte (WIO), *Pocillopora* colonies from 0 to 15 m depth on the outer reef slope suffered a severe bleaching event with the exception of these colonies presenting this fluffy aspect (H. Magalon, University of Reunion Island, pers. com.; Appendix S9), indicating that they might represent a distinct genetic species or holobiont. Anecdotally, considering their zooxanthellae, they were genotyped using ITS2 and did not differ from other *Pocillopora* colonies present at the same site and collected the same day (ITS2 type C1; data not shown), indicating that resistance to bleaching may be due more to the host genetic identity rather than to the symbionts alone.

Concerning SSH16, which was represented by 19 colonies from the TSP, colonies were poorly assigned but close to SSH13c, which is exclusively composed of colonies from the TSP. The weak number of SSH16 colonies might have biased the assignment test, not allowing clear clustering and thus impeding us to conclude on this species status.

Finally, concerning the SSHs absent from our sampling, a supplementary sampling effort is needed to confirm their presence in the ecoregions we studied. However, SSH02 [corresponding to type 6 (Pinzón et al., 2013) and Clade IIIb (Marti-Puig et al., 2014)], SSH03 [corresponding to type δ (Schmidt-Roach et al., 2013)] and SSH11 (Hsu et al., 2014) were not found at all overall our whole sampling (943 colonies sequenced nor over the 8168 *Pocillopora* colonies genotyped by microsatellites for another purpose). So they might be extremely rare in our sampling area, suggesting that they might represent endemic species from Hawaii, Eastern Australia and Taiwan, respectively, till the day they will be discovered elsewhere.

***Pocillopora* species identification: nothing but genetics?**

As previously shown (Marti-Puig et al., 2014; Pinzón and LaJeunesse, 2011), our results confirm that morphology and genetics are not always congruent. Indeed, macromorphology cannot be used as a solely character, notably because of the huge plasticity in this genus and also because some colonies belonging to different SSHs are morphologically close (e.g. SSH06 with SSH09, Appendix S6g-h and Appendix S6j or SSH04 with SSH10, Appendix S6b-d and Appendix S6k). All in all, colonies within SSHs were morphologically similar, even if, in

extensively sampled SSHs, some colonies sometimes still play tricks [e.g. some individuals sampled in Madagascar were morphologically close to *P. damicornis* type α (SSH04) or β (SSH05) while they belonged to SSH09 (Appendix S6j) due to particular environmental conditions (reef crest and rough hydrodynamism conditions)], and others looked like weird cases [e.g. two *P. eydouxi/meandrina*-like colonies belonging to SSH09c sampled in natural environment in Chesterfield Islands (TSP) and showing a pattern like heterophylly in plants, with branches and verrucae showing completely different morphs without fusion of colonies; Appendix S10, as Figure 8 in Marti-Puig et al. (2014)]. In addition, we found several morphs within the same SSH (e.g. SSH09 with five distinct morphs at least), and several SSHs composed of colonies that could be classified in the same morph (e.g. SSH04, SSH05 and SSH13 for *P. damicornis*-like morph, and SSH05 and SSH13 for *P. verrucosa*-like morph) as in (Pinzón and LaJeunesse, 2011) who found, among colonies from the TEP, several morphotypes corresponding to one genetic cluster and *vice versa*. As an illustration, we genotyped again, with the 13 microsatellites, all the *P. meandrina* colonies (used in Magalon et al., 2005) and sequenced some of them. During this previous study, colonies were identified based on morphoanatomical criteria (some of them by M. Pichon). The present result of the clustering analysis assigned them into three SSHs: SSH14 (72%), SSH13c (8%), SSH09b (4%), and 16% were not assigned in any SSH (individual probability to belong to any cluster < 0.7). Populations Mo2, Mo3 (Moorea), Ta (Tahiti) and Bo (Bora-Bora) were homogeneous at 99% and all colonies (except one) belonged to SSH14, confirming the absence of structuring among these populations. But population Mo1 (Moorea) was composed of a mix of colonies belonging to SSH14 and SSH09b, explaining possibly the observed differentiation with all the other populations, even at a small scale. Similarly, populations from Tonga (To1 and To2) were composed of a mix of colonies belonging to several SSHs (SSH13c, SSH16, SSH04, SSH14) with often weak assignment probabilities for each cluster. So the differentiation pattern between these studied populations is due to a mixing of colonies belonging to distinct evolutionary lineages, demonstrating the necessity to define correctly the entity we wish to work on.

This high morphological variability among *Pocillopora* species may be due to recombination and hybridization, distorting identification based on sole morphology. Indeed, this phenomenon does not seem anecdotal in corals: it has been shown in several genera [e.g. *Pocillopora*, (Combosch et al., 2008; Combosch and Vollmer, 2015), *Madracis*, (Frade et al., 2010), *Montastraea* (Szmant et al., 1997)] and sometimes could produce “new species” as for *Acropora prolifera*, hybrid between *Acropora cervicornis* and *Acropora palmata* (van Oppen et al., 2000). In this way, Combosch and Vollmer (2015) revealed hybridization between

Pocillopora type 1 and type 3 using RADseq and highlighted the fact that macromorphology is not maternally inherited. More confusing, they studied *P. damicornis*-like colonies exhibiting an ORF type 1, a haplotype usually attributed to the *P. eydouxi/meandrina* complex in our study and previous ones (Pinzón et al., 2013; Schmidt-Roach et al., 2013; Schmidt-Roach et al., 2014). Rigorously, as for *A. cervicornis/palmata/prolifera* complex, *Pocillopora* type 1, type 3, and their hybrids should not be considered as distinct species considering the biological concept, but as a unique species or syngameon (even if names remain valid) as they exchange genes and as the evolutionary fate of one species is not independent from the other (i.e. a beneficial mutation can be shared by both species through hybridization). More generally, if *Pocillopora* colonies are able to reproduce with close species, the biological species concept (no reproduction between species) might not be an appropriate criterion to delineate *Pocillopora* species. If frequent events of hybridization are confirmed by genomics studies and several observations, and if hybrids are viable, it will be necessary to re-evaluate what represents a species in corals, as already suggested in Pinzón et al. (2013) (for a review, see also Willis et al., 2006). However, for the SSHs that present a sufficient number of colonies, microsatellite data did not show a pattern of mixing indicating possible hybrid species such as in *Acropora* complex. But this hypothesis cannot be ruled out for SSHs that did not present enough colonies and for which assignment methods are not applicable due to the small sample size. Nevertheless, hybrids can be present but in lesser proportions in our populations indicating that this phenomenon must be rare.

Species delimitation is thus especially challenging in corals due to an extreme phenotypic plasticity and possible hybridization. Moreover, a recent study evidenced cases of mosaicism and chimerism studying *Acropora florida*, *A. hyacinthus*, *A. sarmentosa*, the *Pocillopora* species complex, and *Porites australiensis* (Schweinsberg et al., 2015), implying the presence of multiple genomes at the intra-colony level. Thus we need to insure that the colonies' fragments from which DNA is extracted are composed of cells harboring a unique genome, above all in case of mosaicism which appears to be more frequent than chimerism and where cells from different genomes are juxtaposed within the tissue.

Conclusion

Species delimitation methods (ABGD, GMYC and PTP) based on mitochondrial markers and assignment tests based on microsatellite data helped us in disentangling the puzzle

of *Pocillopora* species. Obviously, each method taken independently gave slightly different results, but their combination might approach reality, confirming the necessity to use different methods to properly assess species boundaries. With our approach, we define at least 18 Secondary Species Hypotheses among the genus, confirming the difficulty to define *Pocillopora* species in general and the lack of confidence in macromorphological characters. Our study rejoins the growing body of literature supporting the need of an integrative framework to identify reliable criteria in order to reassess the taxonomy of Cnidarians, above all of *Pocillopora* brain-teaser. For this latter, it becomes necessary that taxonomists reach an agreement on species naming to not feel like an archeologist when comparing different studies and to avoid confusing interpretation for non-warned readers. Next generation sequencing, combined with other parameters (i.e. microstructure, zooxanthellae identification, ecology even at a micro-scale, resistance and resilience ability to bleaching) will be the next step towards *Pocillopora* species delimitation, which will have profound implications for ecological studies, such as studying biodiversity, response to global warming and symbiosis.

Acknowledgments

Corals sampling in New Caledonia (HM) was carried out during COBELO (doi: <http://dx.doi.org/10.17600/14003700>), BIBELOT (doi: <http://dx.doi.org/10.17600/13100100>) and CHEST (<http://dx.doi.org/10.17600/15004500>) oceanographic campaigns on board of RV Alis (IRD). Sampling in Reunion Island (HM, PG) was supported by program CONPOCINPA (LabEx CORAIL fund); in Madagascar (HM) and in Rodrigues Island (HM) supported by project Biodiversity (POCT FEDER fund); in Europa, Juan de Nova and Glorioso Islands (HM) by program BIORECIE (financial supports from INEE, INSU, IRD, AAMP, FRB, TAAF, and the foundation Veolia Environnement); in Tromelin Island (HM) by program ORCIE (INEE), and in Mayotte (HM) by program SIREME (FED). HM thanks all the buddies who helped in photographs during diving (J. Butscher, S. Andréfouët, L. Bigot, M. Pinault). We acknowledge the Plateforme Gentyane of the Institut National de la Recherche Agronomique (INRA, Clermont-Ferrand, France) for guidance and technical support. PG was financially supported by a PhD contract from the LabEx CORAIL.

Supplementary material

Appendix S1. Recap chart of all identified haplotypes compared to literature data. For each haplotype are indicated the associated publications, the corresponding name given in each publication (Type), the corresponding macromorphology given in each publication (Morpho-species), the corresponding location (AG: Arabian Gulf, AND: Andaman, BOR: Bora-Bora, CHA: Chagos, CHE: Chesterfield Islands, CLI: Clipperton Island, COK: Cook Islands, EA: Eastern Australia, GAL: Galapagos, GLO: Glorioso Islands, HAW: Hawaii, HER: Heron

Island, HOW: Howland Island, JDN: Juan de Nova Island, JPN: Japan, KEN: Kenya, LIZ: Lizard Island, MAU: Mauritius, MAD: Madagascar, MAY: Mayotte, MOO: Moorea, NCL: New Caledonia, PHO: Phoenix Island, PLW: Palau, REU: Reunion Island, ROD: Rodrigues Island, RS: Red Sea, TAH: Tahiti, THA: Thailand, TRO: Tromelin Island, TON: Tonga, TWN: Taiwan, TZN: Tanzania, WA: Western Australia, ZAN: Zanzibar) and the GenBank accession numbers.

Appendix S2. GenBank accession numbers for Dloop and ITS2 haplotypes used in this study.

Appendix S3. Summary of the putative species limits as obtained through different methods. The phylogenetic tree is the ORF-Dloop Bayesian tree; black nodes indicated posterior probability (PP) > 0.95 and bootstrap proportion (BS) > 0.70, gray nodes indicated $0.50 < PP < 0.95$ and $BS < 0.70$ (PP/BS), and white nodes indicated $PP > 0.95$ and $BS < 0.70$. For nodes with $PP < 0.50$ and $BS < 0.70$, PP and BS values are not indicated. ORF-Dloop_{single} means that one sequence per haplotype was kept, ORF-Dloop_{single with gap} (a) meant that the gap was kept and ORF-Dloop_{single gap recoded} (b) meant that the gap was recoded. For each haplotype, are indicated the number of colonies in parentheses. Among one repartition, the bars sharing the same color (except black) correspond to the same group. ABGD results are presented for 4, 5 and 11 putative species in (a) and for 5 and 11 putative species in (b). GMYC results are presented for single and multiple threshold model (sGMYC and mGMYC, respectively). bGMYC results are presented at probabilities of 0.7 [bGMYC(70)] and 0.8 [bGMYC(80)] to be conspecific. PTP/bPTP results are presented for the different trees: u_PTP/u_bPTP for ultrametric tree and bay_PTP/bay_bPTP for the Bayesian tree. The Primary Species Hypotheses (PSH) are obtained according to the most frequent partition of all the delimitation methods.

Appendix S4. ABGD graphs: a. ORF_{single}, b. ORF₁₀, c. ORF_{all}, d. ORF-Dloop_{single} and e. ORF-Dloop_{all}.

Appendix S5. a. Haploweb constructed by adding manually connection between haplotypes found in one colony, single locus fields for recombination (sl-FFR) are represented by the gray subsets. b. haplotypes are grouped considering the Secondary Species Hypotheses.

Appendix S6. Miscellany of colonies' pictures for each Secondary Species Hypothesis (SSH) representing *corallum* macromorphology diversity.

Appendix S7. a. Ln (K) and b. Δ(K) obtained from STRUCTURE HARVESTER for the Evanno's method and c. the BIC distribution for the DAPC analysis.

Appendix S8. a. Structure plots from K =2 to K =18, b. DAPC plots from K =2 to K =18

Appendix S9. Picture of the reef of Mayotte (May 2016) just after the bleaching event of March-April 2016 in the Western Indian Ocean. Colonies belonging to SSH13b (circled on the picture) were the only ones to be alive among *Pocillopora* colonies.

Appendix S10. Pictures of weird cases from colonies belonging to SSH09c, showing a pattern like heterophylly in plants.

Appendix S11. Number of individuals per ORF haplotype (Hap) and per locality, found in the Western Indian Ocean (WIO), the Tropical Southeast Pacific (TSP) and Southeastern Polynesia (SEP). GLO: Glorioso Islands, MAY: Mayotte, JDN: Juan de Nova Island, MAD: Madagascar, EUR: Europa Island, REU: Reunion Island, ROD: Rodrigues Island, TRO: Tromelin Island, CHE: Chesterfield Islands, NCL: New Caledonia (Grande Terre), LOY: Loyalty Islands, TON: Tonga, BOR: Bora-Bora, MOO: Moorea, TAH: Tahiti.

Appendix S12. Haplotype combination of the three markers (ORF, Dloop, ITS2) found among individuals.

Appendix S13. Genetic differentiation between SSHs. Values in the lower-left matrix are estimated with Weir and Cockerham's F_{ST} ; values in the upper-right matrix are estimated with Jost's D_{est} .

References

- Appeltans, W., Ahyong, S.T., Anderson, G., Angel, M.V., Artois, T., Bailly, N., Bamber, R., Barber, A., Bartsch, I., Berta, A., Błażewicz-Paszkowycz, M., Bock, P., Boxshall, G., Boyko, C.B., Brandão, S.N., Bray, R.A., Bruce, N.L., Cairns, S.D., Chan, T.-Y., Cheng, L., Collins, A.G., Cribb, T., Curini-Galletti, M., Dahdouh-Guebas, F., Davie, P.J.F., Dawson, M.N., De Clerck, O., Decock, W., De Grave, S., de Voogd, N.J., Domning, D.P., Emig, C.C., Erséus, C., Eschmeyer, W., Fauchald, K., Fautin, D.G., Feist, S.W., Fransen, C.H.J.M., Furuya, H., Garcia-Alvarez, O., Gerken, S., Gibson, D., Gittenberger, A., Gofas, S., Gómez-Daglio, L., Gordon, D.P., Guiry, M.D., Hernandez, F., Hoeksema, B.W., Hopcroft, R.R., Jaume, D., Kirk, P., Koedam, N., Koenemann, S., Kolb, J.B., Kristensen, R.M., Kroh, A., Lambert, G., Lazarus, D.B., Lemaitre, R., Longshaw, M., Lowry, J., Macpherson, E., Madin, L.P., Mah, C., Mapstone, G., McLaughlin, P.A., Mees, J., Meland, K., Messing, C.G., Mills, C.E., Molodtsova, T.N., Mooi, R., Neuhaus, B., Ng, P.K.L., Nielsen, C., Norenburg, J., Opresko, D.M., Osawa, M., Paulay, G., Perrin, W., Pilger, J.F., Poore, G.C.B., Pugh, P., Read, G.B., Reimer, J.D., Rius, M., Rocha, R.M., Saiz-Salinas, J.I., Scarabino, V., Schierwater, B., Schmidt-Rhaesa, A., Schnabel, K.E., Schotte, M., Schuchert, P., Schwabe, E., Segers, H., Self-Sullivan, C., Shenkar, N., Siegel, V., Sterrer, W., Stöhr, S., Swalla, B., Tasker, M.L., Thuesen, E.V., Timm, T., Todaro, M.A., Turon, X., Tyler, S., Uetz, P., van der Land, J., Vanhoorne, B., van Ofwegen, L.P., van Soest, R.W.M., Vanaverbeke, J., Walker-Smith, G., Walter, T.C., Warren, A., Williams, G.C., Wilson, S.P., Costello, M.J., 2012. The Magnitude of Global Marine Species Diversity. *Curr. Biol.* 22, 2189-2202.
- Barley, A.J., Thomson, R.C., 2016. Assessing the performance of DNA barcoding using posterior predictive simulations. *Mol. Ecol.* 25, 1944-1957.
- 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.

- 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.
- Darriba, D., Taboada, G.L., Doallo, R., Posada, D., 2012. jModelTest 2: more models, new heuristics and parallel computing. *Nat. Methods* 9, 772-772.
- De Queiroz, K., 2007. Species concepts and species delimitation. *Syst. Biol.* 56, 879-886.
- Drummond, A.J., Rambaut, A., 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol. Biol.* 7.
- Earl, D.A., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv. Genet. Resour.* 4, 359-361.
- Erixon, P., Sennblad, B., Britton, T., Oxelman, B., 2003. Reliability of Bayesian posterior probabilities and bootstrap frequencies in phylogenetics. *Syst. Biol.* 52, 665-673.
- Esselstyn, J.A., Evans, B.J., Sedlock, J.L., Anwarali Khan, F.A., Heaney, L.R., 2012. Single-locus species delimitation: a test of the mixed Yule-coalescent model, with an empirical application to Philippine round-leaf bats. *Proc. R. Soc. Lond., B, Biol. Sci.* 279, 3678-3686.
- Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* 14, 2611-2620.
- Ezard, T., Fujisawa, T., Barraclough, T., 2014. splits: SPecies' LLimits by Threshold Statistics.
- Fauvelot, C., Bernardi, G., Planes, S., 2003. Reductions in the mitochondrial DNA diversity of coral reef fish provide evidence of population bottlenecks resulting from Holocene sea-level change. *Evolution* 57, 1571-1583.
- Fennessy, J., Bidon, T., Reuss, F., Kumar, V., Elkan, P., Nilsson, Maria A., Vamberger, M., Fritz, U., Janke, A., Multi-locus analyses reveal four giraffe species instead of one. *Curr. Biol.* 26, 1-7.
- Flot, J.-F., Couloux, A., Tillier, S., 2010. Haplowebs as a graphical tool for delimiting species: a revival of Doyle's "field for recombination" approach and its application to the coral genus *Pocillopora* in Clipperton. *BMC Evol. Biol.* 10, 1-14.
- 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. *C. R. Biol.* 331, 239-247.
- Flot, J.F., Tillier, S., 2006. Molecular phylogeny and systematics of the scleractinian coral genus *Pocillopora* in Hawaii. *Proceedings of 10th international coral reef symposium.*
- 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.
- Fontaneto, D., Iakovenko, N., Eyres, I., Kaya, M., Wyman, M., Barraclough, T.G., 2011. Cryptic diversity in the genus *Adineta* Hudson & Gosse, 1886 (Rotifera: Bdelloidea: Adinetidae): a DNA taxonomy approach. *Hydrobiologia* 662, 27-33.
- Forsman, Z., Johnston, E., Brooks, A., Adam, T., Toonen, R., 2013. Genetic evidence for regional isolation of *Pocillopora* Corals from Moorea. *Oceanography* 26, 153-155.
- Frade, P.R., Reyes-Nivia, M.C., Faria, J., Kaandorp, J.A., Luttikhuisen, P.C., Bak, R.P., 2010. Semi-permeable species boundaries in the coral genus *Madracis*: introgression in a brooding coral system. *Mol. Phylogenet. Evol.* 57, 1072-1090.
- Gompert, Z., Nice, C.C., Fordyce, J.A., Forister, M.L., Shapiro, A.M., 2006. Identifying units for conservation using molecular systematics: the cautionary tale of the Karner blue butterfly. *Mol. Ecol.* 15, 1759-1768.

- Gorospe, K.D., Karl, S.A., 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.
- Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., Léger, P., Lepais, O., Lepoittevin, C., Malausa, T., Revardel, E., Salin, F., Petit, R.J., 2011. Current trends in microsatellite genotyping. *Mol. Ecol. Resour.* 11, 591-611.
- Guindon, S., Dufayard, J.F., Lefort, V., Anisimova, M., Hordijk, W., Gascuel, O., 2010. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst. Biol.* 59, 307-321.
- Hayashibara, T., Shimoike, K., 2002. Cryptic species of *Acropora digitifera*. *Coral Reefs* 21, 224-225.
- Hebert, P.D.N., Cywinska, A., Ball, S.L., DeWaard, J.R., 2003. Biological identifications through DNA barcodes. *Proc R Soc Lond B Biol Sci* 270.
- Hsu, C.-M., de Palmas, S., Kuo, C.-Y., Denis, V., Chen, C.A., 2014. Identification of scleractinian coral recruits using fluorescent censusing and DNA barcoding techniques. *PLoS One* 9, e107366.
- Jakobsson, M., Rosenberg, N.A., 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23, 1801-1806.
- Jombart, T., 2008. *adegenet*: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403-1405.
- Jombart, T., Collins, C., 2015. A tutorial for discriminant analysis of principal components (DAPC) using adegenet 2.0. 0.
- 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.
- Katoh, K., Kuma, K.-i., 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., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Mentjies, P., Drummond, A., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 28, 1647-1649.
- Keshavmurthy, S., Yang, S.-Y., Alamaru, A., Chuang, Y.-Y., Pichon, M., Obura, D., Fontana, S., De Palmas, S., Stefani, F., Benzoni, F., MacDonald, A., Noreen, A.M.E., Chen, C., Wallace, C.C., Pillay, R.M., Denis, V., Amri, A.Y., Reimer, J.D., Mezaki, T., Sheppard, C., Loya, Y., Abelson, A., Mohammed, M.S., Baker, A.C., Mostafavi, P.G., Suharsono, B.A., Chen, C.A., 2013. DNA barcoding reveals the coral “laboratory-rat”, *Stylophora pistillata* encompasses multiple identities. *Sci. Rep.* 3, 1520.
- Kimura, M., 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16, 111-120.
- Kingman, J.F.C., 1982. On the genealogy of large populations. *J. Appl. Probab.* 19, 27-43.
- LaJeunesse, T.C., Pinzón, J.H., 2007. Screening intragenomic rDNA for dominant variants can provide a consistent retrieval of evolutionarily persistent ITS (rDNA) sequences. *Mol. Phylogenet. Evol.* 45, 417-422.
- Librado, P., Rozas, J., 2009. DnaSP v5: a software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25, 1451-1452.
- Magalon, H., Adjeroud, M., Veuille, M., 2005. Patterns of genetic variation do not correlate with geographical distance in the reef building coral *Pocillopora meandrina* in the South Pacific. *Mol. Ecol.* 14, 1861-1868.

- 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.
- Márquez, L.M., Van Oppen, M.J., Willis, B., Reyes, A., Miller, D.J., 2002. The highly cross-fertile coral species, *Acropora hyacinthus* and *Acropora cytherea*, constitute statistically distinguishable lineages. *Mol. Ecol.* 11, 1339-1349.
- Marti-Puig, P., Forsman, Z.H., Haverkort-Yeh, R.D., Knapp, I.S.S., Maragos, J.E., Toonen, R.J., 2014. Extreme phenotypic polymorphism in the coral genus *Pocillopora*; micro-morphology corresponds to mitochondrial groups, while colony morphology does not. *Bull. Mar. Sci.* 90, 211-231.
- Mayfield, A.B., Chen, C.-H., Chen, C.-S., Bruckner, A.W., 2015. A survey of pocilloporids and their endosymbiotic dinoflagellate communities in the Austral and Cook Islands of the South Pacific. *Platax* 12, 1-17.
- McFadden, C.S., Brown, A.S., Brayton, C., Hunt, C.B., van Ofwegen, L.P., 2014. Application of DNA barcoding in biodiversity studies of shallow-water octocorals: molecular proxies agree with morphological estimates of species richness in Palau. *Coral Reefs* 33, 275-286.
- Mesak, F., Tatarenkov, A., Earley, R.L., Avise, J.C., 2014. Hundreds of SNPs vs. dozens of SSRs: which dataset better characterizes natural clonal lineages in a self-fertilizing fish? *Front. Ecol. Evol.* 2, 74.
- Modrich, P., Lahue, R., 1996. Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu. Rev. Biochem.* 65, 101-133.
- Monaghan, M.T., Wild, R., Elliot, M., Fujisawa, T., Balke, M., Inward, D.J., Lees, D.C., Ranaivosolo, R., Eggleton, P., Barraclough, T.G., 2009. Accelerated species inventory on Madagascar using coalescent-based models of species delineation. *Syst. Biol.* 58.
- Mora, C., Tittensor, D.P., Adl, S., Simpson, A.G.B., Worm, B., 2011. How many species are there on earth and in the ocean? *PLoS Biol.* 9, e1001127.
- Moritz, C., 2002. Strategies to protect biological diversity and the evolutionary processes that sustain it. *Syst. Biol.* 51.
- Nakajima, Y., Nishikawa, A., Iguchi, A., Sakai, K., 2012. The population genetic approach delineates the species boundary of reproductively isolated corymbose acroporid corals. *Mol. Phylogenet. Evol.* 63, 527-531.
- Padial, J.M., Miralles, A., De la Riva, I., Vences, M., 2010. The integrative future of taxonomy. *Frontiers in Zoology* 7, 1-14.
- 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.
- Paradis, E., Claude, J., Strimmer, K., 2004. APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20.
- Park, E., Hwang, D.S., Lee, J.S., Song, J.I., Seo, T.K., Won, Y.J., 2012. Estimation of divergence times in cnidarian evolution based on mitochondrial protein-coding genes and the fossil record. *Mol. Phylogenet. Evol.* 62, 329-345.
- Paz-García, D.A., Aldana-Moreno, A., Cabral-Tena, R.A., García-De-León, F.J., Hellberg, M.E., Balart, E.F., 2015. Morphological variation and different branch modularity across contrasting flow conditions in dominant *Pocillopora* reef-building corals. *Oecologia* 178, 207-218.
- Pentinsaari, M., Vos, R., Mutanen, M., 2016. Algorithmic single-locus species delimitation: effects of sampling effort, variation and non-monophyly in four methods and 1870 species of beetles. *Mol. Ecol. Resour.*, doi:10.1111/1755-0998.12557.
- Pfenninger, M., Schwenk, K., 2007. Cryptic animal species are homogeneously distributed among taxa and biogeographical regions. *BMC Evol. Biol.* 7, 1-6.

- 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., Maggs, 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.
- Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S., Kamoun, S., Sumlin, W.D., Vogler, A.P., 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Syst. Biol.* 55, 595-609.
- Posada, D., 2008. jModelTest: phylogenetic model averaging. *Mol. Biol. Evol.* 25.
- Postaire, B., Magalon, H., Bourmaud, C.A.F., Bruggemann, J.H., 2016. Molecular species delimitation methods and population genetics data reveal extensive lineage diversity and cryptic species in Aglaopheniidae (Hydrozoa). *Mol. Phylogenet. Evol.* 105, 36-49.
- Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 954-959.
- Puillandre, N., Lambert, A., Brouillet, S., Achaz, G., 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Mol. Ecol.* 21, 1864-1877.
- R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Raj, A., Stephens, M., Pritchard, J.K., 2014. fastSTRUCTURE: Variational inference of population structure in large SNP datasets. *Genetics*.
- Rambaut, A., Drummond, A.J., 2013. TreeAnnotator v1. 7.5.
- Rambaut, A., Suchard, M.A., Xie, D., Drummond, A.J., 2014. Tracer v1.6, Available from <http://beast.bio.ed.ac.uk/Tracer>.
- Reid, N., 2014. bGMYC: A Bayesian MCMC implementation of the general mixed Yule-coalescent model for species delimitation.
- Reid, N.M., Carstens, B.C., 2012. Phylogenetic estimation error can decrease the accuracy of species delimitation: a Bayesian implementation of the general mixed Yule-coalescent model. *BMC Evol. Biol.* 12, 1-11.
- Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D.L., Darling, A., Höhna, S., Larget, B., Liu, L., Suchard, M.A., Huelsenbeck, J.P., 2012. MrBayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst. Biol.* 61, 539-542.
- Rosenberg, N.A., 2004. DISTRUCT: a program for the graphical display of population structure. *Mol. Ecol. Notes.* 4, 137-138.
- Schlick-Steiner, B.C., Steiner, F.M., Seifert, B., Stauffer, C., Christian, E., Crozier, R.H., 2009. Integrative taxonomy: a multisource approach to exploring biodiversity. *Annu. Rev. Entomol.* 55, 421-438.
- Schmidt-Roach, S., Lundgren, P., Miller, K.J., Gerlach, G., Noreen, A.M.E., Andreakis, N., 2013. Assessing hidden species diversity in the coral *Pocillopora damicornis* from Eastern Australia. *Coral Reefs* 32, 161-172.
- Schmidt-Roach, S., Miller, K.J., 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.
- Schmidt-Roach, S., Miller, K.J., Woolsey, E., Gerlach, G., Baird, A.H., 2012. Broadcast spawning by *Pocillopora* species on the Great Barrier Reef. *PLoS One* 7, e50847.
- Schweinsberg, M., Weiss, L.C., Striewski, S., Tollrian, R., Lampert, K.P., 2015. More than one genotype: how common is intracolony genetic variability in scleractinian corals? *Mol. Ecol.* 24, 2673-2685.

- Shearer, T.L., Van Oppen, M.J.H., Romano, S.L., Wörheide, G., 2002. Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). *Mol. Ecol.* 11, 2475-2487.
- Sites, J.W., Marshall, J.C., 2004. Operational criteria for delimiting species. *Ann. Rev. Ecol. Evol. Syst.* 35, 199-227.
- Slatkin, M., 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139, 457-462.
- Souter, P., 2010. Hidden genetic diversity in a key model species of coral. *Mar. Biol.* 157, 875-885.
- Spalding, M.D., Fox, H.E., Allen, G.R., Davidson, N., Ferdaña, Z.A., Finlayson, M., Halpern, B.S., Jorge, M.A., Lombana, A., Lourie, S.A., Martin, K.D., McManus, E., Molnar, J., Recchia, C.A., Robertson, J., 2007. Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *Bioscience* 57, 573-583.
- Starger, C.J., Yeoh, S.S., Dai, C.F., Baker, A.C., Desalle, R., 2007. Ten polymorphic STR loci in the cosmopolitan reef coral, *Pocillopora damicornis*. *Mol. Ecol. Resour.* 8, 619-621.
- Stefani, F., Benzoni, F., Yang, S.Y., Pichon, M., Galli, P., Chen, C.A., 2011. Comparison of morphological and genetic analyses reveals cryptic divergence and morphological plasticity in *Stylophora* (Cnidaria, Scleractinia). *Coral Reefs* 30, 1033-1049.
- Stephens, M., Donnelly, P., 2003. A comparison of bayesian methods for haplotype reconstruction from population genotype data. *Am. J. Hum. Genet.* 73, 1162-1169.
- Stephens, M., Smith, N.J., Donnelly, P., 2001. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* 68, 978-989.
- Szmant, M.A., Weil, E., Miller, W.M., Colón, E.D., 1997. Hybridization within the species complex of the scleractinian coral *Montastraea annularis*. *Mar. Biol.* 129, 561-572.
- Talavera, G., Dincă, V., Vila, R., Paradis, E., 2013. Factors affecting species delimitations with the GMYC model: insights from a butterfly survey. *Methods Ecol. Evol.* 4, 1101-1110.
- Tang, C.Q., Humphreys, A.M., Fontaneto, D., Barraclough, T.G., 2014. Effects of phylogenetic reconstruction method on the robustness of species delimitation using single-locus data. *Methods Ecol. Evol.* 5, 1086-1094.
- Taquet, C., 2007. Diversité et différenciation génétiques des populations de tortues vertes (*Chelonia mydas*) dans les sites de ponte et d'alimentation du sud-ouest de l'océan Indien : Application aux stratégies de conservation de l'espèce. *Ocean, Atmosphere. Université de La Réunion*, pp. 226.
- Terraneo, T.I., Benzoni, F., Arrigoni, R., Berumen, M.L., 2016. Species delimitation in the coral genus *Goniopora* (Scleractinia, Poritidae) from the Saudi Arabian Red Sea. *Mol. Phylogenet. Evol.* 102, 278-294.
- Todd, P.A., 2008. Morphological plasticity in scleractinian corals. *Biol. Rev.* 83, 315-337.
- Torda, G., Schmidt-Roach, S., Peplow, L.M., Lundgren, P., van Oppen, M.J.H., 2013. A rapid genetic assay for the identification of the most common *Pocillopora damicornis* genetic lineages on the great barrier reef. *PLoS One* 8, e58447.
- van Oppen, M.J., McDonald, B.J., Willis, B., Miller, D.J., 2001. The evolutionary history of the coral genus *Acropora* (Scleractinia, Cnidaria) based on a mitochondrial and a nuclear marker: reticulation, incomplete lineage sorting, or morphological convergence? *Mol. Biol. Evol.* 18, 1315-1329.
- van Oppen, M.J., Willis, B.L., Vugt, H.W., Miller, D.J., 2000. Examination of species boundaries in the *Acropora cervicornis* group (Scleractinia, cnidaria) using nuclear DNA sequence analyses. *Mol. Ecol.* 9, 1363-1373.
- Vaughan, T.W., Wells, J.W., 1943. Revision of the suborders, families and genera of the Scleractinia. *Geol Soc Am Bull Special Papers* 44, 1-363.
- Verity, R., Nichols, R.A., 2016. Estimating the number of subpopulations (K) in structured populations. *Genetics* 203, 1827-1839.

- Veron, J.E.N., 2000. Corals of the world. Australian Institute of Marine Science.
- Veron, J.J., Pichon, M.M., 1976. Scleractinia of Eastern Australia. Part I: Families Thamnasteriidae, Astrocoeniidae, Pocilloporidae. In: Science, A.I.o.M. (Ed.).
- Wallace, C.C., 1999. Staghorn corals of the world: a revision of the coral genus *Acropora*. CSIRO Publishing, Melbourne.
- Warner, P.A., van Oppen, M.J.H., Willis, B.L., 2015. Unexpected cryptic species diversity in the widespread coral *Seriatopora hystrix* masks spatial-genetic patterns of connectivity. *Mol. Ecol.* 24, 2993-3008.
- Wiens, J.J., 2007. Species delimitation: new approaches for discovering diversity. *Syst. Biol.* 56, 875-878.
- Willis, B.L., van Oppen, M.J.H., Miller, D.J., Vollmer, S.V., Ayre, D.J., 2006. The role of hybridization in the evolution of reef corals. *Ann. Rev. Ecol. Evol. Syst.* 37, 489-517.
- Yule, G.U., 1925. A mathematical theory of evolution, based on the conclusions of Dr. J. C. Willis, F.R.S. *Philos. Trans. Roy. Soc. Lond. B Biol. Sci.* 213, 21-87.
- Zhang, J., Kapli, P., Pavlidis, P., Stamatakis, A., 2013. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29, 2869-2876.
- Zinger, L., Philippe, H., 2016. Coalescing molecular evolution and DNA barcoding. *Mol. Ecol.* 25, 1908-1910.

Appendix S1

ORF haplotype	Type	Morphos-species	Location	Study	GenBank Accession number
ORF01	type 2	effusus	CLI	Pinzon et al 2013	HQ378759
	Clade IIIa	eydouxi, zelli	HAW, PHO	Marti-Puig et al 2014	KF583923, KF583931
	type B		HAW	Flot et al 2010	FR729327
	ORF01	fungiformis, ligulata	REU, MAD, CHE	This study	KX538981
ORF02	Clade IIIb	damicornis	HAW	Marti-Puig et al 2014	KF583937
ORF03	type 6a	ligulata	HAW	Pinzon et al 2013	JX994081
	Clade IIIb	ligulata, capitata, molokensis, damicornis	HAW	Marti-Puig et al 2014	KF583920, KF583921, KF583922, KF583924, KF583933
		ligulata	HAW	Flot et al 2010	EU374253
		damicornis	HAW	Gorospe et al 2013	KP698586
ORF04	Clade IIIb		HAW	Marti-Puig et al 2014	KF583941
ORF05	type 6b	ligulata	HAW	Pinzon et al 2013	JX994082
	Clade IIIb	molokensis	HAW (deep)	Marti-Puig et al 2014	KF583947
ORF06	sigma	damicornis	EA	Schmidt-Roach et al 2012	JX985614
ORF07	type 4c	damicornis	EA, JPN	Pinzon et al 2013	JX994077
ORF08	alpha	damicornis	COK, MOO	Mayfield et al (unpublished)	KJ720219
ORF09	type 4a	damicornis	TWN, EA, HER, LIZ, JPN	Pinzon et al 2013	JX994086
	alpha	damicornis	EA	Schmidt-Roach et al 2013	JX624991
	Clade Ib	damicornis	HAW	Marti-Puig et al 2014	KF583925, KF583946
		damicornis	HAW	Flot et al 2008	EU374235
		damicornis	MOO	Mayfield et al (unpublished)	KJ720218
		damicornis	HAW	Gorospe et al 2013	KP698587
			TWN	Hsu et al 2014	KM215098

	ORF09	damicornis	NCL, CHE	This study	KX538982
ORF10		damicornis	MOO	Mayfield et al (unpublished)	KJ720226
ORF11	ORF11	damicornis	NCL	This study	KX538983
ORF12	Clade Ib	damicornis	HAW	Marti-Puig et al 2014	KF583950
ORF13	type 4b	damicornis	HER	Pinzon et al 2013	JX994087
ORF14		damicornis	MOO	Mayfield et al (unpublished)	KJ720235
ORF15	alpha	damicornis	EA	Schmidt-Roach et al 2012	JX985618
ORF16	alpha	damicornis	EA	Schmidt-Roach et al 2013	JX625025
ORF17		damicornis	ROD	This study	KX538984
ORF18	type 5a	damicornis	RS, TZA, MAU, CHA, AND, THA, WA, HER, LIZ, PLW	Pinzon et al 2013	JX994073
	beta	damicornis	EA	Schmidt-Roach et al 2013	JX624999
	Clade Ia	damicornis	HAW, HOW	Marti-Puig et al 2014	KF583928, KF583935
	type a	damicornis	HAW	Flot et al 2008	EU374226
	type F		TZA, KEN	Souter et al 2009	FJ424111
	beta	damicornis	MOO, COK	Mayfield et al (unpublished)	KJ720240
			TWN	Hsu et al 2014	KM215075
		damicornis	HAW	Gorospe et al 2015	KP698585
	ORF18	damicornis, verrucosa, eydouxi	JDN, MAY, MAD, REU, ROD, NCL, CHE	This study	KX538985
ORF19	ORF19	damicornis	MAD	This study	KX538986
ORF20			TWN	Hsu et al 2014	KM215104
ORF21		damicornis	COK	Mayfield et al (unpublished)	KJ720241
ORF22	ORF22	elegans	NCL	This study	KX538987
ORF23	type 8a	damicornis	TWN	Pinzon et al 2013	JX994074
			COK	Mayfield et al (unpublished)	KJ720260
	ORF23	verrucosa, elegans, meandrina	NCL, CHE, MOO	This study	KX538988
ORF24	ORF24		MOO	This study	KX538989
ORF25	ORF25	damicornis	JDN	This study	KX538990

ORF26	ORF26	verrucosa	JDN	This study	KX538991
ORF27	type 1a	eydouxi, meandrina, verrucosa, damicornis	TZA, ZAN, AND, TWN, LIZ, PLW, HAW, PHO, CLI, MEX, PAN, GAL	Pinzon et al 2013	HQ378758
	e/m	eydouxi, meandrina	EA	Schmidt-Roach et al 2012	JX983184, JX985610
	type A	eydouxi, meandrina	HAW	Flot et al 2008	EU374247, EU374256, FR729394
		damicornis, verrucosa, meandrina, inflata		Paz-Garcia et al 2013	KF985974, KF985975, KF985976, KF985980
		damicornis	GAL	Baums et al 2014	KM610241
	clade IIb	eydouxi, zelli, meandrina, molokensis	HOW, JOH, NIH	Marti-Puig et al 2014	KF583930, KF583934, KF583948, KF583945
			TWN	Hsu et al 2014	KM215072
		meandrina	COK	Mayfield et al (unpublished)	KJ720263
	ORF27	eydouxi, woodjonesi, damicornis, zelli, meandrina	EUR, GLO, JDN, MAY, MAD, REU, ROD, TRO, NCL, CHE, MOO	This study	KX538992
ORF28	type 1b		HAW	Pinzon et al 2013	JX994088
ORF29	epsilon	damicornis	EA	Schmidt-Roach et al 2012	JX985611
ORF30		brevicornis	NCL	This study	KX538993
ORF31		damicornis	NCL	This study	KX538994
ORF32		sp	TWN	Hsu et al 2014	KM215121
ORF33		sp	TWN	Hsu et al 2014	KM215069
ORF34	type 7a	damicornis, verrucosa, meandrina	RS, AG, CHA, AND	Pinzon et al 2013	JX994084
	ORF34		TRO	This study	KX538995
ORF35	type 3h	damicornis, verrucosa, meandrina	RS	Pinzon et al 2013	JX994072
	ORF35	verrucosa	NCL	This study	KX538996
ORF36	type 3e	damicornis, verrucosa, meandrina	RS, AG, TZA, TWN, LIZ	Pinzon et al 2013	JX994083
	ORF36	verrucosa, damicornis	EUR, GLO, JDN, MAY, MAD, ROD	This study	KX538997
ORF37	ORF37	eydouxi	MAD	This study	KX538998
ORF38	ORF38	verrucosa	EUR, REU	This study	KX538999
ORF39	type 3c	damicornis, verrucosa, meandrina	RS, TZA, ZAN, CHA	Pinzon et al 2013	JX994075, FJ424129

	type NF	damicornis	TZA, KEN	Souter et al 2009	FJ424129
	ORF39	verrucosa, damicornis, eydouxi	EUR, GLO, JDN, MAY, MAD, REU, ROD, TRO, NCL, CHE	This study	KX539000
ORF40	ORF40	eydouxi	JDN	This study	KX539001
ORF41	ORF41	damicornis	MAD	This study	KX539002
ORF42	ORF42		NCL	This study	KX539003
ORF43	type 3g	damicornis, verrucosa, meandrina	RS, AG	Pinzon et al 2013	JX994076
	ORF43	damicornis	REU, NCL	This study	KX539004
ORF44	ORF44		NCL	This study	KX539005
ORF45	type 3j	damicornis, verrucosa, meandrina	RS	Pinzon et al 2013	JX994080
ORF46	type 3a	damicornis, verrucosa, meandrina	ZAN, CHA, TWN, LIZ, PLW, HAW	Pinzon et al 2013	HQ378760
	gamma	damicornis, verrucosa		Schmidt-Roach et al 2012	JX985590
		molokensis	HAW	Flot et al 2008	EU374261
	type 3a	damicornis	GAL	Baums et al 2014	KM610277
				Hsu et al 2014	KM215107
	ORF46	eydouxi, kelleheri, damicornis	EUR, GLO, JDN, MAY, MAD, REU, ROD, NCL	This study	KX539006
ORF47	type 3b	damicornis, verrucosa, meandrina	ZAN, WA, LIZ, PLW, GAL	Pinzon et al 2013	HQ378761
	Clade IIa	zelli, verrucosa	HOW	Marti-Puig et al 2014	KF583936, KF583943
			TWN	Hsu et al 2014	KM215102
	ORF47	verrucosa, kelleheri, ankeli, damicornis	NCL, CHE, TON	This study	KX539007
ORF48	type 3i	damicornis, verrucosa, meandrina	RS	Pinzon et al 2013	JX994078
ORF49	ORF49		MOO	This study	KX539008
ORF50		verrucosa	COK	Mayfield et al (unpublished)	KJ720271
		verrucosa	MOO	Forsman et al 2013	KF381330
	ORF50		MOO, BOR, TAH	This study	KX539009
ORF51	ORF51	damicornis, verrucosa	NCL	This study	KX539010
ORF52	ORF52	damicornis	NCL	This study	KX539011

ORF53	type 3f	damicornis, verrucosa, meandrina	TWN, LIZ	Pinzon et al 2013	JX994079
			TWN	Hsu et al 2014	KM215070
	ORF53	ankeli, ligulata, kelleheri	NCL, CHE, TON	This study	KX539012
ORF54	type 3d	damicornis, verrucosa, meandrina	AND, LIZ, PLW	Pinzon et al 2013	JX994085
	gamma	damicornis, verrucosa	EA	Schmidt-Roach et al 2012	JX985587, JX985612
			TWN	Hsu et al 2014	KM215094
	ORF54	verrucosa	NCL	This study	KX539013
ORF55	x	verrucosa		Schmidt-Roach et al 2014	KF709244
ORF56			MOO	Forsman et al 2013	KF381329
ORF57			MOO	Forsman et al 2013	KF381328

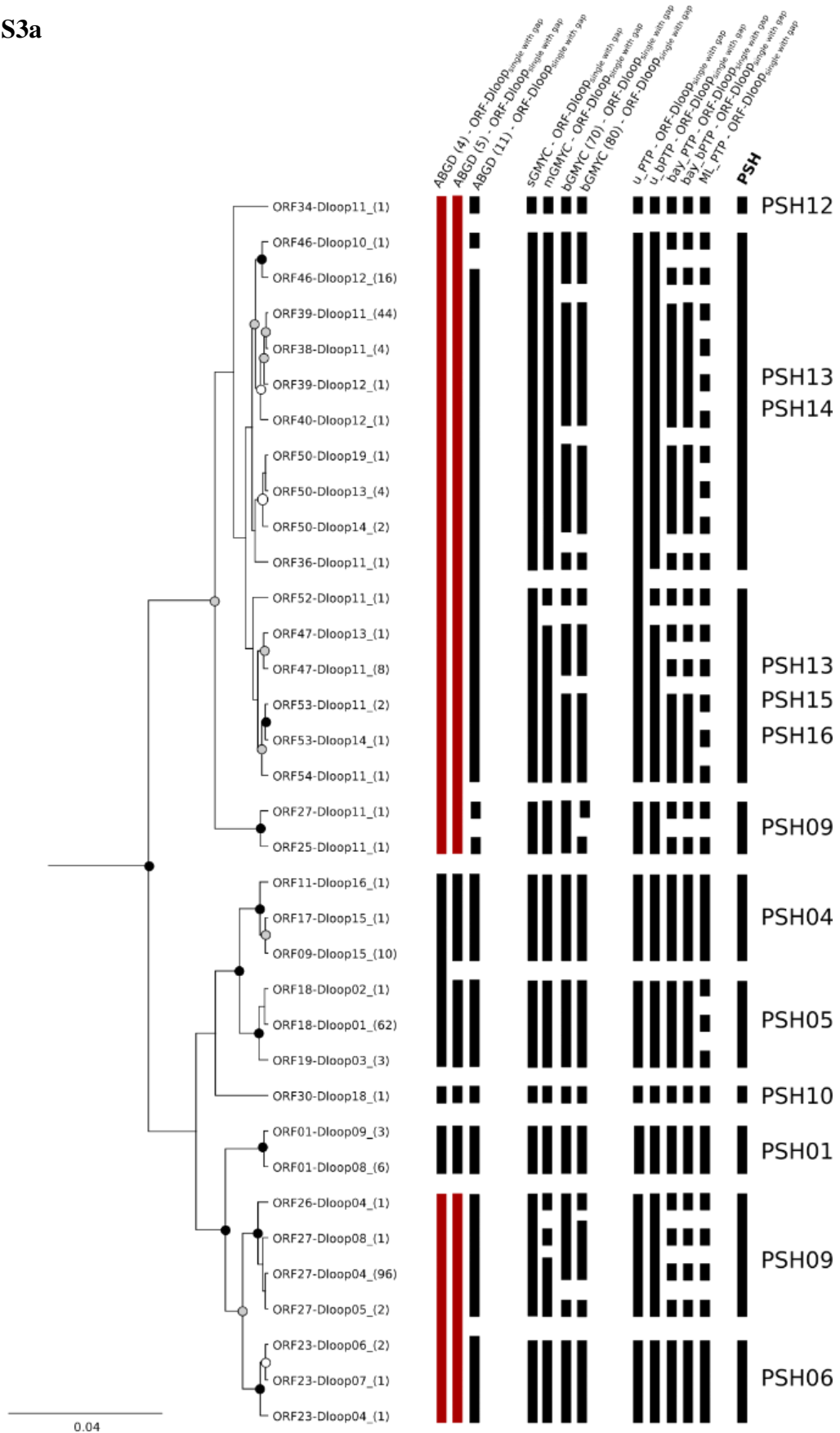
Appendix S2

Marker	Haplotype	Study	Accession
Dloop	Dloop18	Flot et al 2008	EU374262
Dloop	Dloop19	Flot et al 2008	EU374272
Dloop	Dloop20	Flot et al 2008	EU374285
Dloop	Dloop21	Schmidt-Roach et al 2013	JX625105
Dloop	Dloop22	Schmidt-Roach et al 2013	JX625106
Dloop	Dloop23	Schmidt-Roach et al 2013	JX625114
Dloop	Dloop24	Marti-Puig et al 2014	KF583905
Dloop	Dloop25	Marti-Puig et al 2014	KF583907
Dloop	Dloop26	Marti-Puig et al 2014	KF583918
Dloop	Dloop27	Marti-Puig et al 2014	KF583914
Dloop	Dloop28	Marti-Puig et al 2014	KF583910
Dloop	Dloop29	Marti-Puig et al 2014	KF583911
Dloop	Dloop30	Marti-Puig et al 2014	KF583912
Dloop	Dloop31	Marti-Puig et al 2014	KF583919
Dloop	Dloop01	This study	KX539014
Dloop	Dloop02	This study	KX539015
Dloop	Dloop03	This study	KX539016
Dloop	Dloop04	This study	KX539017
Dloop	Dloop05	This study	KX539018
Dloop	Dloop06	This study	KX539019
Dloop	Dloop07	This study	KX539020
Dloop	Dloop08	This study	KX539021
Dloop	Dloop09	This study	KX539022
Dloop	Dloop10	This study	KX539023
Dloop	Dloop11	This study	KX539024
Dloop	Dloop12	This study	KX539025
Dloop	Dloop13	This study	KX539026
Dloop	Dloop14	This study	KX539027
Dloop	Dloop15	This study	KX539028
Dloop	Dloop16	This study	KX539029
Dloop	Dloop17	This study	KX539030
ITS2	ITS2-001	This study	KX539031
ITS2	ITS2-002	This study	KX539032
ITS2	ITS2-003	This study	KX539033
ITS2	ITS2-004	This study	KX539034
ITS2	ITS2-005	This study	KX539035
ITS2	ITS2-006	This study	KX539036
ITS2	ITS2-007	This study	KX539037
ITS2	ITS2-008	This study	KX539038
ITS2	ITS2-009	This study	KX539039
ITS2	ITS2-010	This study	KX539040
ITS2	ITS2-011	This study	KX539041
ITS2	ITS2-012	This study	KX539042
ITS2	ITS2-013	This study	KX539043
ITS2	ITS2-014	This study	KX539044
ITS2	ITS2-015	This study	KX539045
ITS2	ITS2-016	This study	KX539046
ITS2	ITS2-017	This study	KX539047
ITS2	ITS2-018	This study	KX539048
ITS2	ITS2-019	This study	KX539049
ITS2	ITS2-020	This study	KX539050
ITS2	ITS2-021	This study	KX539051
ITS2	ITS2-022	This study	KX539052
ITS2	ITS2-023	This study	KX539053
ITS2	ITS2-024	This study	KX539054
ITS2	ITS2-025	This study	KX539055
ITS2	ITS2-026	This study	KX539056
ITS2	ITS2-027	This study	KX539057
ITS2	ITS2-028	This study	KX539058
ITS2	ITS2-029	This study	KX539059
ITS2	ITS2-030	This study	KX539060
ITS2	ITS2-031	This study	KX539061
ITS2	ITS2-032	This study	KX539062

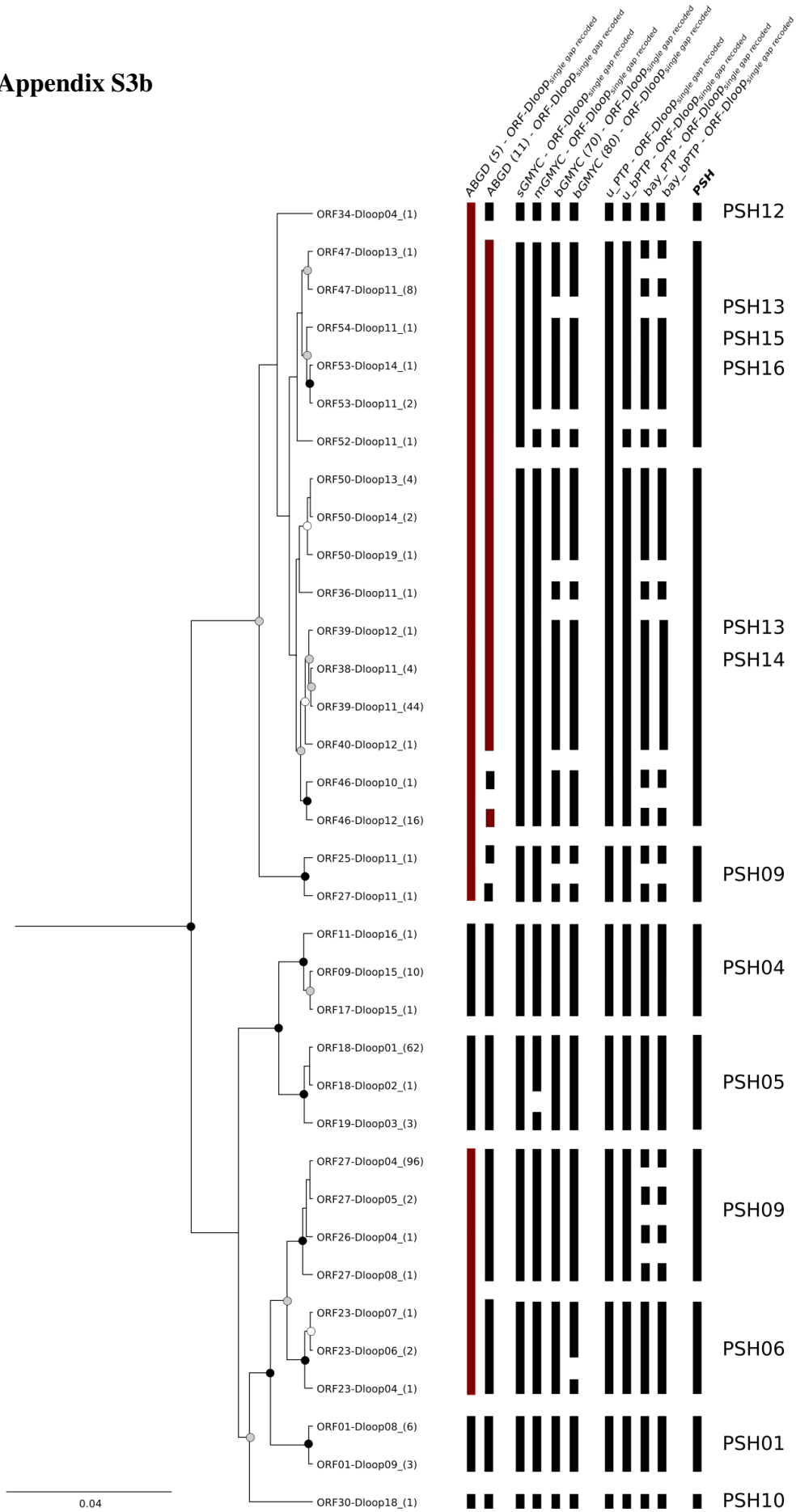
ITS2	ITS2-033	This study	KX539063
ITS2	ITS2-034	This study	KX539064
ITS2	ITS2-035	This study	KX539065
ITS2	ITS2-036	This study	KX539066
ITS2	ITS2-037	This study	KX539067
ITS2	ITS2-038	This study	KX539068
ITS2	ITS2-039	This study	KX539069
ITS2	ITS2-040	This study	KX539070
ITS2	ITS2-041	This study	KX539071
ITS2	ITS2-042	This study	KX539072
ITS2	ITS2-043	This study	KX539073
ITS2	ITS2-044	This study	KX539074
ITS2	ITS2-045	This study	KX539075
ITS2	ITS2-046	This study	KX539076
ITS2	ITS2-047	This study	KX539077
ITS2	ITS2-048	This study	KX539078
ITS2	ITS2-049	This study	KX539079
ITS2	ITS2-050	This study	KX539080
ITS2	ITS2-051	This study	KX539081
ITS2	ITS2-052	This study	KX539082
ITS2	ITS2-053	This study	KX539083
ITS2	ITS2-054	This study	KX539084
ITS2	ITS2-055	This study	KX539085
ITS2	ITS2-056	This study	KX539086
ITS2	ITS2-057	This study	KX539087
ITS2	ITS2-058	This study	KX539088
ITS2	ITS2-059	This study	KX539089
ITS2	ITS2-060	This study	KX539090
ITS2	ITS2-061	This study	KX539091
ITS2	ITS2-062	This study	KX539092
ITS2	ITS2-063	This study	KX539093
ITS2	ITS2-064	This study	KX539094
ITS2	ITS2-065	This study	KX539095
ITS2	ITS2-066	This study	KX539096
ITS2	ITS2-067	This study	KX539097
ITS2	ITS2-068	This study	KX539098
ITS2	ITS2-069	This study	KX539099
ITS2	ITS2-070	This study	KX539100
ITS2	ITS2-071	This study	KX539101
ITS2	ITS2-072	This study	KX539102
ITS2	ITS2-073	This study	KX539103
ITS2	ITS2-074	This study	KX539104
ITS2	ITS2-075	This study	KX539105
ITS2	ITS2-076	This study	KX539106
ITS2	ITS2-077	This study	KX539107
ITS2	ITS2-078	This study	KX539108
ITS2	ITS2-079	This study	KX539109
ITS2	ITS2-080	This study	KX539110
ITS2	ITS2-081	This study	KX539111
ITS2	ITS2-082	This study	KX539112
ITS2	ITS2-083	This study	KX539113
ITS2	ITS2-084	This study	KX539114
ITS2	ITS2-085	This study	KX539115
ITS2	ITS2-086	This study	KX539116
ITS2	ITS2-087	This study	KX539117
ITS2	ITS2-088	This study	KX539118
ITS2	ITS2-089	This study	KX539119
ITS2	ITS2-090	This study	KX539120
ITS2	ITS2-091	This study	KX539121
ITS2	ITS2-092	This study	KX539122
ITS2	ITS2-093	This study	KX539123
ITS2	ITS2-094	This study	KX539124
ITS2	ITS2-095	This study	KX539125
ITS2	ITS2-096	This study	KX539126

ITS2	ITS2-097	This study	KX539127
ITS2	ITS2-098	This study	KX539128
ITS2	ITS2-099	This study	KX539129
ITS2	ITS2-100	This study	KX539130
ITS2	ITS2-101	This study	KX539131
ITS2	ITS2-102	This study	KX539132
ITS2	ITS2-103	This study	KX539133
ITS2	ITS2-104	This study	KX539134
ITS2	ITS2-105	This study	KX539135
ITS2	ITS2-106	This study	KX539136
ITS2	ITS2-107	This study	KX539137
ITS2	ITS2-108	This study	KX539138
ITS2	ITS2-109	This study	KX539139
ITS2	ITS2-110	This study	KX539140
ITS2	ITS2-111	This study	KX539141
ITS2	ITS2-112	This study	KX539142
ITS2	ITS2-113	This study	KX539143
ITS2	ITS2-114	This study	KX539144
ITS2	ITS2-115	This study	KX539145
ITS2	ITS2-116	This study	KX539146
ITS2	ITS2-117	This study	KX539147
ITS2	ITS2-118	This study	KX539148
ITS2	ITS2-119	This study	KX539149
ITS2	ITS2-120	This study	KX539150
ITS2	ITS2-121	This study	KX539151
ITS2	ITS2-122	This study	KX539152
ITS2	ITS2-123	This study	KX539153
ITS2	ITS2-124	This study	KX539154
ITS2	ITS2-125	This study	KX539155
ITS2	ITS2-126	This study	KX539156
ITS2	ITS2-127	This study	KX539157
ITS2	ITS2-128	This study	KX539158

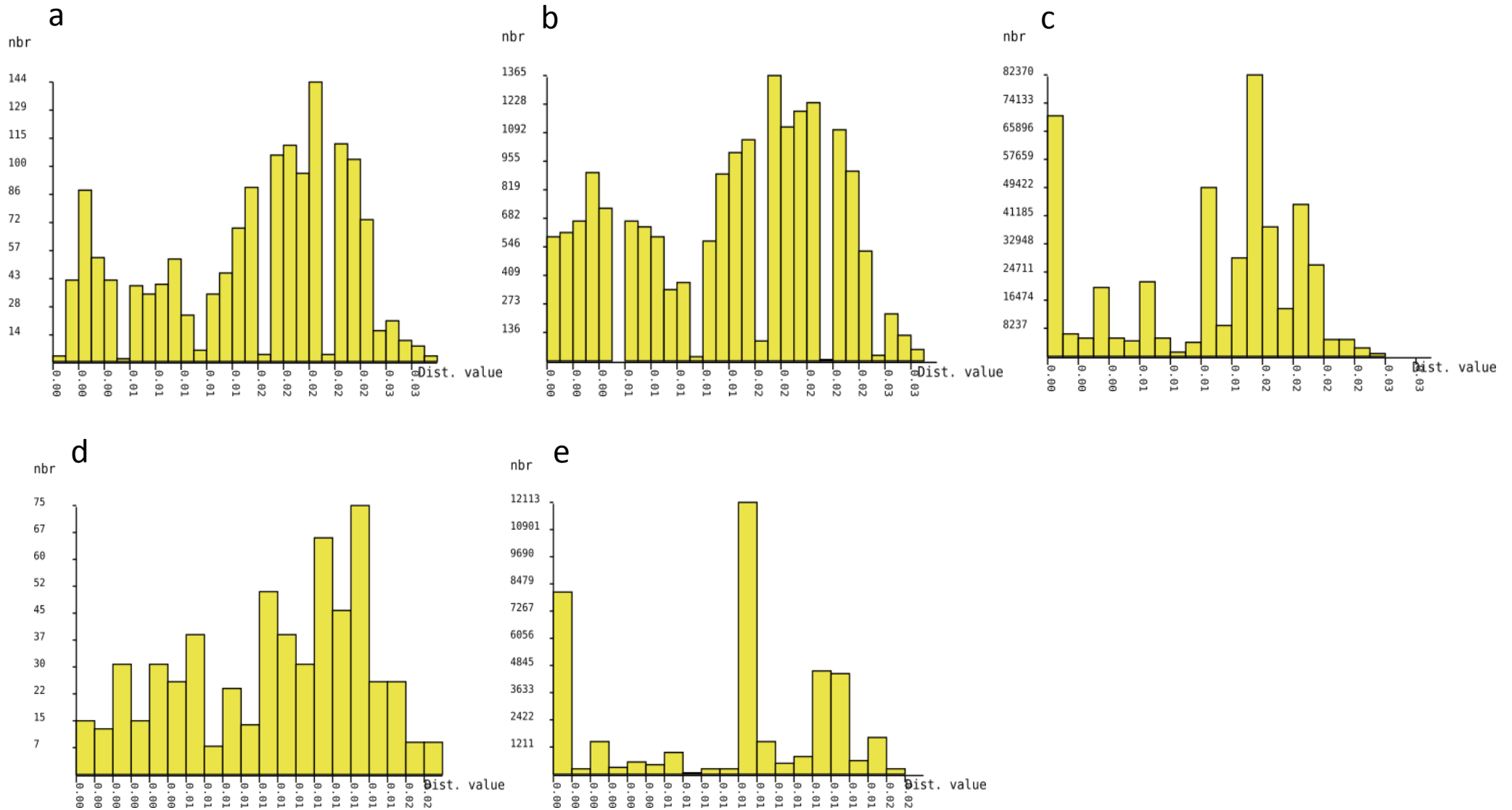
Appendix S3a



Appendix S3b

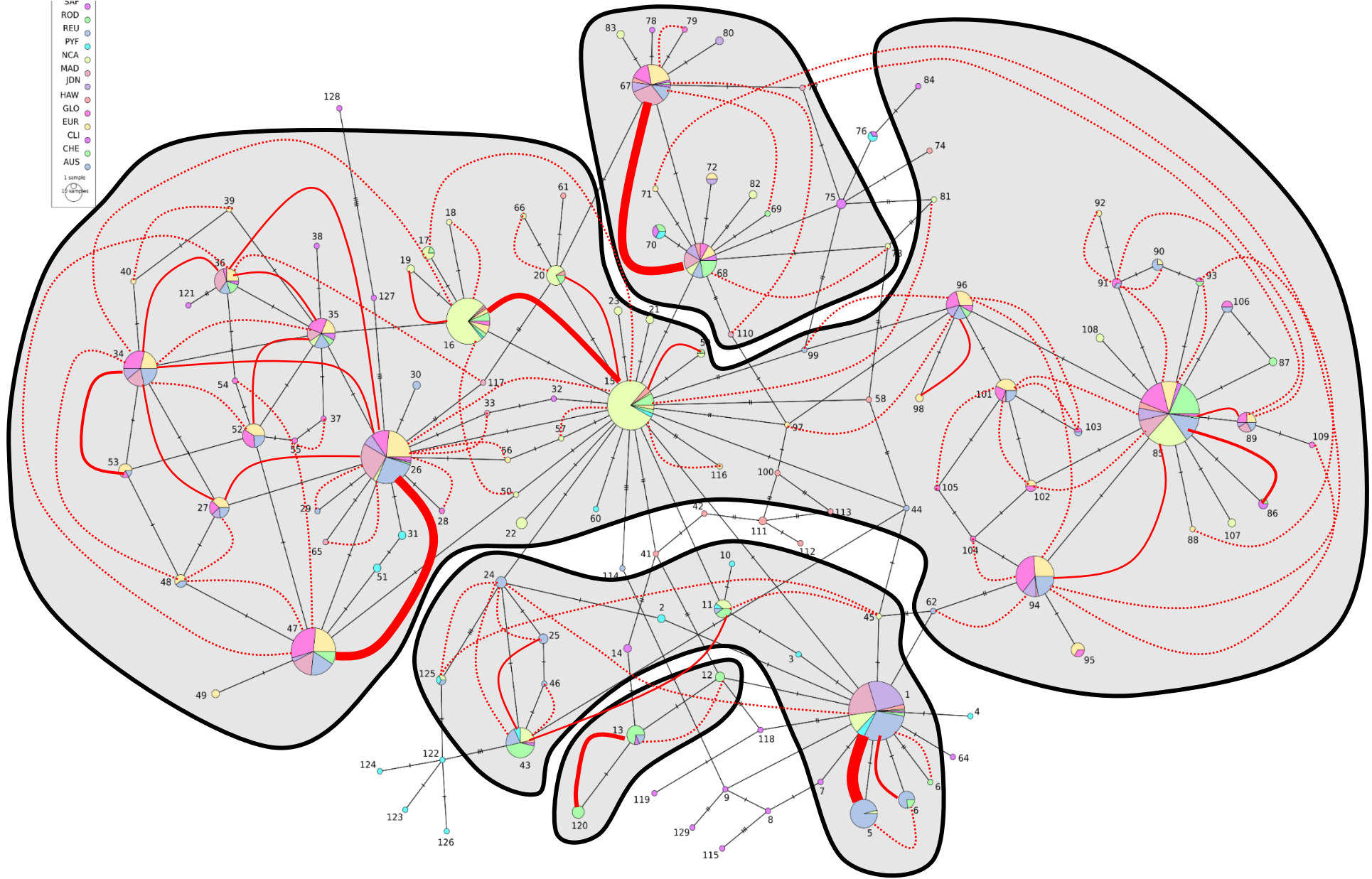


Appendix S4

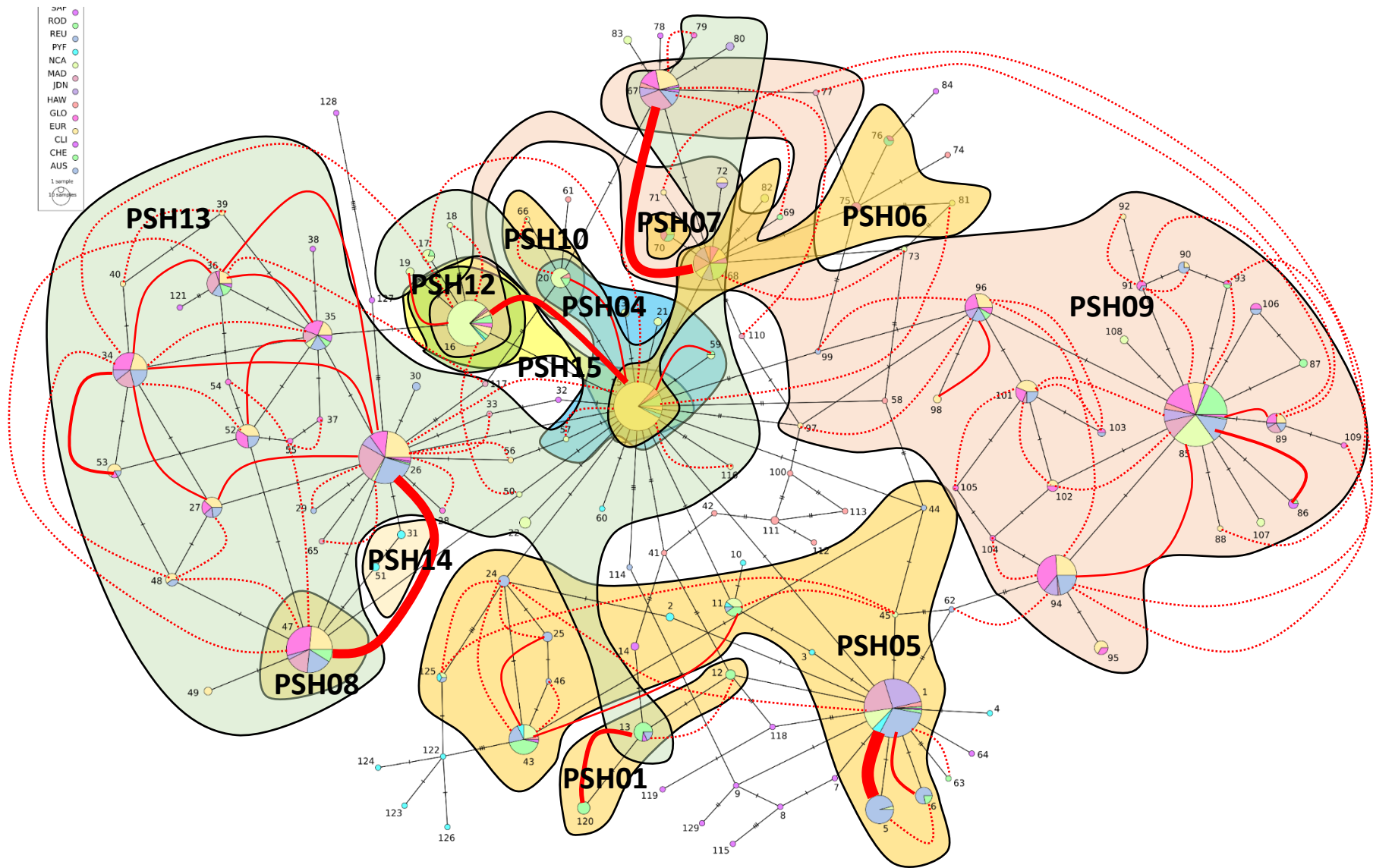


Appendix S5a

- MAF
 - ROD
 - REU
 - PYF
 - NCA
 - MAD
 - JDN
 - HAW
 - GLO
 - EUR
 - CLI
 - CHE
 - AUS
- 1 sample
16 samples



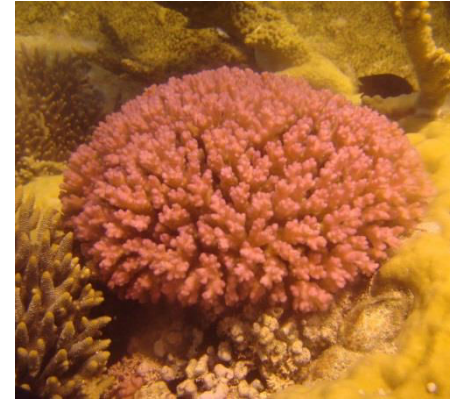
Appendix S5b



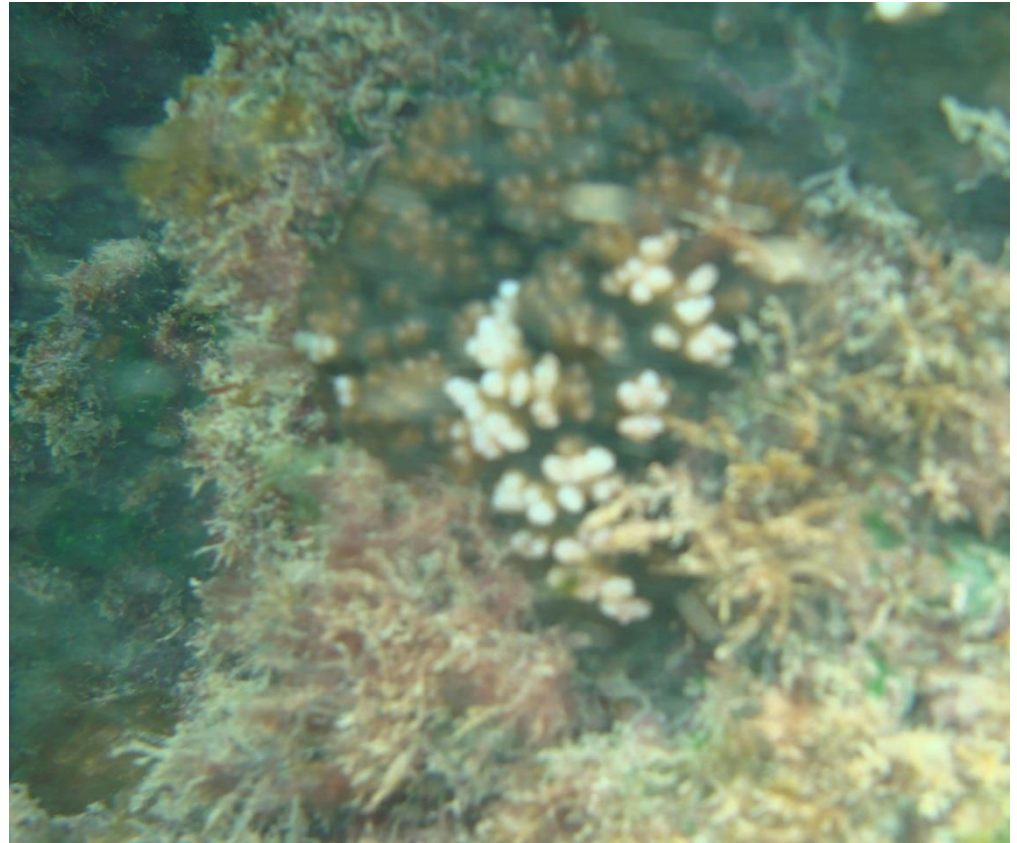
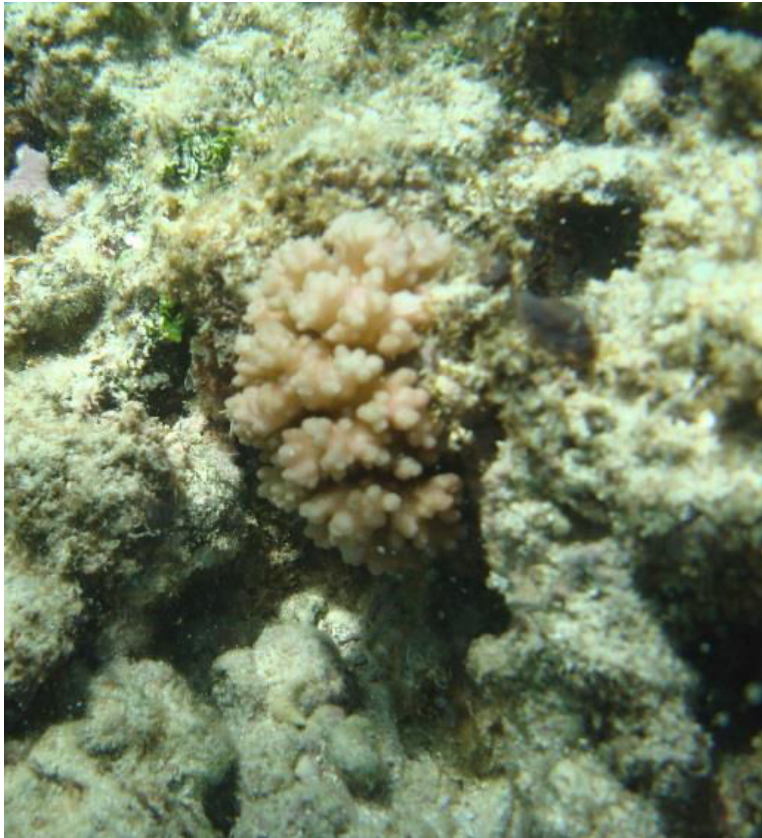
Appendix S6a - SSH01 - ORF01 (10) - *Pocillopora* sp. [undetermined (*P. fungiformis*-like, *P. effusus*-like, *P. ligulata*-like)]



Appendix S6b – SSH04 – ORF09 (98) - *P. damicornis* (type α)



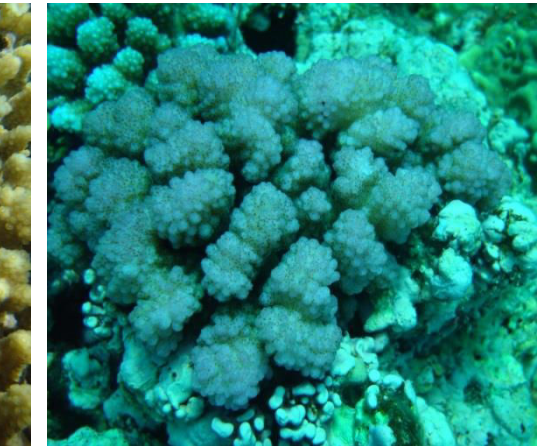
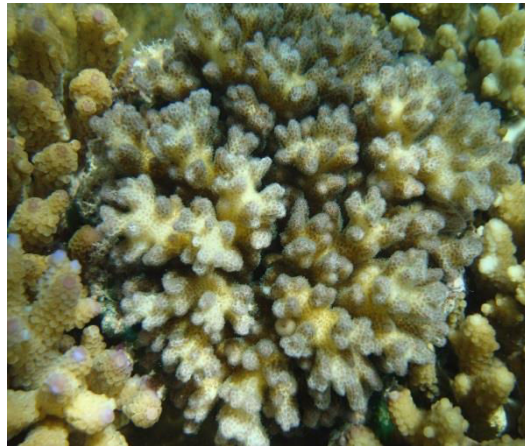
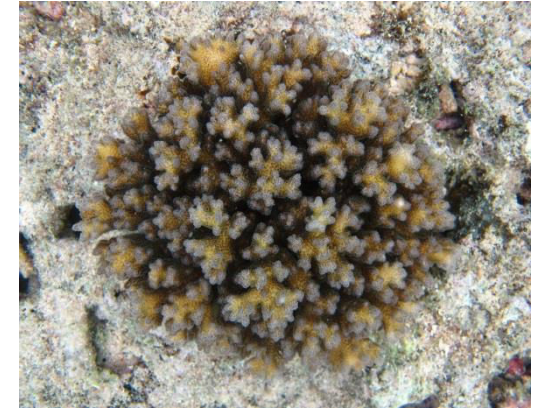
Appendix S6c – SSH04 - ORF11 (2) - *P. damicornis* (type α)



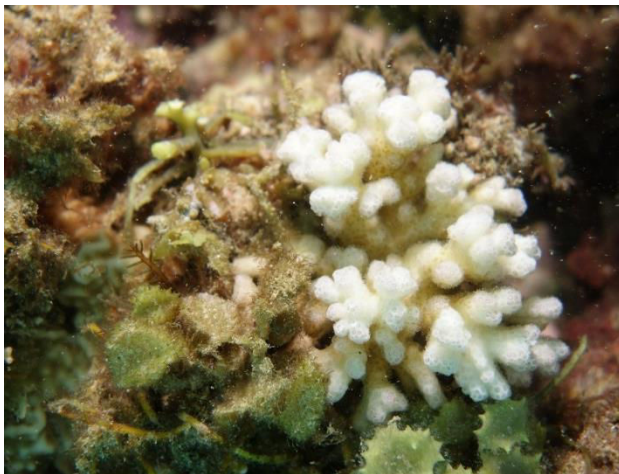
Appendix S6d – SSH04 - ORF17 (2) - *P. damicornis* (type α)



Appendix S6e – SSH05 – ORF18 (208) - *P. acuta* (type β)



Appendix S6f – SSH05 – ORF19 (6) - *P. acuta* (type β)



Appendix S6g – SSH06 – ORF22 (1) - *Pocillopora* sp.



Appendix S6h – SSH06 – ORF23 (8) - *Pocillopora* sp.



Appendix S6i – SSH08 – ORF25 (1) & ORF26 (1) - *Pocillopora* sp.



ORF25

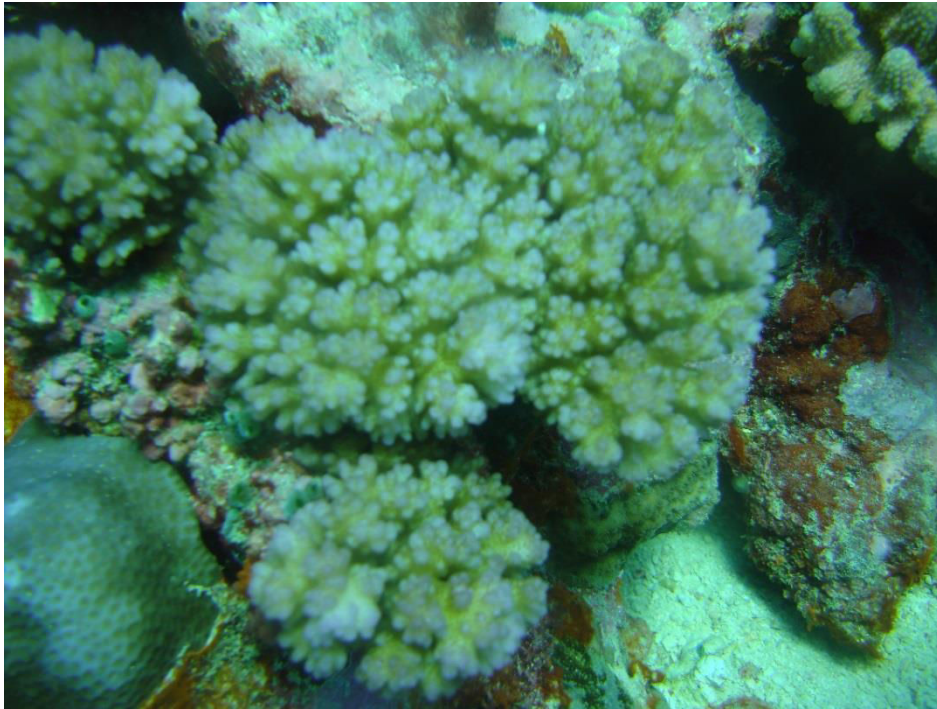


ORF26

Appendix S6j– SSH09 – ORF27 (241) - *P. eydouxi/meandrina* (type 1)



Appendix S6k - SSH10 – ORF30 (1) & ORF31 (1) - *P. brevicornis* (type ϵ)



ORF30



ORF31

Appendix S61 – SSH13a – ORF35 (1)



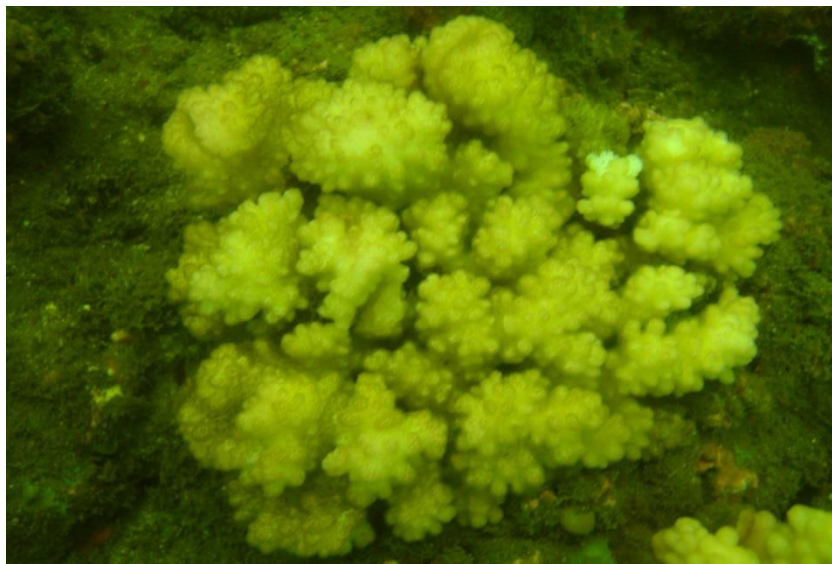
Appendix S6m – SSH13a – ORF36 (11) - *Pocillopora* sp.



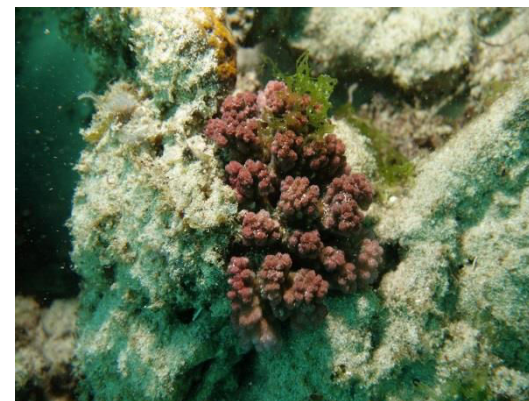
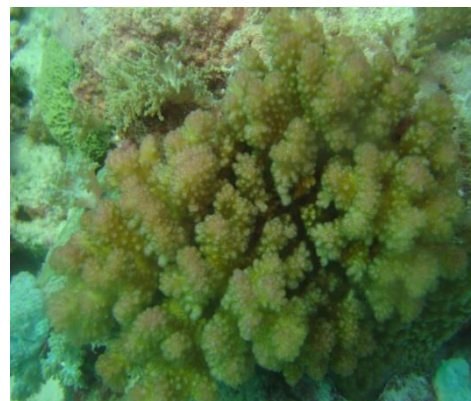
Appendix S6n – SSH13a – ORF37 (1) - *Pocillopora* sp.



Appendix S6o – SSH13a – ORF38 (4) - *Pocillopora* sp.



Appendix S6p – SSH13a – ORF39 (161) - *Pocillopora* sp.



Appendix S6q – SSH13a – ORF40 (1) - *Pocillopora* sp.



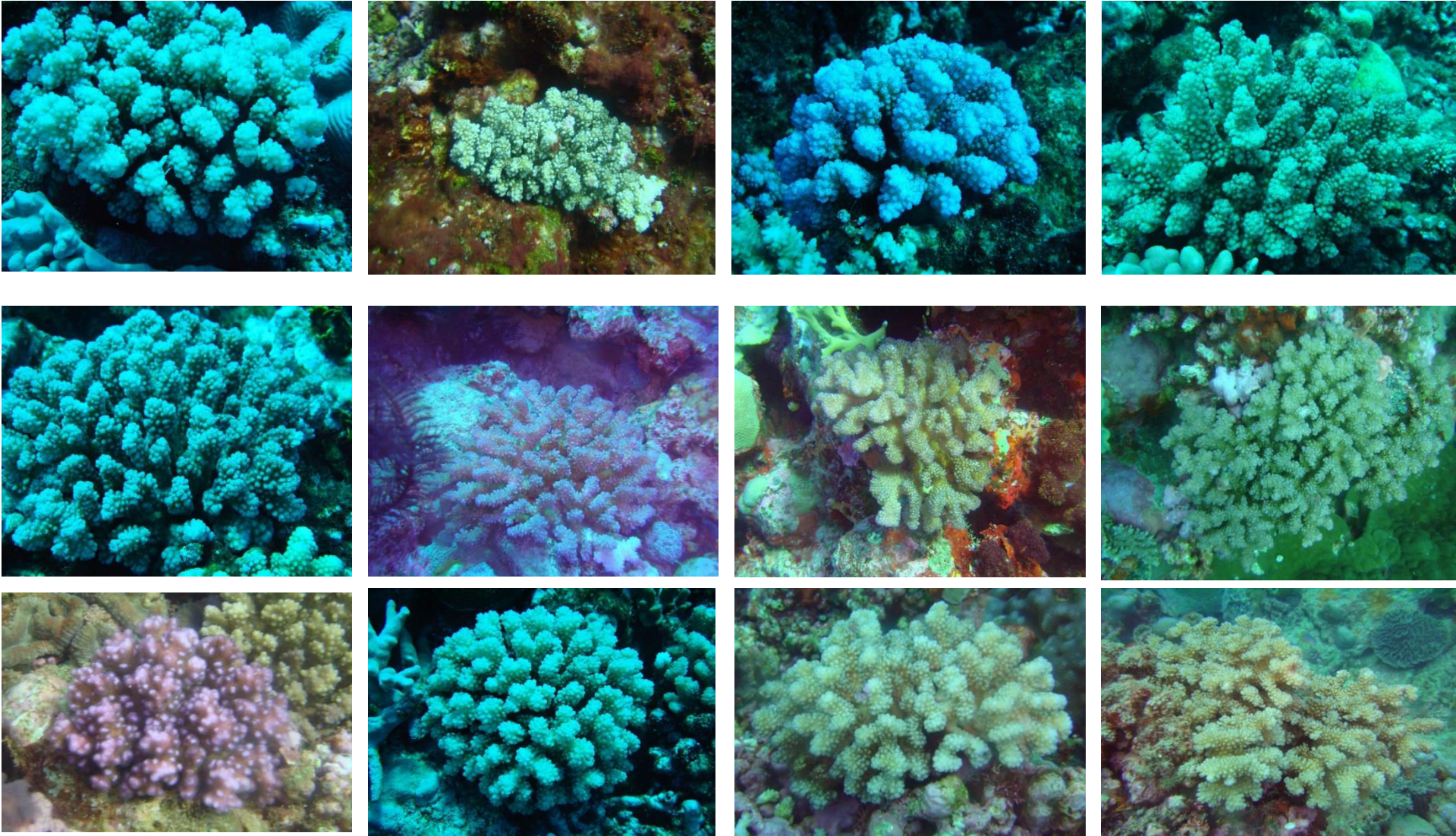
Appendix S6r – SSH13a – ORF41 (1) - *Pocillopora* sp.



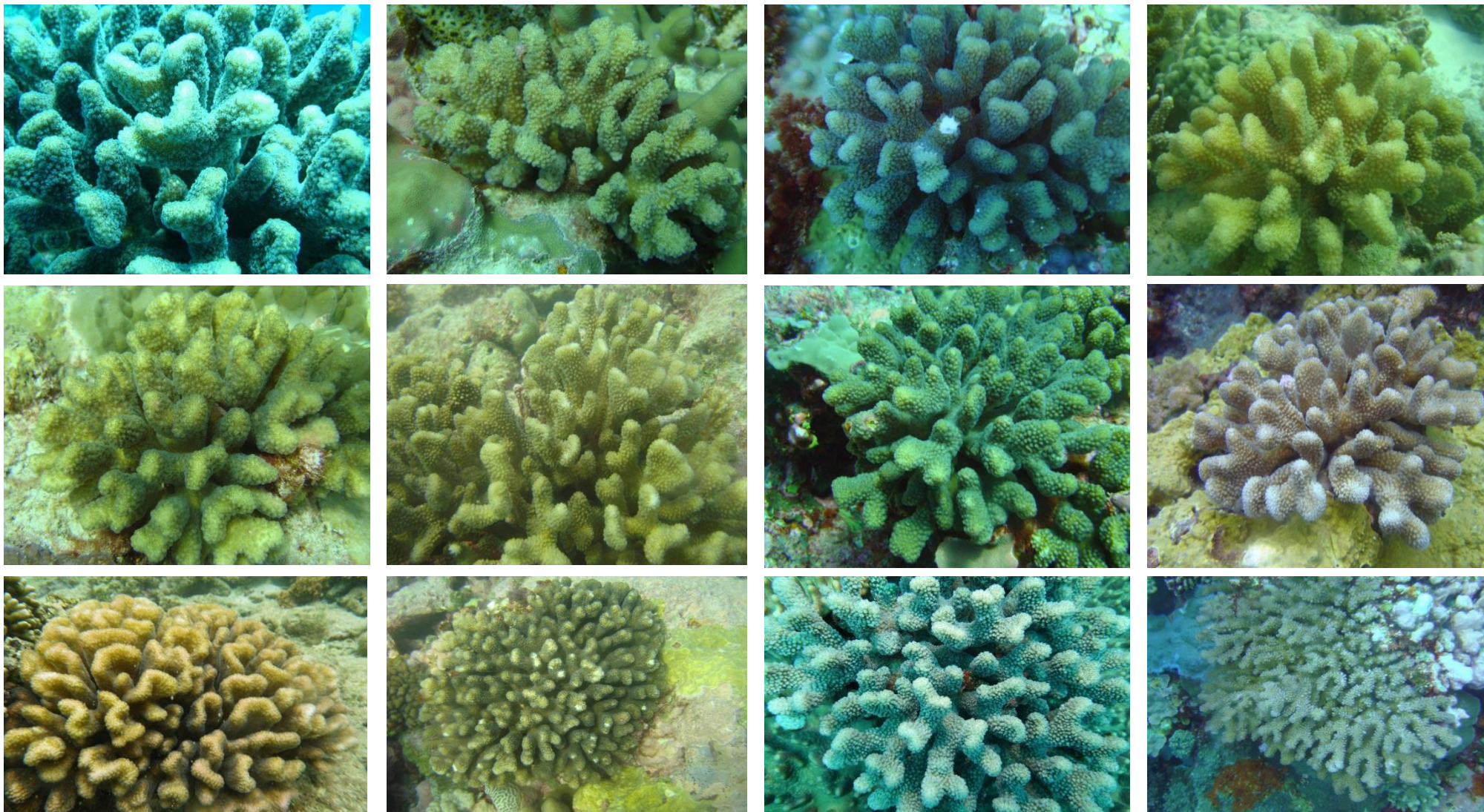
Appendix S6s – SSH13a – ORF42 (1) - *Pocillopora* sp.



Appendix S6t – SSH13c – ORF47 (51) - *Pocillopora* sp.



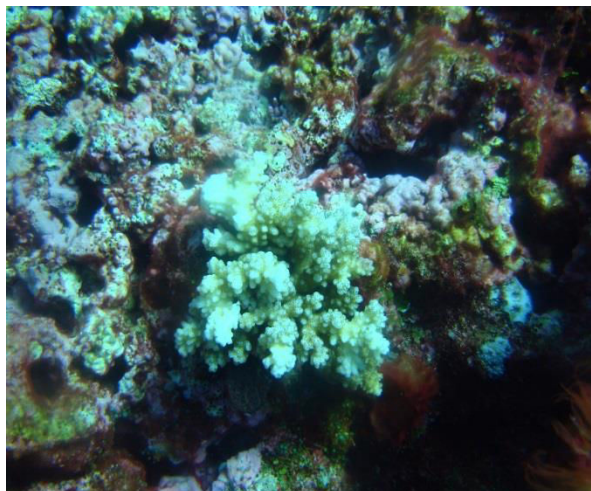
Appendix S6u – SSH13b – ORF46 (42) - *Pocillopora* sp.



Appendix S6v – SSH15 – ORF51 (2) & ORF52 (3)



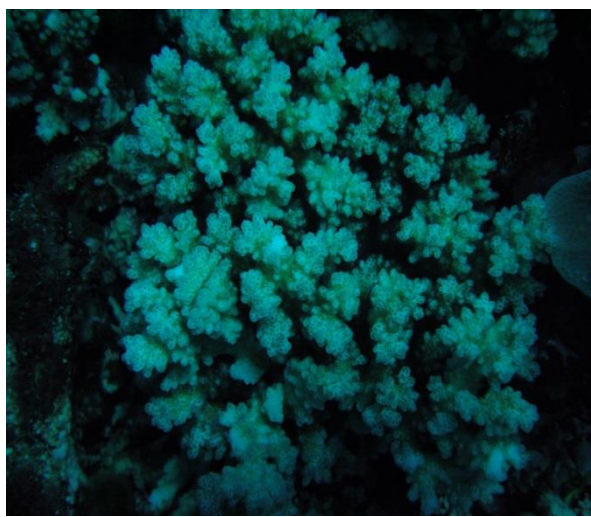
ORF51



ORF51



ORF52

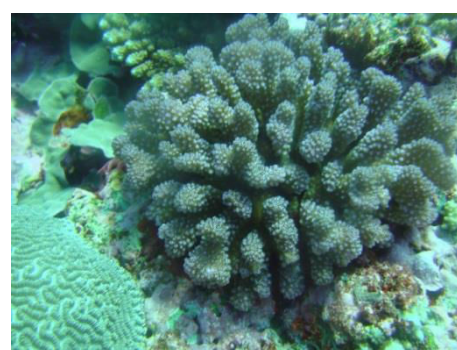
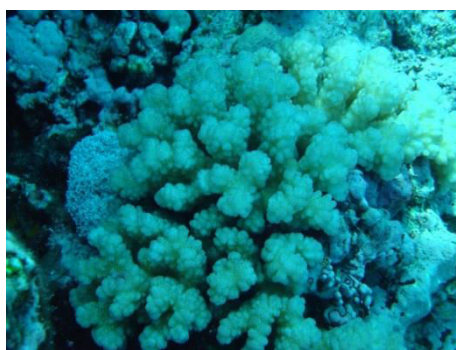
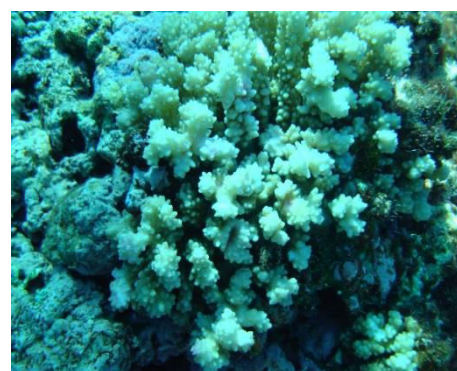
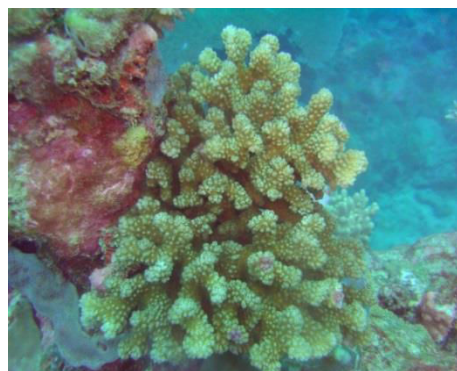


ORF52

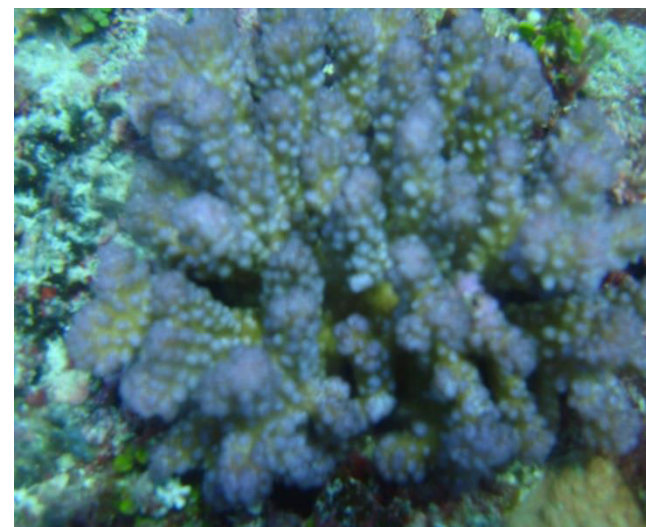
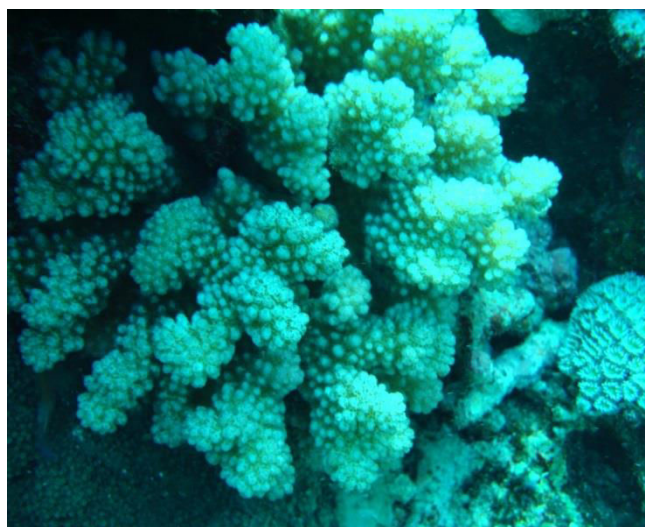


ORF52

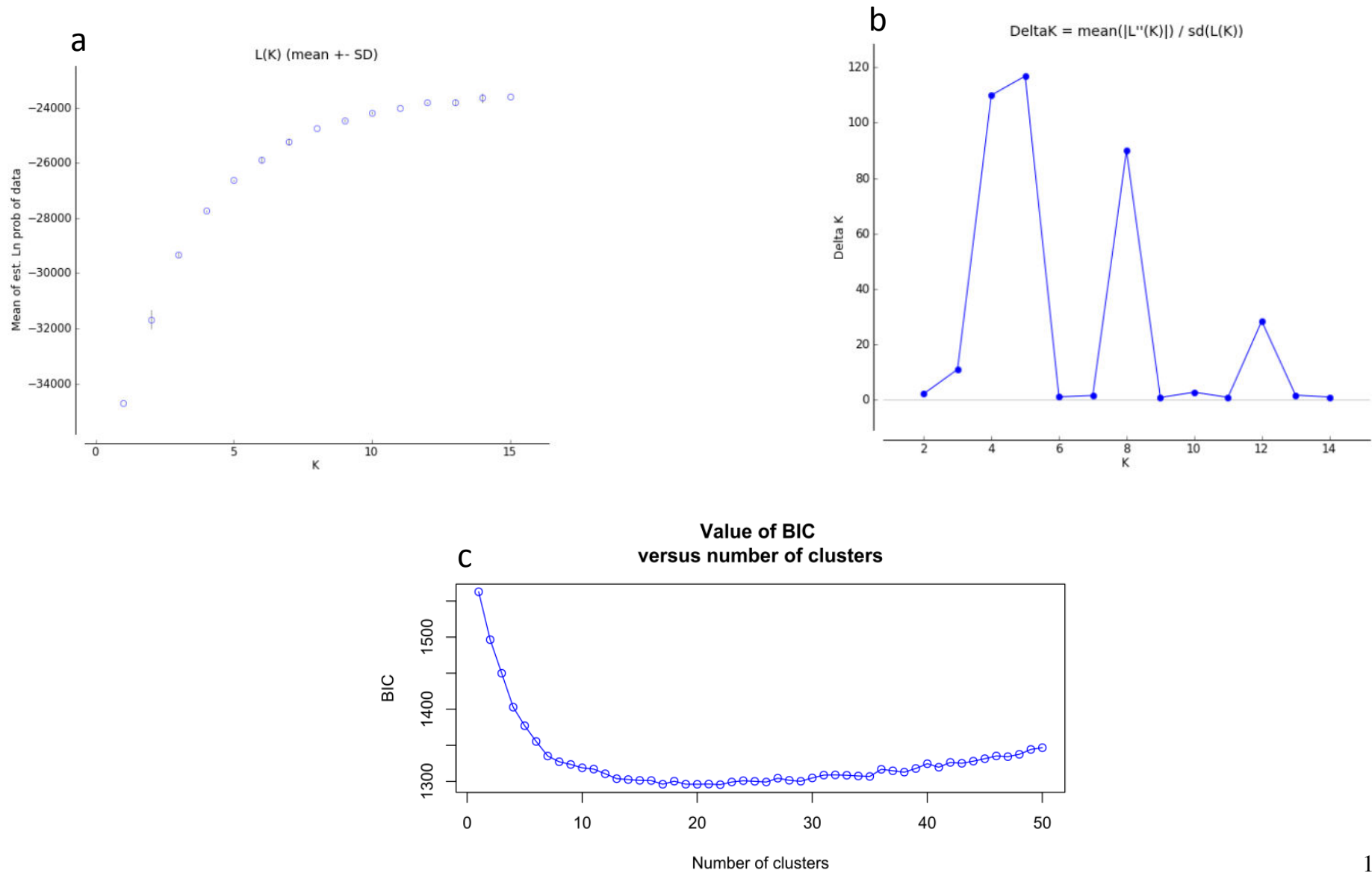
Appendix S6w – SSH16 – ORF53 (14) - *Pocillopora* sp.



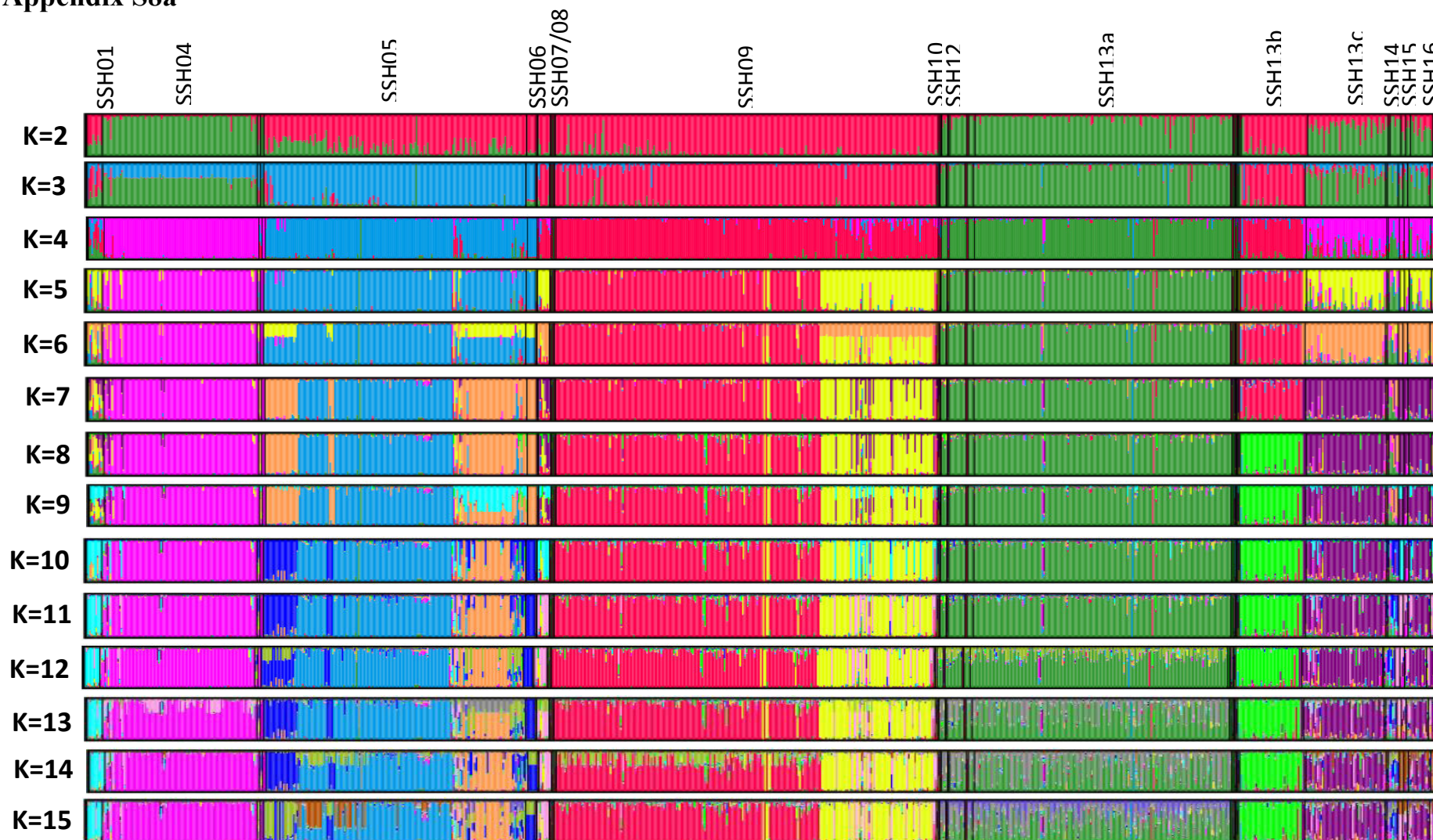
Appendix S6x – SSH16 – ORF54 (5) - *Pocillopora* sp.



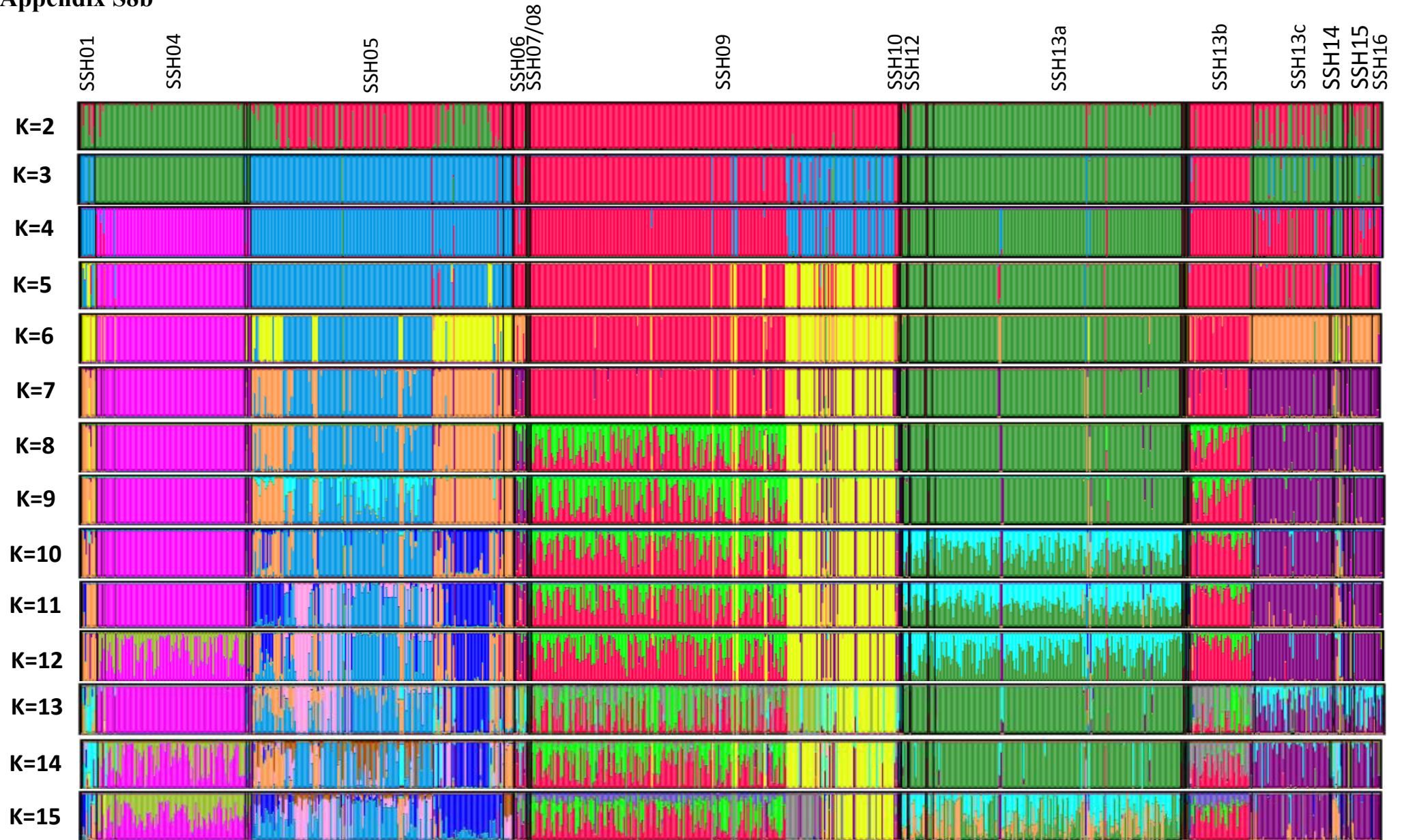
Appendix S7



Appendix S8a



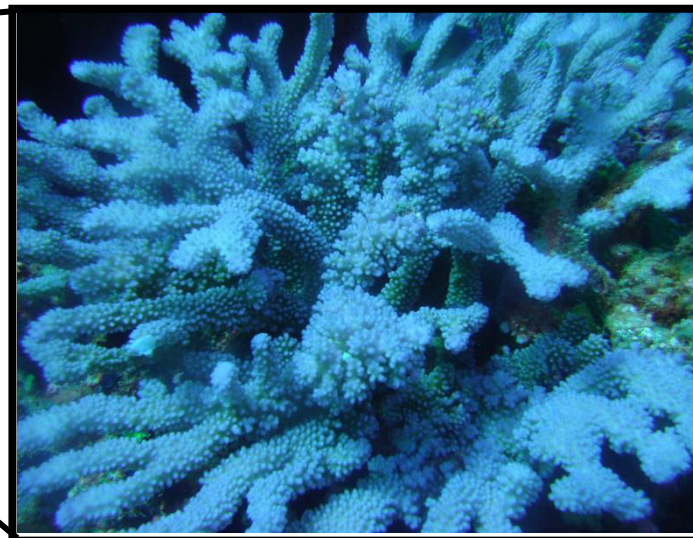
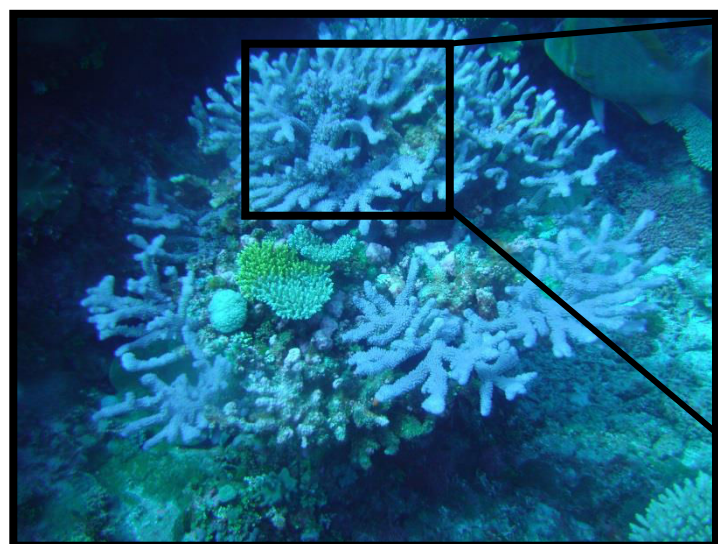
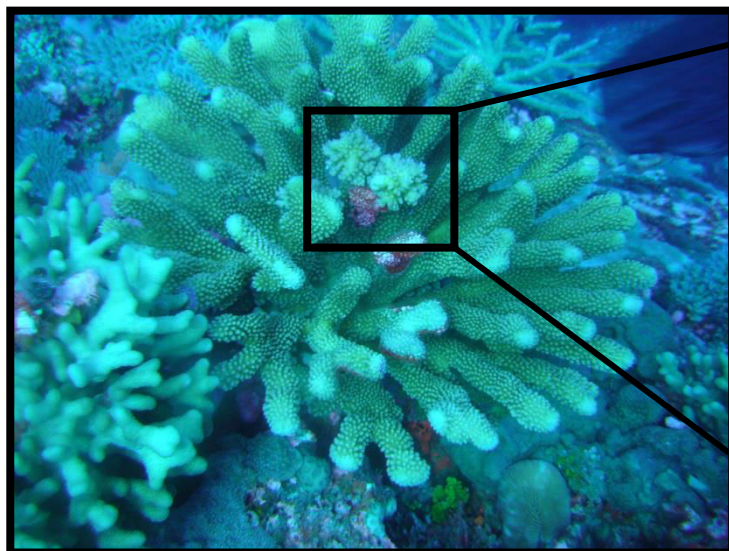
Appendix S8b



Appendix S9



Appendix S10



Appendix S11

Hap	WIO								TSP				SEP			Total
	GLO	MAY	JDN	MAD	EUR	REU	ROD	TRO	CHE	NCL	LOY	TON	BOR	MOO	TAH	
1				1		1			8							10
2																0
3																0
4																0
5																0
6																0
7																0
8																0
9									5	92	1					98
10																0
11										2						2
12																0
13																0
14																0
15																0
16																0
17											2					2
18		12	19	7		99	36	1	46							220
19				6												6
20																0
21																0
22											1					1
23								2	4	1			1			8
24													1			1
25			1													1
26			1													1
27	45	16	17	19	42	34	9	3	20	41	10		2			258
28																0
29																0
30										1						1
31										1						1
32																0
33																0
34								4								4
35									1							1
36	2	2	1	3	4		1									13
37				1												1
38					1	3										4
39	37	9	7	33	43	33	8	2								172
40			1													1
41				1												1
42										1						1
43						1										1
44										1						1

45																		0
46	7	8	6	9	10	5	3		2									50
47									2	39	8	2						51
48																		0
49																	1	1
50														2	3	2		7
51											2							2
52										3								3
53									1	9	3	1						14
54										4	1							5
55																		0
56																		0
57																		0
Total	91	47	53	80	100	176	59	9	39	245	29	3		2	8	2		943

Appendix S12

ORF haplotype	Dloop haplotype	ITS2_hap1	ITS2_hap2
ORF01	Dloop08	ITS2-031	ITS2-031
ORF01	Dloop08	ITS2-031	ITS2-037
ORF01	Dloop08	ITS2-046	ITS2-046
ORF01	Dloop09	ITS2-031	ITS2-031
ORF01	Dloop09	ITS2-031	ITS2-037
ORF09	Dloop15		
ORF09		ITS2-003	ITS2-003
ORF09		ITS2-028	ITS2-028
ORF09		ITS2-054	ITS2-054
ORF09	Dloop15	ITS2-003	ITS2-003
ORF09	Dloop15	ITS2-003	ITS2-062
ORF09	Dloop15	ITS2-055	ITS2-055
ORF11		ITS2-003	ITS2-028
ORF11	Dloop16	ITS2-003	ITS2-028
ORF17		ITS2-003	ITS2-003
ORF17	Dloop15	ITS2-003	ITS2-091
ORF18	Dloop01		
ORF18	Dloop02		
ORF18	Dloop17		
ORF18		ITS2-001	ITS2-001
ORF18		ITS2-011	ITS2-001
ORF18		ITS2-017	ITS2-017
ORF18		ITS2-017	ITS2-029
ORF18		ITS2-022	ITS2-022
ORF18		ITS2-027	ITS2-001
ORF18		ITS2-027	ITS2-027
ORF18		ITS2-029	ITS2-029
ORF18		ITS2-072	ITS2-050
ORF18	Dloop01	ITS2-001	ITS2-001
ORF18	Dloop01	ITS2-011	ITS2-001
ORF18	Dloop01	ITS2-011	ITS2-011
ORF18	Dloop01	ITS2-011	ITS2-027
ORF18	Dloop01	ITS2-017	ITS2-017
ORF18	Dloop01	ITS2-022	ITS2-022
ORF18	Dloop01	ITS2-022	ITS2-038
ORF18	Dloop01	ITS2-022	ITS2-047
ORF18	Dloop01	ITS2-022	ITS2-073
ORF18	Dloop01	ITS2-027	ITS2-001
ORF18	Dloop01	ITS2-027	ITS2-027
ORF18	Dloop01	ITS2-029	ITS2-029
ORF18	Dloop01	ITS2-038	ITS2-001
ORF18	Dloop01	ITS2-038	ITS2-047
ORF18	Dloop01	ITS2-038	ITS2-050
ORF18	Dloop01	ITS2-074	ITS2-075
ORF18	Dloop01	ITS2-096	ITS2-029
ORF18	Dloop17	ITS2-001	ITS2-001
ORF19		ITS2-001	ITS2-001
ORF19	Dloop03	ITS2-001	ITS2-001
ORF23	Dloop04		

ORF haplotype	Dloop haplotype	ITS2_hap1	ITS2_hap2
ORF23	Dloop06		
ORF23		ITS2-026	ITS2-026
ORF23		ITS2-068	ITS2-068
ORF23		ITS2-084	ITS2-092
ORF23	Dloop07	ITS2-059	ITS2-059
ORF24		ITS2-040	ITS2-040
ORF25	Dloop11	ITS2-004	ITS2-004
ORF26	Dloop04		
ORF27	Dloop04		
ORF27	Dloop05		
ORF27	Dloop11		
ORF27		ITS2-005	ITS2-005
ORF27		ITS2-006	ITS2-006
ORF27		ITS2-006	ITS2-007
ORF27		ITS2-006	ITS2-021
ORF27		ITS2-007	ITS2-007
ORF27		ITS2-007	ITS2-104
ORF27		ITS2-008	ITS2-008
ORF27		ITS2-008	ITS2-106
ORF27		ITS2-015	ITS2-015
ORF27		ITS2-016	ITS2-016
ORF27		ITS2-016	ITS2-048
ORF27		ITS2-021	ITS2-021
ORF27		ITS2-023	ITS2-023
ORF27		ITS2-024	ITS2-008
ORF27		ITS2-025	ITS2-025
ORF27		ITS2-026	ITS2-026
ORF27		ITS2-041	ITS2-102
ORF27		ITS2-042	ITS2-042
ORF27		ITS2-043	ITS2-043
ORF27		ITS2-045	ITS2-045
ORF27		ITS2-063	ITS2-007
ORF27		ITS2-065	ITS2-065
ORF27		ITS2-066	ITS2-025
ORF27		ITS2-071	ITS2-126
ORF27		ITS2-103	ITS2-023
ORF27		ITS2-105	ITS2-007
ORF27		ITS2-107	ITS2-108
ORF27		ITS2-111	ITS2-112
ORF27		ITS2-113	ITS2-114
ORF27		ITS2-118	ITS2-120
ORF27		ITS2-125	ITS2-071
ORF27		ITS2-127	ITS2-128
ORF27	Dloop04	ITS2-006	ITS2-006
ORF27	Dloop04	ITS2-006	ITS2-007
ORF27	Dloop04	ITS2-006	ITS2-021
ORF27	Dloop04	ITS2-006	ITS2-063
ORF27	Dloop04	ITS2-007	ITS2-007
ORF27	Dloop04	ITS2-008	ITS2-008

ORF haplotype	Dloop haplotype	ITS2_hap1	ITS2_hap2
ORF27	Dloop04	ITS2-015	ITS2-015
ORF27	Dloop04	ITS2-016	ITS2-016
ORF27	Dloop04	ITS2-016	ITS2-100
ORF27	Dloop04	ITS2-016	ITS2-101
ORF27	Dloop04	ITS2-021	ITS2-021
ORF27	Dloop04	ITS2-023	ITS2-023
ORF27	Dloop04	ITS2-024	ITS2-008
ORF27	Dloop04	ITS2-024	ITS2-024
ORF27	Dloop04	ITS2-025	ITS2-025
ORF27	Dloop04	ITS2-026	ITS2-026
ORF27	Dloop04	ITS2-041	ITS2-041
ORF27	Dloop04	ITS2-042	ITS2-042
ORF27	Dloop04	ITS2-045	ITS2-045
ORF27	Dloop04	ITS2-048	ITS2-048
ORF27	Dloop04	ITS2-049	ITS2-006
ORF27	Dloop04	ITS2-049	ITS2-041
ORF27	Dloop04	ITS2-064	ITS2-064
ORF27	Dloop04	ITS2-067	ITS2-067
ORF27	Dloop04	ITS2-117	ITS2-119
ORF27	Dloop04	ITS2-123	ITS2-124
ORF27	Dloop05	ITS2-040	ITS2-040
ORF27	Dloop08	ITS2-010	ITS2-095
ORF30		ITS2-028	ITS2-085
ORF31		ITS2-028	ITS2-028
ORF34	Dloop11	ITS2-032	ITS2-032
ORF36		ITS2-002	ITS2-002
ORF36		ITS2-004	ITS2-004
ORF36		ITS2-009	ITS2-009
ORF36		ITS2-030	ITS2-030
ORF36		ITS2-032	ITS2-032
ORF36	Dloop11	ITS2-013	ITS2-013
ORF38	Dloop11	ITS2-002	ITS2-002
ORF38	Dloop11	ITS2-004	ITS2-002
ORF38	Dloop11	ITS2-013	ITS2-009
ORF38	Dloop11	ITS2-044	ITS2-044
ORF39	Dloop11		
ORF39		ITS2-002	ITS2-002
ORF39		ITS2-002	ITS2-009
ORF39		ITS2-002	ITS2-013
ORF39		ITS2-002	ITS2-018
ORF39		ITS2-002	ITS2-076
ORF39		ITS2-002	ITS2-090
ORF39		ITS2-003	ITS2-099
ORF39		ITS2-004	ITS2-002
ORF39		ITS2-004	ITS2-004
ORF39		ITS2-004	ITS2-013
ORF39		ITS2-004	ITS2-035
ORF39		ITS2-004	ITS2-077
ORF39		ITS2-005	ITS2-009
ORF39		ITS2-009	ITS2-009
ORF39		ITS2-013	ITS2-009

ORF haplotype	Dloop haplotype	ITS2_hap1	ITS2_hap2
ORF39		ITS2-013	ITS2-013
ORF39		ITS2-014	ITS2-009
ORF39		ITS2-014	ITS2-013
ORF39		ITS2-014	ITS2-014
ORF39		ITS2-018	ITS2-009
ORF39		ITS2-018	ITS2-018
ORF39		ITS2-019	ITS2-009
ORF39		ITS2-019	ITS2-014
ORF39		ITS2-019	ITS2-019
ORF39		ITS2-030	ITS2-030
ORF39		ITS2-033	ITS2-009
ORF39		ITS2-033	ITS2-033
ORF39		ITS2-035	ITS2-035
ORF39		ITS2-036	ITS2-036
ORF39		ITS2-044	ITS2-044
ORF39		ITS2-057	ITS2-057
ORF39		ITS2-078	ITS2-002
ORF39		ITS2-080	ITS2-081
ORF39		ITS2-088	ITS2-082
ORF39		ITS2-109	ITS2-036
ORF39		ITS2-110	ITS2-030
ORF39	Dloop11	ITS2-002	ITS2-002
ORF39	Dloop11	ITS2-002	ITS2-009
ORF39	Dloop11	ITS2-002	ITS2-097
ORF39	Dloop11	ITS2-004	ITS2-002
ORF39	Dloop11	ITS2-004	ITS2-004
ORF39	Dloop11	ITS2-009	ITS2-009
ORF39	Dloop11	ITS2-014	ITS2-014
ORF39	Dloop11	ITS2-018	ITS2-009
ORF39	Dloop11	ITS2-018	ITS2-018
ORF39	Dloop11	ITS2-019	ITS2-019
ORF39	Dloop11	ITS2-030	ITS2-030
ORF39	Dloop11	ITS2-032	ITS2-032
ORF39	Dloop11	ITS2-033	ITS2-009
ORF39	Dloop11	ITS2-035	ITS2-009
ORF39	Dloop11	ITS2-035	ITS2-018
ORF39	Dloop11	ITS2-051	ITS2-051
ORF39	Dloop11	ITS2-079	ITS2-002
ORF39	Dloop11	ITS2-087	ITS2-019
ORF40	Dloop12		
ORF42		ITS2-056	ITS2-056
ORF46	Dloop12		
ORF46		ITS2-010	ITS2-010
ORF46		ITS2-010	ITS2-020
ORF46		ITS2-012	ITS2-010
ORF46		ITS2-012	ITS2-012
ORF46		ITS2-012	ITS2-020
ORF46		ITS2-012	ITS2-020
ORF46		ITS2-020	ITS2-020
ORF46		ITS2-060	ITS2-060
ORF46		ITS2-070	ITS2-070
ORF46		ITS2-093	ITS2-010

ORF haplotype	Dloop haplotype	ITS2_hap1	ITS2_hap2
ORF46	Dloop10	ITS2-039	ITS2-039
ORF46	Dloop12	ITS2-010	ITS2-010
ORF46	Dloop12	ITS2-010	ITS2-020
ORF46	Dloop12	ITS2-012	ITS2-012
ORF46	Dloop12	ITS2-012	ITS2-020
ORF46	Dloop12	ITS2-020	ITS2-020
ORF46	Dloop12	ITS2-069	ITS2-069
ORF46	Dloop12	ITS2-094	ITS2-020
ORF47	Dloop11		
ORF47		ITS2-003	ITS2-002
ORF47		ITS2-003	ITS2-003
ORF47		ITS2-003	ITS2-005
ORF47		ITS2-005	ITS2-005
ORF47		ITS2-014	ITS2-014
ORF47		ITS2-034	ITS2-034
ORF47		ITS2-039	ITS2-039
ORF47		ITS2-053	ITS2-005
ORF47		ITS2-083	ITS2-005
ORF47	Dloop11	ITS2-003	ITS2-034
ORF47	Dloop11	ITS2-005	ITS2-005
ORF47	Dloop11	ITS2-089	ITS2-005
ORF47	Dloop13	ITS2-005	ITS2-005
ORF50	Dloop13		
ORF50	Dloop14		
ORF50	Dloop13	ITS2-052	ITS2-052
ORF50	Dloop13	ITS2-058	ITS2-058
ORF51		ITS2-003	ITS2-003
ORF52	Dloop11		
ORF53		ITS2-003	ITS2-003
ORF53		ITS2-003	ITS2-005
ORF53		ITS2-005	ITS2-005
ORF53		ITS2-032	ITS2-032
ORF53	Dloop11	ITS2-003	ITS2-003
ORF53	Dloop11	ITS2-005	ITS2-005
ORF53	Dloop14	ITS2-003	ITS2-003
ORF54	Dloop11		
ORF54		ITS2-003	ITS2-003
ORF54		ITS2-005	ITS2-005
		ITS2-008	ITS2-008
		ITS2-025	ITS2-025
	Dloop04	ITS2-006	ITS2-006
	Dloop04	ITS2-008	ITS2-008
	Dloop04	ITS2-015	ITS2-015
	Dloop04	ITS2-025	ITS2-025
	Dloop04	ITS2-049	ITS2-021
	Dloop04	ITS2-122	ITS2-121
	Dloop11	ITS2-002	ITS2-002
	Dloop11	ITS2-002	ITS2-018
	Dloop11	ITS2-004	ITS2-004
	Dloop11	ITS2-004	ITS2-086
	Dloop11	ITS2-009	ITS2-009
	Dloop11	ITS2-014	ITS2-013
	Dloop11	ITS2-019	ITS2-019

Appendix S13

	SSH01	SSH04	SSH05a	SSH05b	SSH05c	SSH05d	SSH09a	SSH09b	SSH09c	SSH13a	SSH13b	SSH13c	SSH16
SSH01	-	0.617***	0.473***	0.530***	0.503***	0.540***	0.549***	0.556***	0.508***	0.480***	0.654***	0.543***	0.546***
SSH04	0.378	-	0.488***	0.526***	0.534***	0.557***	0.636***	0.484***	0.685***	0.507***	0.650***	0.353***	0.336***
SSH05a	0.148*	0.285**	-	0.328***	0.350***	0.277***	0.512***	0.433***	0.545***	0.385***	0.431***	0.429***	0.375***
SSH05b	0.352 ^{NA}	0.418 ^{NA}	0.214 ^{NA}	-	0.522***	0.474***	0.625***	0.656***	0.669***	0.613***	0.666***	0.611***	0.614***
SSH05c	0.244***	0.333***	0.133***	0.318 ^{NA}	-	0.330***	0.583***	0.466***	0.556***	0.562***	0.535***	0.451***	0.427***
SSH05d	0.218**	0.327***	0.088***	0.285 ^{NA}	0.173***	-	0.614***	0.541***	0.561***	0.482***	0.562***	0.556***	0.501***
SSH09a	0.249*	0.377**	0.214***	0.362 ^{NA}	0.251***	0.269***	-	0.374***	0.416***	0.601***	0.329***	0.471***	0.483***
SSH09b	0.226	0.304	0.148*	0.377 ^{NA}	0.203***	0.230***	0.154*	-	0.378***	0.496***	0.454***	0.308***	0.276***
SSH09c	0.268*	0.414*	0.233***	0.392 ^{NA}	0.294***	0.276***	0.225**	0.225*	-	0.605***	0.557***	0.587***	0.606***
SSH13a	0.215**	0.301***	0.165***	0.283 ^{NA}	0.260***	0.219***	0.276***	0.191**	0.288***	-	0.534***	0.390***	0.374***
SSH13b	0.293	0.422*	0.205***	0.391 ^{NA}	0.261***	0.278***	0.123**	0.209	0.328**	0.270**	-	0.426***	0.439***
SSH13c	0.224	0.223	0.167*	0.339 ^{NA}	0.203**	0.229**	0.206*	0.091	0.286	0.171**	0.195	-	0.038
SSH16	0.206	0.213*	0.128**	0.360 ^{NA}	0.178***	0.203***	0.196**	0.064	0.290*	0.160**	0.184*	0.004	-

*: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$

CHAPITRE 2. Structure des populations à fine échelle, cas du corail *Pocillopora damicornis* à La Réunion

Résumé

Le corail *Pocillopora damicornis* type β est connu pour présenter plusieurs modes de reproduction: la reproduction sexuée permet de créer des nouvelles combinaisons d'allèles et donc de nouveaux génotypes, tandis que la reproduction asexuée permet de propager rapidement les génotypes. Dans le but d'estimer la part relative de chaque mode de reproduction dans les populations de *P. damicornis* type β de l'île de La Réunion (océan Indien), la propagation clonale a été étudiée sur la côte Ouest de l'île entre quatre sites présentant un gradient de distances (de 2 à 40 km). Les colonies ont été échantillonnées de manière exhaustive, aléatoire ou au hasard en fonction des caractéristiques de chacun des sites. La diversité génotypique a été estimée à partir de 13 microsatellites sur 510 colonies de *P. damicornis* type β dont le type a été vérifié *a posteriori* grâce au séquençage du marqueur mitochondrial ORF (840 pb). Sur l'ensemble des 510 colonies de *P. damicornis* type β étudiées, 47 % d'entre-elles présentent le même génotype multi-locus (MLG), correspondant à un « superclone » et suggérant que la reproduction asexuée est extrêmement présente à La Réunion. Au sein de chacun des sites étudiés, plusieurs MLGs sont partagés par plusieurs colonies, suggérant une propagation clonale localisée, probablement résultant de phénomènes de fragmentation. Par ailleurs, certains MLGs sont identifiés sur plusieurs sites séparés jusqu'à 40 km. Alors qu'il paraît peu probable que la propagation de certains clones sur plusieurs dizaines de kilomètres soit le résultat de la fragmentation des colonies, les résultats suggèrent l'existence d'une production de larves parthénogénétiques. Malgré le fait que certains génotypes soient retrouvés sur plusieurs sites, les tests d'assignement révèlent deux clusters génétiques très différenciés entre eux le long de la côte Ouest de La Réunion, soulignant l'intérêt de mettre en place localement des mesures de protection adaptées.

Les données acquises au cours de ce chapitre ont permis de rédiger un article qui est en révision pour publication dans *PLOS ONE*.

Superclone expansion, long-distance clonal dispersal and local genetic structuring in the coral *Pocillopora damicornis* type β in Reunion Island, South Western Indian Ocean

Short title: Superclone expansion and genetic structure in *Pocillopora damicornis* type β

Gélin P¹, Fauvelot C², Mehn V¹, Bureau S¹, Rouzé H², Magalon H¹

¹ UMR ENTROPIE (IRD, Université de La Réunion, CNRS), Laboratoire d'excellence-CORAIL; Université de La Réunion, St Denis, La Réunion

² UMR ENTROPIE (IRD, Université de La Réunion, CNRS), Laboratoire d'excellence-CORAIL; centre IRD de Nouméa, New Caledonia

Corresponding author: pauline.gelin@univ-reunion.fr; Université de La Réunion, UMR ENTROPIE Faculté des Sciences et Technologies, 15 Bd René Cassin, CS 92003, 97744 St Denis Cedex 09, La Réunion

Abstract

The scleractinian coral *Pocillopora damicornis* type β (*sensu* [1]) is known to present a mixed reproduction mode: through sexual reproduction, new genotypes are created, while asexual reproduction insures their propagation. In order to investigate the relative proportion of each reproduction mode in *P. damicornis* type β populations from Reunion Island, Indian Ocean, clonal propagation along the west coast was assessed through four sampling sites with increasing geographical distance between sites. Coral colonies were sampled either exhaustively, randomly or haphazardly within each site, and genotypic diversity was assessed using 13 microsatellite loci over a total of 510 *P. damicornis* type β determined *a posteriori* from their mtDNA haplotype (a 840 bp sequenced fragment of the Open Reading Frame). Overall, 47% of all the sampled colonies presented the same multi-locus genotype (MLG), a superclone, suggesting that asexual propagation is extremely important in Reunion Island. Within each site, numerous MLGs were shared by several colonies, suggesting local clonal propagation through fragmentation. Moreover, some of these MLGs were found to be shared among several sites located 40 km apart. While asexual reproduction by fragmentation seems unlikely over long distances, our results suggest a production of parthenogenetic larvae. Despite shared MLGs, two differentiated clusters were enclosed among populations of the west coast of Reunion Island, revealing the necessity to set up appropriate managing strategies at a local scale.

Introduction

Asexual reproduction allows high population growth rates, successful genotypes propagation and energy savings due to the lack of mate search (*e.g.*, [2]). However, contrary to sexual reproduction [3] which is more costly [4], asexual reproduction does not offer the opportunity to (1) produce recombinant types that might be advantageous [5, 6], (2) flush out deleterious mutations [7, 8], and (3) keep up with environmental modifications by increasing genetic diversity in populations [9]. Nevertheless, even with efficient DNA repair mechanisms (*e.g.*, in bdelloid rotifers; [10]), most of asexual lineages need some occasional sexual recombination to counter the disadvantage of not reproducing sexually in order to persist over long time scales. Considering this, the theoretical best model would correspond to a mixed strategy: several taxa from several kingdoms show an alternation from sexual to asexual mode [11].

Most reef-building corals present various mixed strategies for reproduction (reviewed in [12]). Generally, for marine benthic organisms and particularly for scleractinian corals, sexual reproduction is achieved by either spawning gametes with external fertilization (spawners), or by internal fertilization and brooding larvae inside the coral polyp (brooders) [12]. In both cases, the formed planulae disperse, settle and metamorphose to form new colonies. Additionally, some corals have been described to reproduce asexually by vegetative propagation in different ways: fragmentation [13], budding or polyp expulsion [14], or release of asexual larvae, which has been evidenced in few scleractinians such as *Pocillopora* species (*e.g.*, [15, 16]) and *Tubastrea* species [17]. Indeed, parthenogenetic larvae production has been found to be an alternate reproductive mode for corals when sexual reproduction leads to high failure in fertilization rates in cases of sperm limitation [18].

The common branching coral *P. damicornis* (Linnaeus, 1758) is found in lagoons from Red Sea, Indian and Pacific Oceans in shallow habitats [19]. It has been described to present (1) a huge morphological plasticity linked to environmental conditions and depth [20] and (2) sexual and asexual reproduction modes with abilities to be either brooder or spawner [21]. Parthenogenetic larvae production in *P. damicornis* has been reported for thirty years using different genetic tools: isozymes [15], allozymes (*e.g.*, [16]) or microsatellites (*e.g.*, [22]). However, these findings are to be taken with a grain of salt as the taxonomy of *Pocillopora* species has been recently revised using the mitochondrial Open Reading Frame (ORF) sequences (840 bp), showing five distinct lineages (α , β , γ , δ and ϵ) under the *P. damicornis*

designation [1]. Concerning our study, we focused on *P. damicornis* type β , also designated as *P. acuta* (*sensu* [23]), *Pocillopora* type 5 [24], *P. damicornis* type F [25] and *Pocillopora* Clade Ia [26]. Without judging these different designations, we chose to use the name *P. damicornis* type β (referred to *P. damicornis* β hereafter for lightening the writing) to refer to the lineage identity herein under study, as it seems to us the more explicit considering the knowledge on species delimitation in this genus to date. Thus, former studies dealing with *P. damicornis sensu lato* should be reconsidered in the light of these new taxonomic insights, as molecular identification of colonies is revealed as an absolute necessity. As an example, Pinzón et al [27] revealed high clonal propagation over small distances (*i.e.*, ≤ 10 km) in colonies from ORF type 1 that the authors designated as *P. damicornis*. Nevertheless, in Pinzón et al [24], the authors showed that type 1 colonies form a clade corresponding to various morphs (*damicornis*, *verrucosa*, *elegans*, *meandrina*, *capitata*, *eydouxi*), which Schmidt-Roach et al [23] attributed to *P. eydouxi/meandrina* complex. Similarly, Gorospe and Karl [28] studied clonal propagation in *P. damicornis* in Hawaii, showing an over-representation of one clone at a scale of 1.5 km: from their sampling, they found two lineages (one represented by 98.8% of their samples), identified *a posteriori* as *P. damicornis* lineages α , β or type 6 (*sensu* [24]) [29], but the authors did not indicate which lineage was the most represented in their first study. Likewise, in Combosch and Vollmer [22], it seems that two lineages were present as the planulation date was different from one colony to another. Indeed, some spawned at the new moon and others at the full moon, indicating that they may be from lineages β and α , respectively (see discussion in [1]). Concerning the earlier studies focusing on *P. damicornis* clonal propagation [15, 16, 30-33], they did not provide the ORF sequences nor indications about planulation date, so it is not possible to determine which *P. damicornis* lineage they dealt with and to extrapolate about clonal propagation of *P. damicornis* β . Contrary to all the studies cited above, few studies presented no doubt about *P. damicornis* identification. Indeed, the studied colonies from Souter et al [25] were assessed *a posteriori* to *P. damicornis* β . This latter found clonal propagation over 60 km on the East Coast of Africa. Moreover, in Polynesia, high clonal propagation was found in *P. damicornis* β and clones were found to be extended over very large distances (*ca.* 200 km; [34]). Conversely, in the Great Barrier Reef, Torda et al [35] found very short distance of clonal propagation (not more than 230 m) in *P. damicornis* β .

Given the new taxonomic insights in the *Pocillopora* genus and the few studies for which *P. damicornis* lineages were clearly identified, the aim of our study was to investigate, in Reunion Island (Western Indian Ocean), the clonal propagation and population structure in the coral *P. damicornis* β . For this purpose, we used 13 microsatellite loci to identify genetically

sampled colonies and we chose four study sites presenting contrasted environmental conditions (hydrodynamics, temperature, tourism activities,...) along the west coast of Reunion Island to assess the relative proportions of each reproductive modes in this species and the distance of clonal propagation at the island scale.

Material and methods

Sampling design

Colonies of *P. damicornis* β were sampled in lagoons, at four sites located along the west coast of Reunion Island (Fig 1), situated in the South Western Indian Ocean, 700 km east of Madagascar. These sites, chosen for their contrasted environmental conditions and differences in coral density, were accessible by snorkeling in one meter depth and were named, from North to South, REU1, -2, -3, -4. REU1 and REU2 are located on the north part of the west coast of Reunion Island, where the lagoon is at its widest (approximately 500 m) and where touristic activities induce disturbances on the reef. REU4 and in lesser proportions REU3, the southernmost sites, are exposed to strong swells and waves during the southern winter and are less frequented, contrasting with REU1 and REU2 sites which are located in less exposed reefs with weaker hydrodynamics [36].

REU1 (S21°04'56.49", E55°13'22.19") and REU2 (S21°05'47.94", E55°14'01.68"), separated by a shallow channel, are located approximately 2 km apart in the lagoon complex of La Saline-L'Ermitage. REU3 (S21°16'15.17", E55°19'56.31") and REU4 (S21°20'44.68", E55°28'21.66") are located in separated lagoons, REU3 approximately 25 km south of REU1-REU2 and REU4 approximately 40 km south of REU1-REU2 (Fig 1).

Exhaustive sampling

In REU1, colonies were distributed along with *Acropora* patches separated by sand strips. As the colonies distribution was not uniform, an exhaustive sampling was set up. We sampled all colonies ($N_{total} = 183$) found in a 12 m radius circle centered on a patch and recorded the X-Y coordinates using two tape measures placed perpendicularly to materialize the axes. A third tape measure combined to a 1 m x 1 m quadrat was used to move along both axes.

Random sampling

In REU2, as colonies were distributed uniformly on the substrate and easily accessible without damaging the reef, colonies were sampled following a random sampling method using four nested circles of 2, 4, 8 and 12 m radius, respectively ($N_{total} = 297$; Fig 2). Within each strip, 50 colonies were randomly collected, except in the central circle (2 m radius) where an exhaustive sampling was conducted to study micro-scale clone dispersion (from 0 to 12 m). Practically, graphic coordinates had been previously randomly generated in the different circles. Then, in the field, a colony situated approximately in the middle of the *P. damicornis* β patch was selected as origin of the plan. The same plan as for REU1 was set up for recording the coordinates.

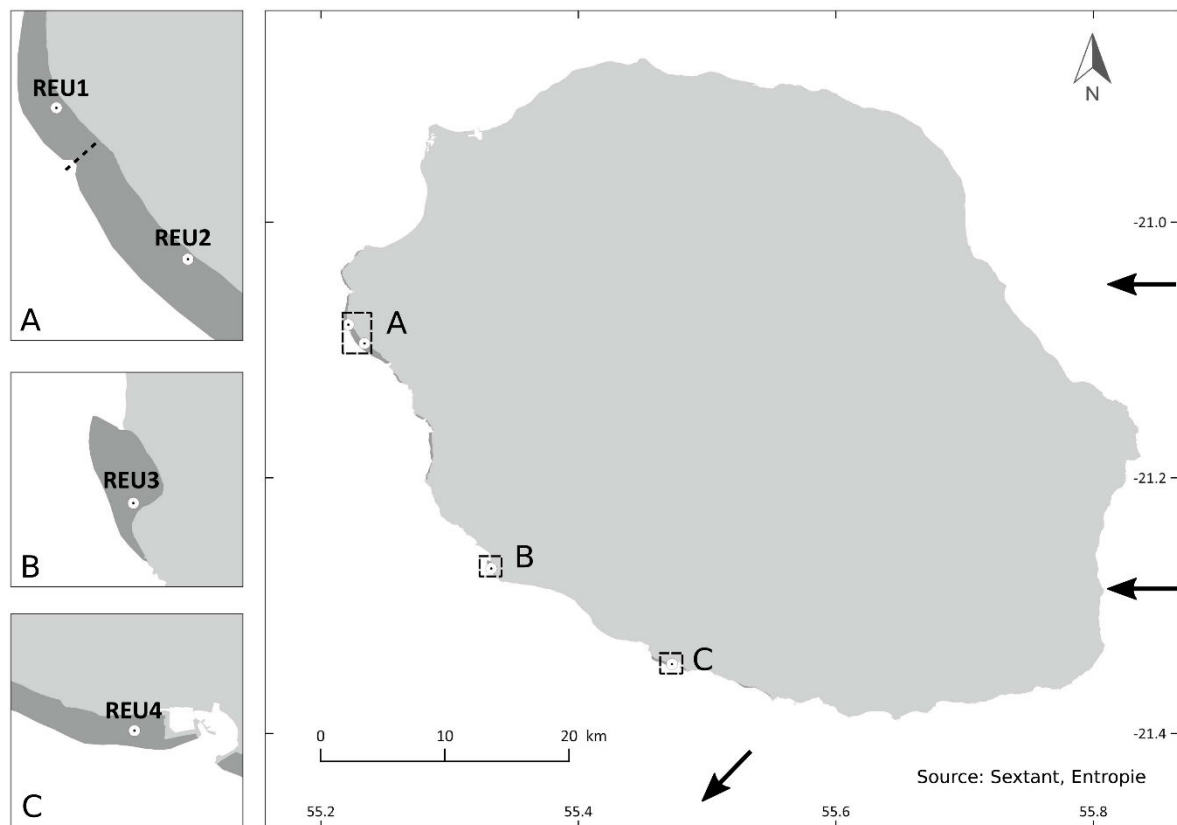


Fig 1: Map of the sampling sites of *Pocillopora damicornis* type β in Reunion Island. Lagoons are represented in dark grey. The channel between REU1 and REU2 is symbolized with a dash line and the direction of the South Equatorial Current around Reunion Island is symbolized with black arrows (from [37]). (Open access data: http://sextant.ifremer.fr/fr/web/ocean_indien/geoportail/sextant#/metadata/f3587329-b5b7-4ee1-822d-5cda9cef4d09).

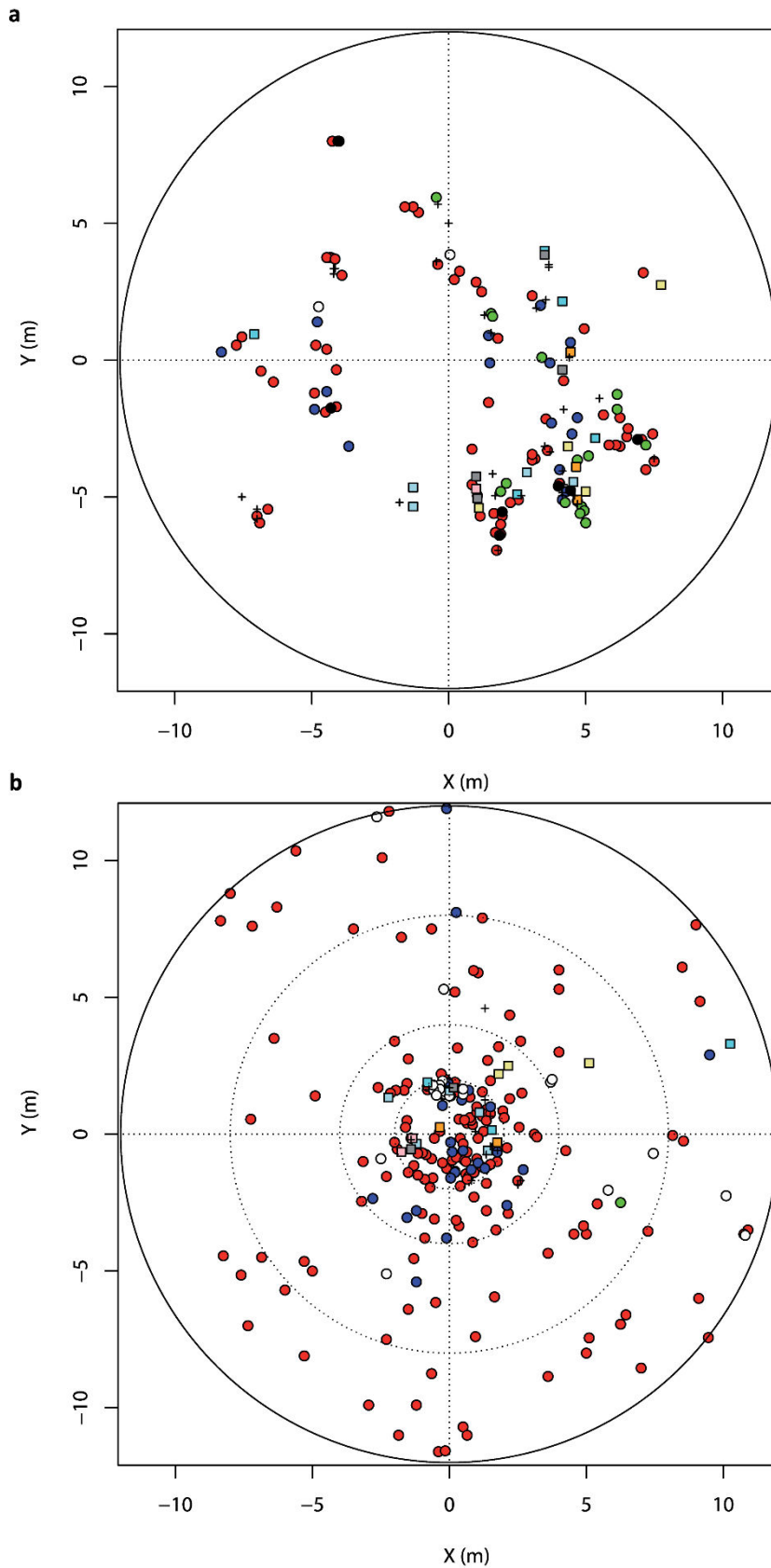


Fig 2: Spatial representation of *Pocillopora damicornis* type β colonies sampled in (a) REU1 ($N = 158$) and (b) REU2 ($N = 264$). Each symbol represents a colony. Multi-locus genotypes (MLG) occurring only once are represented with “+”. Colonies sharing the same MLG are represented with the same color and each repeated MLG is represented by a different color. MLGs shared by both sites are represented with circles (MLG01: red, MLG02: blue, MLG03: white, MLG04: green, MLG05: black) and MLGs specific to a unique site with squares.

Haphazard sampling

For REU3 and REU4, a random sampling scheme was unfortunately not applicable. Indeed, in REU3, the high coral cover combined with low depth prevented us from setting up the same kind of sampling scheme without damaging the reef. Also, in REU4, *P. damicornis* β colonies were too scattered to sample enough colonies in a 12 m radius circle. Thus, colonies from these two sites were sampled haphazardly while snorkeling, without reporting coordinates ($N_{total} = 50$ and 57 , respectively). All colonies were photographed and their color reported (pink or cream). A small fragment (1–3 cm) was removed from each colony, placed into a numbered zip-lock bag, fixed in 90% ethanol and stored at room temperature. Hereafter, population refers to all the colonies sampled at one sampling site.

DNA extraction, sequencing, and microsatellite genotyping

From the 587 sampled colonies, DNA was extracted using DNeasy Blood & Tissue kit (Qiagen™). Colonies were genotyped using 13 microsatellite loci (Appendix S1). Forward primers were indirectly fluorochrome labelled (6-FAM, VIC, NED, PET) by adding a universal M13 tail at the 5'-end and were multiplexed post-PCR in three panels (Appendix S1). Each amplification reaction was performed as in Postaire et al [38]. PCR products were genotyped using an ABI 3730 genetic analyser (Applied Biosystems) and allelic sizes were determined with GeneMapper V4.0 (Applied Biosystems) using an internal size standard (Genescan LIZ-500, Applied Biosystems). Individuals showing peak profiles either faint or with more than two peaks were processed again and, when remaining ambiguous, designated as missing data to avoid dealing with more than two alleles (resulting either from non-specific amplifications or somatic mutation). However, it should be noted that keeping the third allele, particularly for Pd3-EF65, did not interfere with multi-locus genotypes (MLG) identification (PG pers. obs.; [28]). Finally, 519 colonies without missing data were kept for further analyses. Because colonies were sampled based on their *corallum* macromorphology, *P. damicornis* β lineage identity was verified by amplifying a 840 bp fragment of the mitochondrial open reading frame (ORF) region with the FATP6.1 and the RORF primers and using the same conditions as described in Flot and Tillier [39], for 108 colonies selected *a posteriori* based on their MLG at microsatellite loci (see below). Amplifications were sent for sequencing to Genoscreen (Lille, France) and edited with Geneious R7 [40].

Analysis of clonal structure

The occurrence of identical MLGs among the sampled colonies was assessed using GenClone V2.0 [45]. The probability of obtaining the same MLG more than once from distinct random reproductive events was further estimated using $P_{SEX} (F_{IS})$, which considers possible departure from Hardy-Weinberg equilibrium in order to obtain a less biased estimator [46]. Colonies sharing the same MLG resulting from a unique reproductive event (one unique zygote) were considered as belonging to the same clone. Since somatic mutations [47] or scoring errors may overestimate the true number of clones, pairwise genetic distances were calculated among all colonies using GenoType [48] based on mutational steps under the stepwise mutational model. A threshold in the genetic distance distribution was determined under which distinct MLGs were considered as the same multi-locus lineage or MLL (*i.e.*, clones genetically close, *sensu* [45]).

For each population, the clonal richness R was estimated [49], as well as the genotypic diversity G , estimated as G_O/G_E , where G_O is the observed genotypic diversity as described by [50] and G_E is the expected genotypic diversity as described in Baums et al [51]. G reflects the reproduction mode, varying from 1 in sexual populations to 0 in clonal ones. The average number of alleles per locus (allelic richness; \hat{A}) was estimated with the truncated dataset (*i.e.*, keeping one representative per MLG per population) with FSTAT V2.9.3 [52]. Moreover, to assess the relevance of the loci used (number and variability), we estimated the probability that two randomly sampled MLGs share the exact same alleles over all loci just by chance rather than being the result of asexual reproduction (probability of identity P_{ID} , [53]) with GIMLET V1.3.3 [54].

With the aim to assess clonal heterogeneity and evenness of all our populations, the parameter β of the Pareto distribution and the Simpson's evenness (ED^*) were estimated with GenClone V2.0 [45] within each population. Moreover, in order to investigate the spatial distribution of genotypes in each population for which spatial coordinates were available (REU1 and REU2), the aggregation index (Ac) and edge effect (Ee) were estimated and their significance tested with 10^3 permutations in GenClone V2.0 [45]. Moran's spatial autocorrelation index (I) was further estimated with SPAGeDi V1.5 [55] to assess whether genetic relatedness is related to geographic distance between colonies.

Populations' genetic differentiation

The genotypic disequilibrium among pairs of loci was assessed with FSTAT V2.9.3 [52], on the truncated dataset. The presence of null allele was tested using Micro-Checker V2.2.3 [56]. Departures from Hardy-Weinberg equilibrium (HWE) for each population (F_{IS}) were estimated using GENETIX software V4.05.2 [57] with 10^4 permutations.

Population differentiation indices F_{ST} [58] and D_{est} [59] were estimated both on the entire dataset (*i.e.*, keeping all individuals) and on the truncated one. Indeed, as some genotypes might present a higher fitness than others, keeping all the colonies sharing these genotypes might bias population differentiation estimation, as well as keeping only one representative per genotype. Thus, fitness of each genotype is taken into account when all repetitions of each MLG are kept while the truncated dataset considers all fitnesses equally (as in [60]). For both cases, hypotheses are probably wrong and the real differentiation between populations is likely intermediate between both scenarios.

Additionally, to determine the most likely number of genetically homogenous clusters (K) within our dataset, a Bayesian clustering analysis was performed using InStruct [61], which considers partial self-fertilization or inbreeding. Conditions were set to 10^6 chain length after a burn-in of 10^5 and 10 chains were run for each K varying from 1 to 10. The analysis was performed both on the entire and the truncated datasets for comparisons. The optimal K was chosen based on lower values of the Deviance Information Criterion (DIC, [62]) and the Evanno's method [63]. For both cases, graphs were drawn using Distruct V1.1 [64] after combining the different replicates for each K in Clumpp V1.1.2 [65]. Additionally, in order to explore the diversity structure without any prior hypothesis regarding our populations (HWE, LD), we performed a Principal Components Analysis (PCA) based on a dissimilarity matrix using Dice distance between MLGs with Darwin V6.0.9 [66] and constructed a network based on shared allele distance using EDENetwork V2.18 [67].

Results

All 13 loci used were polymorphic in each population except locus PV7 in REU2. The allelic richness ranged from 2.46 for Pd04 to 6.07 for Poc40 (Appendix S1). The combination and high variability of the 13 microsatellite loci used here resulted in a very low probability ($P_{ID} = 1.337 \times 10^{-5}$) that two colonies sharing the same MLG were identical by chance, so that shared MLGs were considered to be the result of asexual reproduction. The use of these 13 loci therefore allows uncovering the entire genotypic diversity of the dataset.

Over the 519 genotypes without missing data, 108 distinct MLGs were detected. For one representative of each MLG, the mitochondrial ORF was sequenced (840 bp; accession numbers from KT879932 to KT880039) and compared to GenBank sequences identified as *P. damicornis* type β , *P. damicornis* type α , *P. damicornis* type σ , *P. damicornis* type ε , *P. verrucosa*, *P. meandrina* and *P. eydouxi* (JX625045 to JX625077, JX983183 to JX983186 and JX985610 to JX985613; [1, 23]). Over our 108 sequenced colonies, nine were identified as belonging to the *P. meandrina/eydouxi* complex (one from REU1 and five from REU4) and *P. verrucosa* (one from REU1 and two from REU4) and were removed from the dataset. Finally, for further analyses, were kept 510 *P. damicornis* β colonies (all sharing the exact same mtDNA haplotype), corresponding to 99 different MLGs.

The hypothesis of multiple reproduction events being responsible for the presence of several colonies sharing the same MLG was rejected in all cases (all $P_{SEX}(F_{IS}) < 1.24 \times 10^{-5}$), indicating that colonies with identical MLG belonged to the same clone. Moreover, the first gap in the distribution of pairwise allelic differences among the 99 MLGs was found between 0 and two mutation steps. Thus, to compensate for genotyping errors and possible somatic mutations, we chose a threshold of one mutation step to differentiate MLLs, as found in other studies (e.g., [68]). Since pairwise MLGs differed by at least two mutation steps, all distinct MLGs were considered as distinct MLLs. To avoid redundancy, we further kept the term MLG rather than MLL for the 99 different MLGs found among our sampling.

Over the 99 MLGs, 27 occurred more than once, being shared by several colonies within sites, as well as between sites for seven of these MLGs (Fig 3). Interestingly, one MLG (MLG01) was over-represented (47% of the overall colonies sampled), especially in REU1 and REU2 presenting 82% and 81% of the sampled colonies, respectively (Fig 2). Furthermore, three MLGs (including the two most highly common ones, MLG01 and MLG02; Fig 3) were observed both in REU1 and REU2, separated by 2 km. More surprisingly, MLGs were found in much more distant sites, between REU1 and REU3 (MLG01 and MLG02; Fig 3) approximately 25 km apart and between REU1 and REU4 (MLG03), located approximately 40 km apart. Furthermore, it is noteworthy that colonies sharing a same MLG did not always display the same colour (pink or cream).

Within populations, the clonal richness varied from 0.095 in REU2 to 0.596 in REU4 (Table 1). Additionally, the genotypic diversity G was low for all populations and ranged from 0.006 in REU2 to 0.264 in REU4. Concerning Simpson's evenness index and the parameter β , the lowest values were found for REU2 (very low slope; Appendix S2), revealing the presence of few MLGs but highly and non-evenly repeated (Table 1, Appendix S2). Conversely, in REU4, there

were numerous MLGs repeated a few times and more evenly (steepest slope for Pareto distribution; Appendix S2). REU1 and REU3 presented an intermediate slope between the two others.

Table 1: Summary statistics for the four *Pocillopora damicornis* type β populations

Population	REU1	REU2	REU3	REU4
N_{total}	183	297	50	57
N	158	264	42	46
N_{MLG}	44	25	13	27
R	0.274	0.095	0.293	0.596
G	0.016	0.006	0.129	0.264
ED^*	0.681	0.514	0.755	0.706
B	0.212	0.085	0.375	0.625
Ac	0.109***	0.162***	-	-
Ee	-0.476	-0.472	-	-
Moran's I	0	-0.024	-	-
$F_{IS\ entire}$	-0.058**	-0.262***	0.055 ^{NS}	0.182***
$F_{IS\ truncated}$	0.110**	-0.037 ^{NS}	0.213***	0.279***
H_E	0.485 (0.182)	0.428 (0.215)	0.559 (0.171)	0.586 (0.141)
H_O	0.432 (0.307)	0.443 (0.329)	0.444 (0.252)	0.424 (0.292)

N_{total} : number of colonies sampled; N : number of colonies presenting no missing data in their multi-locus genotypes (MLG); N_{MLG} : number of MLG; R : clonal richness; G : genotypic diversity; ED^* : Simpson's evenness; β : parameter of the Pareto distribution (-1*regression slope); Ac : aggregation index; Ee : edge effect; Moran's I : spatial autocorrelation index, $F_{IS\ entire}$: values are estimated on the entire dataset; $F_{IS\ truncated}$: values are estimated on the truncated dataset (one representative per MLG and per population); H_E (s.d.) and H_O (s.d.): expected and observed heterozygosities respectively, values are estimated using the truncated dataset. ^{NS}: non-significant; ** $P < 0.01$; *** $P < 0.001$.

Spatial analysis examining MLG distribution were performed within REU1 and REU2 sites. The aggregation coefficient was significantly positive in both sites ($Ac = 0.109$, $P = 0.004$ and $Ac = 0.162$, $P < 10^{-3}$, respectively), suggesting that colonies belonging to the same clone were spatially closer than colonies belonging to different clones. However, spatial autocorrelation analysis did not show a significantly higher relatedness among geographically close MLGs as compared to distant ones (REU1: $I = 0.000$, $P = 1.000$; REU2: $I = -0.024$, $P = 0.611$). Also, no significant edge effect was detected in any site (REU1: $Ee = -0.018$, $P = 0.620$; REU2: $Ee = -0.472$, $P = 0.991$), suggesting that new MLGs were not especially detected on the edge of the sampling zone (Fig 2).

Keeping only one representative per MLG, no evidence of linkage disequilibrium was detected between loci; none of the 78 tests were significant at the 0.05 level after correction for multiple

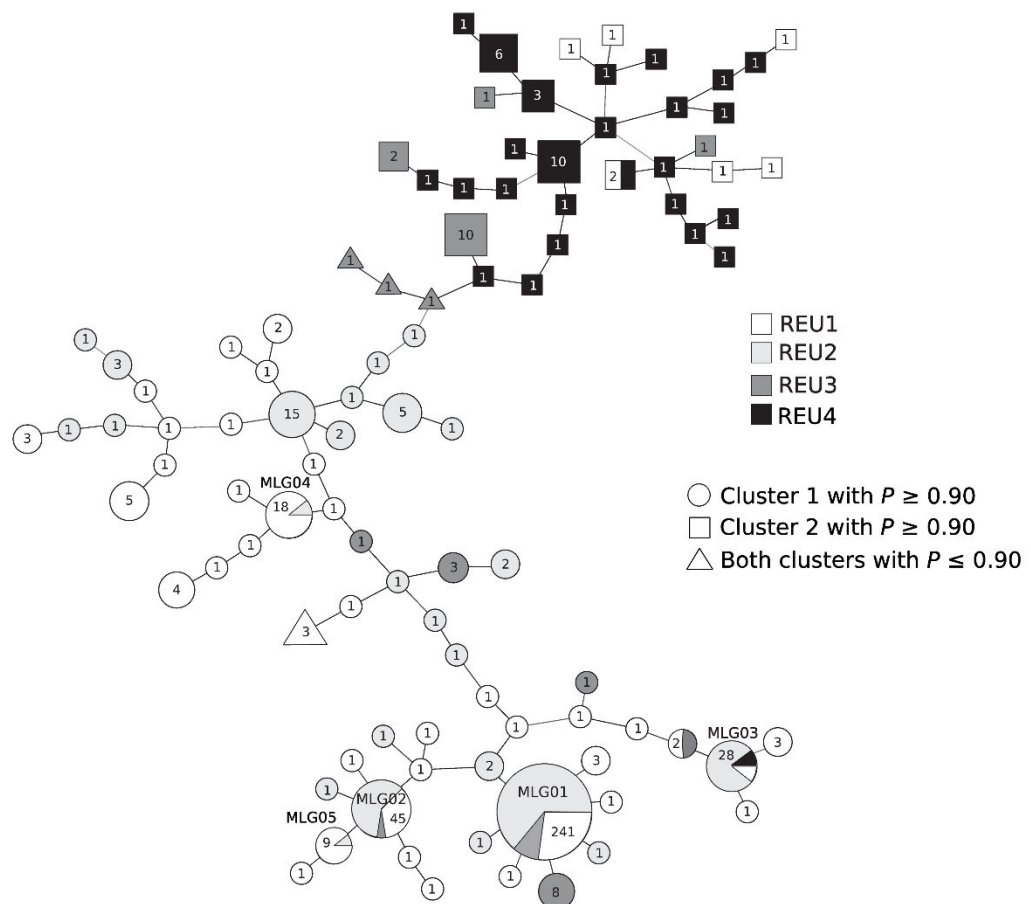


Fig 3: Network of multi-locus genotypes (MLG) based on shared allele distance. The number of colonies sharing each MLG is indicated inside each node and the geographic origin of the colonies is represented by colors. The node shape represents the result of the clustering analysis: circles, MLGs assigned to cluster 1 (see Fig 4) with a probability ≥ 0.90 ; squares, MLGs assigned to cluster 2 (see Fig 4) with a probability ≥ 0.90 ; triangles, MLGs assigned to one cluster or the other with a probability ≤ 0.90 .

testing (Bonferroni). Then, all loci were interpreted as independent loci. Two loci (Pd2-001 and

Poc40) evidenced null alleles in all populations; they were kept for analysis of clonal diversity as they allowed discriminating different MLGs, but removed for the following population structure analyses. F_{IS} values (Table 1) estimated on the entire dataset ($N = 510$) were either significantly negative for two populations REU1 and REU2 (-0.058 , $P < 10^{-2}$; -0.262 , $P < 10^{-3}$, respectively) or significantly positive for REU4 (0.182 ; $P < 10^{-3}$), or not significantly different from zero for REU3. When estimations were performed on the truncated dataset (*i.e.*, one representative per MLG and per population kept; $N = 109$), F_{IS} values increased and became significantly positive except for REU2 (-0.037 , $P > 0.05$; Table 1).

When including all colonies, each population was highly differentiated from the others, with pairwise F_{ST} estimates ranging from 0.016 ($P < 10^{-3}$) between REU1 and REU2, to 0.428 ($P < 10^{-3}$) between REU2 and REU4 (Table 2). Noteworthy, D_{est} estimates were of the same order of magnitude, ranging from 0.012 ($P < 10^{-3}$) between REU1 and REU2, to 0.376 ($P < 10^{-3}$) between REU2 and REU4 (Table 2). These results remained nearly identical when F_{ST} were estimated on the truncated dataset, though they were approximately reduced by a 1.5 order (Table 2).

Table 2: Genetic differentiation between *Pocillopora damicornis* type β populations estimated with Weir and Cockerham's F_{ST} [58] and with Jost's D_{est} (in parentheses, [59]). Values in the lower matrix are estimated on the entire data set; values in the upper matrix are estimated on the truncated data set. Values in bold are significant ($P < 0.05$).

	REU1	REU2	REU3	REU4
REU1	-	0.011 (0.010)	0.068 (0.092)	0.202 (0.293)
REU2	0.016 (0.012)	-	0.094 (0.088)	0.262 (0.339)
REU3	0.090 (0.067)	0.126 (0.066)	-	0.081 (0.185)
REU4	0.348 (0.349)	0.428 (0.376)	0.181 (0.250)	-

Results of the clustering analyses led to similar results according to the DIC method or the Evanno's one (Appendix S3). Both methods gave an optimal K of 2 for both datasets, the MLGs

being assigned to the same cluster whatever the dataset used, naturally implying that colonies sharing the same MLG were assigned to the same cluster (Figs 4a and 4b). The PCA (based on the information from either 13 or 11 microsatellites), when keeping the two first axes, gave the same partition in two groups corresponding to the two clusters identified by the Bayesian analysis (Fig 4c). Thus, with few individual exceptions, the two clusters segregated the northern populations (REU1 and REU2) from the southernmost population (REU4) (Fig 4). Indeed, the two northernmost populations (REU1 and REU2) were almost exclusively composed by

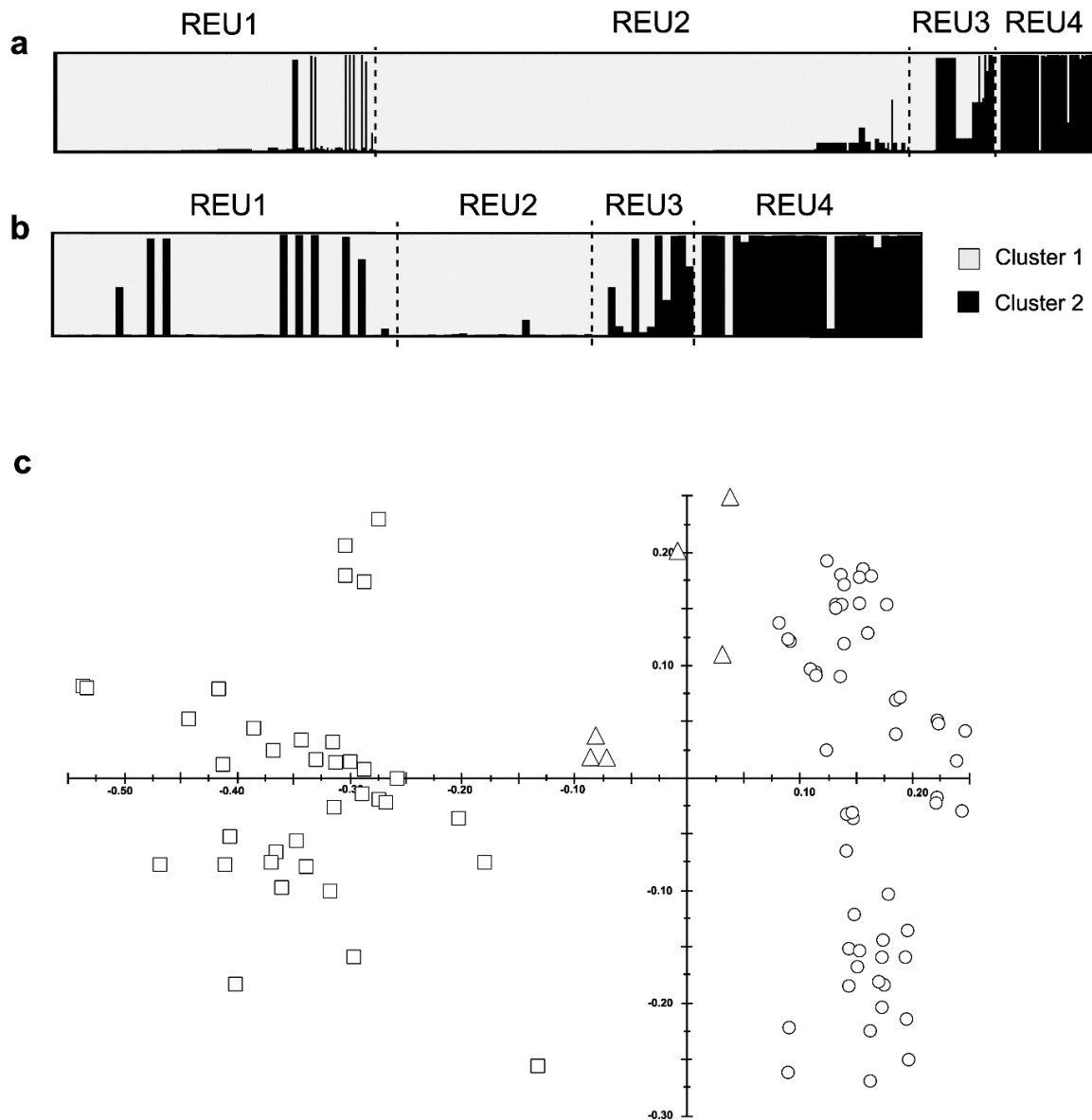


Fig 4: Results of the clustering analysis for $K = 2$, (a) for the entire dataset ($N = 510$) and (b) for the truncated dataset containing one representative per multi-locus genotype and per site ($N = 109$). For each plot, populations are separated with dash lines. (c) Principal Component Analysis estimated on the truncated dataset (based on 13 microsatellites). Axes 1 (horizontal) and 2 (vertical) represent 31.56% and 10.88% of the total variance, respectively. Symbols are the same than in Fig 3.

colonies assigned to cluster 1, and nearly all colonies of REU4 were assigned to cluster 2. REU3, located nearly half-distance between REU2 and REU4, was characterized by a mix of both clusters: over 13 MLGs, six were assigned to the northern cluster (cluster 1; probability ≥ 0.90), four to the southern cluster (cluster 2; probability ≥ 0.90) and three for which the probability to be assigned to one or the other cluster was less than 0.68 (Figs 3 and 4b).

Discussion

Our results revealed a high proportion of repeated MLGs per reef, on all studied sites, suggesting high rate of asexual reproduction in *Pocillopora damicornis* β from Reunion Island reefs. One superclone (representing 47% of all studied colonies) was even found in all sites, suggesting that long-distance dispersal (40 km) of some clones is not uncommon. Despite, this long dispersal, a small scale genetic differentiation was revealed between Northern and Southern populations.

The use of a random sampling scheme is theoretically the most pertinent way to assess clonality [46] as every possible sampling unit would present equal probability of being included in the population. It also allows to minimize bias in estimation of diversity and richness indices [46]. However, it is the most difficult to proceed practically and this scheme could not be applied in all our four study sites, but only in REU2. Beyond the sampling scheme, the sampling effort is also determinant to reduce bias in index estimations [46], which was taken into account in our sampling scheme. Despite all precautions, the exhaustive sampling only in the 2 m radius central circle in REU2, as well as the colony patchiness in REU1 may have biased our estimated indices. From another angle, edge effect and spatial autocorrelation indices indicated that new MLGs were not especially detected on the edge of the sampling scheme and were randomly distributed, which indicates that the bias is minimized.

Superclone expansion

In the lagoon of La Saline–L’Ermitage complex where REU1 and REU2 sites are located, we found the co-existence of one “giant” clone (represented by a large number of colonies sharing MLG01) and several less represented ones, with 71% of our sampled colonies presenting one of the seven most abundant MLGs (out of 99) (as revealed by the slow values

of the parameter β of the Pareto distribution and the Simpson index). Such result of over-representation of few clones has already been observed in *P. damicornis sensu lato* on a patch reef in Kane'ohe Bay, O'ahu (Hawaii) [28], where 70% of the sampled colonies represented only 10% of the MLGs (7 out of 78). This over-representation of one genotype could be the result of asexually produced larvae [22, 31]. Nevertheless, Torda et al [69], studying *P. damicornis* type β on the Great Barrier Reef, found that clonal larvae only represented 7% of the brooded larvae. Moreover, considering that *P. damicornis* β presents a rather thin morphology combined with a breakable skeleton, the high dominance of one genotype in this northernmost lagoon is likely the result of fragmentation, mainly induced by the exposition to waves and touristic activities (confirmed by the significantly positive aggregation coefficient in REU1 and REU2 sites). Also, no significant edge effect was detected in any site (REU1: $Ee = -0.018$, $P = 0.620$; REU2: $Ee = -0.472$, $P = 0.991$), suggesting that new MLGs were not especially detected on the edge of the sampling zone (Fig 2). Indeed, the very low depth (approx. 1 m) of this lagoon makes it highly frequented by bathers and snorkelers and we observed some living fragments not yet fixed on the substratum. Additionally, the occurrence of clones may as well result from polyp bail-out [14], a well-known phenomenon by aquarium hobbyists dealing with *P. damicornis sensu lato* but poorly documented.

The expansion of one superclone might result from a combined effect of asexual propagation with high fitness related to particular environmental conditions [28, 70]. Such phenomena start to be investigated in corals with some studies reporting superclone dominance in different species, like the monoclonal population of *Acropora palmata* in a Caribbean reef [71] or for *P. damicornis sensu lato* in Hawaii [28] and for *Pocillopora* type 1 in the Tropical Eastern Pacific [27]: in these cases, the dominant clone is likely well-adapted to local environmental conditions and its expansion reduces the genotypic diversity on reefs. Indeed, this large superclone dominance could lead to a negative fate of the population as illustrated in two stable reefs from Florida Keys which presented allelic and genotypic losses over four years in highly clonal populations of *A. palmata* [72]. Moreover, in case of perturbation (e.g., warming, pollution or pathogen emergence), the over-represented genotype might not be longer well-adapted to the new conditions and the population could collapse, while a population presenting a higher genotypic diversity is more likely to resist such a perturbation, as it presents a higher probability that one or some of its genotypes could cope [73].

The higher genotypic diversity and clonal richness found in REU4, the southernmost site, may be due to its exposition to stronger swells and waves during the southern winter,

contrasting with REU1 and REU2 sites which are located in less exposed reefs with weaker hydrodynamics [36]. Indeed, genotypic diversity is found to be related to high disturbance history [74]. In REU4, colonies were more compact and the branches shorter and tighter than those from the other sites, resulting in a less breakable skeleton and lower clonal propagation. The presence of these particular *Pocillopora* morphotypes explains also the misidentification with *P. verrucosa* and *P. eydouxi* that occurred during sampling.

Long-distance dispersal

Around the island, we found seven MLGs shared by several populations, suggesting a strong clonal propagation. The two furthestmost populations sharing identical MLGs were located nearly 40 km apart (REU1 and REU4). To date, two studies found a similar result, with clones shared among populations 40 km apart [25] and 200 km apart [34]. As the estimated probability that two colonies sharing the same MLG were identical by chance (*i.e.*, P_{ID}) is less than 10^{-4} (recommended threshold when using more than 10 loci presenting high allelic richness [53]), it is most likely that the identical MLGs found result from asexual propagation. Moreover, fragmentation, while being an obvious explanation at the intra-reef levels, becomes a weak hypothesis to explain the occurrence of clones over such large distances. Indeed, over higher distances, it seems unlikely that coral fragments could break off from one reef, drift for several kilometres and be fixed again on another reef, especially in our case where the two reefs in question are separated by high water depths and “no-reef” zones. Polyp bail-out seems also unlikely at this scale, for the same reasons exposed previously and because of their likely negative buoyancy and limited power of mobility, though this hypothesis cannot be ruled out as free polyps of *Seriatopora hystrix* can leave up to 9 days before re-attaching to the bottom [14].

An alternative hypothesis is that distant clones could be the result of parthenogenetic larvae [22, 31] which appears to be a more solid explanation in our case. Indeed, it has been hypothesized that asexually produced larvae (larger and better-provisioned) may be more effective to disperse than sexual progenies issued from broadcast spawning [75], these latter tending to settle locally [76]. A mixed reproduction mode has moreover already been shown for *P. damicornis sensu lato* [15, 22], with parthenogenetic larvae production dealing with sperm limitation, a general issue for broadcast sperm-dependent marine animals [18]. Anyhow, neither parthenogenetic larvae nor released polyps hypotheses can be verified directly in our

study. Additional experiments in controlled environments are needed to fully conclude on their role in clonal propagation.

Two highly differentiated lineages

The two genetically distinct clusters revealed among all sampled colonies by the Bayesian assignment test more or less differentiate northern from southern populations, whatever the dataset used (truncated and entire datasets). We acknowledge that keeping all individuals (entire dataset) for the genetic structuring analysis may introduce a bias, but no more than keeping one representative per MLG (truncated dataset). Nonetheless, our analyses using both datasets gave exactly the same individual assignment suggesting that the pattern we observed here is robust. Despite the large genetic divergence among the two clusters ($F_{ST} = 0.209$), all colonies included in this analysis shared the exact same mtDNA ORF haplotype, assigned to *P. damicornis* β . Pairwise F_{ST} estimates involving comparisons between northern and southern populations were consequently high, whatever the dataset used. Such estimates are in the order of magnitude of what could be expected between different species using these markers, like for example between two sympatric *Pocillopora* types, based on seven microsatellite loci (type 1 and type 5; $F_{ST} = 0.346$; [24]). Other studies conducted on *P. damicornis sensu lato* showed similar high levels of genetic differentiation, such as found among Pacific populations (*e.g.*, [30]), or in the Gulf of Panama between populations separated by approximately 15 km ($F_{ST} > 0.22$; [76]), but in these examples, high values could result from a mix of different lineages of *P. damicornis sensu lato* in different proportions. All in all, the high F_{ST} estimates reflect the allele frequencies variance between two genetically distant clonal lineages (one in the North and one in the South), each composed of several genetically close MLGs (Fig 3).

Altogether, our results showed on one side, the detection of shared clones among distant sites indicates that larvae dispersal among reefs is not uncommon, while on the other side, the significant genetic structure among the four sites suggests that gene flows appeared limited along the west coast of Reunion Island. Our result contrasts with the genetic structure to be expected from a species with a larval survivorship estimated to reach 100 days [77, 78]. Indeed, with such long larval duration, one would expect to find limited population genetic differentiation due to potential high larval dispersal abilities [79, 80], which is not what we found, nor what was observed in a previous study conducted in the same region, South Western

Indian Ocean, and on the same lineage, *P. damicornis* β ([24]; named *Pocillopora* type 5 therein). Additionally, at very small scale (approximately 1 km), our results are concordant with another study conducted in Hawaii [28] revealing no genetic differentiation, as between REU1 and REU2, separated by 2 km.

Apart from geographic distances, various factors are known to modify the predicted dispersal abilities, such as ocean currents, planktonic environment or larval behaviour and settlement ability [81]. In the Indian Ocean, the South Equatorial Current (SEC), circulating from east to west between 7-8°S and 23-25°S and bathing the South coast of Reunion Island [82] is likely to create a geographic barrier between the northern and southern populations, showing different hydrodynamic conditions. Thus the fact that REU4, the southernmost population, appeared as a group genetically distinct from the three other populations could therefore result (1) from a particular connection of the southern part of Reunion Island with other reefs from Mauritius or Rodrigues through the SEC and (2) from the flushing out of larvae broadcasted from REU4 to the west (towards Madagascar), without going to the north of Reunion Island. Besides, on the west coast of Reunion Island (REU1, REU2 and REU3), oceanic currents are complex, due to local scale eddies created by the SEC residues [37]. Then, larvae from elsewhere might have some issues to reach the northern reefs and larvae produced on the west coast might be retained by local currents. On the other hand, this structuring pattern could also reflect the presence of two different lineages partially reproductively isolated in the *P. damicornis* β , as recently highlighted in *Acropora* corals from Western Australia [83], that may result from asynchronous spawning. Reviewing studies focusing on *P. damicornis sensu lato* reproduction in light of new taxonomic insights Schmidt-Roach et al [84] concluded that *P. damicornis* β broods its larvae and releases them after the new moon, produces parthenogenetic larvae and spawns male gametes. To our knowledge, asynchronous release of gametes or larvae has not been reported yet for *P. damicornis* β . However, two *P. damicornis* populations of the Great Barrier Reef located 13 km apart, characterised *a posteriori* as type α by Schmidt-Roach et al [1], showed differences in larvae release dates: from October to November in One Tree Island [31] and from September to February in Heron Island [85] underlying the possibility of asynchronous spawning events over small spatial scales.

In Reunion Island, although the spawning period of *P. damicornis* β is unknown, the hypothesis of asynchronous spawning cannot be ruled out to explain the presence of these two genetically distinct clusters whose distributions overlap in REU3. Indeed, one could imagine that, considering the short gamete lifespan in the water column in scleractinians (*e.g.*, < 24h for *Leptoria*, *Favites*; [86]), asynchronous spawning could occur at short temporal scale (*i.e.*,

several days) or between two consecutive moon cycles (28 days). Spawning events are determined by a combination of environmental conditions, including sea temperature or accumulated degree days, an empirical measure of heat requirements for growth and development, largely used in agronomy for crops or insect pests, or in aquaculture for egg incubation or gamete maturation (*e.g.*, in mussels; [87]). In order to test whether distinct genetic clusters could result from asynchronicity related to different sea temperature regimes, we investigated the difference in cumulative degree days between the two remote sites (REU1 and REU4) from 6th June 2015 to 18th December 2015 (Appendix S4a; data provided by GIP RNMR/Marex). Over this time period, we estimated a mean difference of temperature of + 0.4°C per day between REU1 (warmer; mean temperature (\pm s.e.) = 25.0°C \pm 0.087) and REU4 (cooler; mean temperature = 24.6°C \pm 0.087). Over the 196 days of survey, the cumulative temperature difference between REU1 and REU4 reached 78.4°C. When extrapolating this difference over one year (taking the mean temperature for each site as a proxy for daily temperature for the 169 remaining days, *i.e.*, a temperature difference of + 0.4°C per day between REU4 and REU1), we found a cumulative difference of 150.1°C between both sites (Appendix S4b). Assuming that an equal cumulative temperature threshold is needed for gametes to mature in both populations, this difference, for a given threshold, would correspond to a 6-days delay in gamete maturation between REU1 and REU4. Would this temperature difference, considering all the other interacting parameters, be sufficient to create a discrepancy in phenology and lead to a partial reproductive isolation between these two sympatric divergent lineages? To answer this question, more information on *P. damicornis* β reproduction in this part of the world (threshold temperature for gamete maturation, maturation start and duration, spawning event date) is obviously needed.

While clonality has been documented for *P. damicornis* β , here we evidenced its clonal propagation for the first time in Reunion Island, South Western Indian Ocean. Indeed, around this island, asexual propagation plays a predominant role in populations of *P. damicornis* β , reproducing asexually by fragmentation and likely through the production of parthenogenetic larvae for long-distance dispersal (40 km). Thus, in case of high environmental and anthropogenic disturbances or global environmental changes, *P. damicornis* β populations from Reunion Island exhibiting low genotypic richness, such as those of La Saline-L'Ermitage lagoon, might be extremely vulnerable, implying dramatic consequences for the reef structure functioning.

Acknowledgments

This research was conducted with permission of the regional authorities of marine affairs (decision 14-2013/DEAL/SEB/UBIO) and the marine park authorities of Reunion Island (RNMR). We gratefully acknowledge Christophe Menkès and Francesca Benzoni for helpful discussion and the Plateforme Gentyane of the Institut National de la Recherche Agronomique (INRA, Clermont-Ferrand, France) for genotyping support.

Supporting Information

Appendix S1: List of the loci used in this study. \hat{A} , the allelic richness, was estimated as the average number of alleles per locus on the basis of the smallest sample size (REU3, $N_{MLG} = 13$).

Appendix S2: Pareto distribution of *Pocillopora damicornis* β multi-locus genotypes for each site.

Appendix S3: Plots of the mean Deviation Information Criterion (DIC) and the $\Delta(K)$ are presented both for the truncated dataset (a. and b., respectively) and for the entire dataset (c. and d., respectively). In each case, an arrow indicates the most likely number of clusters.

Appendix S4: (a) Water temperature in REU1 and REU4 lagoons from 6th June 2015 to 18th December 2015 and (b) cumulative day temperature for water lagoons in REU1 and REU4 over one year focusing on the last 40 days (from 7th November 2015 to 18th December 2015).

Reference list

1. Schmidt-Roach S, Lundgren P, Miller KJ, Gerlach G, Noreen AME, Andreakis N. Assessing hidden species diversity in the coral *Pocillopora damicornis* from Eastern Australia. *Coral Reefs*. 2013;32(1):161-72.
2. Holman EW. Recognizability of sexual and asexual species of rotifers. *Syst Biol*. 1987;36(4):381-6.
3. Otto SP, Lenormand T. Resolving the paradox of sex and recombination. *Nat Rev Genet*. 2002;3(4):252-61.
4. Crow JF. The omnipresent process of sex. *J Evol Biol*. 1999;12:1023-5.
5. Fisher RA. *The genetical theory of natural selection: a complete variorum edition*: Oxford University Press; 1930.
6. Muller HJ. Some genetic aspects of sex. *Am Nat*. 1932;66(703):119-38.
7. Muller HJ. The relation of recombination to mutational advance. *Mutat Res*. 1964;1(1):2-9.
8. Kondrashov AS. Deleterious mutations and the evolution of sexual reproduction. *Nature*. 1988;336(6198):435-40.

9. Crow JF. Advantages of sexual reproduction. *Dev Genet.* 1994;15(3):205-13.
10. Flot J-F, Hespeels B, Li X, Noel B, Arkhipova I, Danchin EGJ, et al. Genomic evidence for ameiotic evolution in the bdelloid rotifer *Adineta vaga*. *Nature.* 2013;500(7463):453-7.
11. Neiman M, Sharbel TF, Schwander T. Genetic causes of transitions from sexual reproduction to asexuality in plants and animals. *J Evol Biol.* 2014;27(7):1346-59.
12. Harrison PL. Sexual reproduction of scleractinian corals. *Coral reefs: an ecosystem in transition: Springer Netherlands;* 2011. p. 59-85.
13. Highsmith RC. Reproduction by fragmentation in corals. *Mar Ecol Prog Ser.* 1982;7:207-26.
14. Sammarco PW. Polyp bail-out: an escape response to environmental stress and new means of reproduction in corals. *Mar Ecol Prog Ser.* 1982;10:57-65.
15. Stoddart JA. Asexual production of planulae in the coral *Pocillopora damicornis*. *Mar Biol.* 1983;76:279-84.
16. Miller KJ, Ayre DJ. The role of sexual and asexual reproduction in structuring high latitude populations of the reef coral *Pocillopora damicornis*. *Heredity.* 2004;92(6):557-68.
17. Ayre DJ, Resing JM. Sexual and asexual production of planulae in reef corals. *Mar Biol.* 1986;90:187-90.
18. Yund PO. How severe is sperm limitation in natural populations of marine free-spawners? *Trends Ecol Evol.* 2000;15(1):10-3.
19. Veron JEN. *Corals of the world: Australian Institute of Marine Science;* 2000.
20. Veron JEN, Pichon MM. Scleractinia of Eastern Australia. Part I: Families Thamnasteriidae, Astrocoeniidae, Pocilloporidae. In: *Science AIoM, editor.* 1976.
21. Ward S. Evidence for broadcast spawning as well as brooding in the scleractinian coral *Pocillopora damicornis*. *Mar Biol.* 1992;112(4):641-6.
22. Combosch DJ, Vollmer SV. Mixed asexual and sexual reproduction in the Indo-Pacific reef coral *Pocillopora damicornis*. *Ecol Evol.* 2013;3(10):3379-87.
23. Schmidt-Roach S, Miller KJ, Lundgren P, Andreakis N. 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.* 2014;170(1):1-33.
24. Pinzón JH, Sampayo E, Cox E, Chauka LJ, Chen CA, Voolstra CR, et al. Blind to morphology: genetics identifies several widespread ecologically common species and few endemics among Indo-Pacific cauliflower corals (*Pocillopora*, Scleractinia). *J Biogeogr.* 2013;40(8):1595-608.
25. Souter P, Henriksson O, Olsson N, Grahn M. Patterns of genetic structuring in the coral *Pocillopora damicornis* on reefs in East Africa. *BMC Ecol.* 2009;9:19.
26. Marti-Puig P, Forsman ZH, Haverkort-Yeh RD, Knapp ISS, Maragos JE, Toonen RJ. Extreme phenotypic polymorphism in the coral genus *Pocillopora*; micro-morphology corresponds to mitochondrial groups, while colony morphology does not. *Bull Mar Sci.* 2014;90(1):211-31.
27. Pinzón JH, Reyes-Bonilla H, Baums IB, LaJeunesse TC. Contrasting clonal structure among *Pocillopora* (Scleractinia) communities at two environmentally distinct sites in the Gulf of California. *Coral Reefs.* 2012;31(3):765-77.
28. Gorospe KD, Karl SA. Genetic relatedness does not retain spatial pattern across multiple spatial scales: dispersal and colonization in the coral, *Pocillopora damicornis*. *Mol Ecol.* 2013;22(14):3721-36.
29. Gorospe KD, Karl SA. Depth as an organizing force in *Pocillopora damicornis*: intra-reef genetic architecture. *PLoS One.* 2015;10(3):e0122127.
30. Ayre DJ, Hughes TP. Genotypic diversity and gene flow in the brooding and spawning corals along the Great Barrier Reef, Australia. *Evolution.* 2000;54:1590-605.

31. Ayre DJ, Miller KJ. Where do clonal coral larvae go? Adult genotypic diversity with reproductive effort in the brooding coral *Pocillopora damicornis*. *Mar Ecol Prog Ser*. 2004;277:95-105.
32. Sherman CDH, Ayre DJ, Miller KJ. Asexual reproduction does not produce clonal populations of the brooding coral *Pocillopora damicornis* on the Great Barrier Reef, Australia. *Coral Reefs*. 2005;25(1):7-18.
33. Starger CJ, Barber PH, Ambariyanto, Baker AC. The recovery of coral genetic diversity in the Sunda Strait following the 1883 eruption of Krakatau. *Coral Reefs*. 2010;29(3):547-65.
34. Adjeroud M, Guérécheau A, Vidal-Dupiol J, Flot J-F, Arnaud-Haond S, Bonhomme F. Genetic diversity, clonality and connectivity in the scleractinian coral *Pocillopora damicornis*: a multi-scale analysis in an insular, fragmented reef system. *Mar Biol*. 2014;161:531-41.
35. Torda G, Lundgren P, Willis BL, van Oppen MJ. Revisiting the connectivity puzzle of the common coral *Pocillopora damicornis*. *Mol Ecol*. 2013;22(23):5805-20.
36. Denis V. Capacités et modalités d'adaptation de deux espèces de coraux zooxanthellés aux perturbations climatiques et anthropiques (île de la Réunion, Sud-Ouest de l'océan Indien): Université de la Réunion; 2010.
37. Pous S, Lazure P, André G, Dumas F, Halo I, Penven P. Circulation around La Réunion and Mauritius islands in the south-western Indian Ocean: A modeling perspective. *J Geophys Res Oceans*. 2014;119(3):1957-76.
38. Postaire B, Aurelle D, Bourmaud CAF, Bruggemann JH, Magalon H. Isolation and characterisation of 26 microsatellite loci from a widespread tropical hydrozoan, *Macrorhynchia phoenicea* (Leptothecata, Aglaopheniidae), and cross-amplification in closely related species. *Biochem Syst Ecol*. 2015;62:137-41.
39. Flot JF, Tillier S. The mitochondrial genome of *Pocillopora* (Cnidaria: Scleractinia) contains two variable regions: the putative D-loop and a novel ORF of unknown function. *Gene*. 2007;401(1-2):80-7.
40. Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, et al. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics*. 2012;28(12):1647-9.
41. Starger CJ, Yeoh SS, Dai CF, Baker AC, Desalle R. Ten polymorphic STR loci in the cosmopolitan reef coral, *Pocillopora damicornis*. *Mol Ecol Resour*. 2007;8(3):619-21.
42. Pinzón JH, LaJeunesse TC. Species delimitation of common reef corals in the genus *Pocillopora* using nucleotide sequence phylogenies, population genetics and symbiosis ecology. *Mol Ecol*. 2011;20(2):311-25.
43. Magalon H, Samadi S, Richard M, Adjeroud M, Veuille M. Development of coral and zooxanthella-specific microsatellites in three species of *Pocillopora* (Cnidaria, Scleractinia) from French Polynesia. *Mol Ecol Notes*. 2004;4(2):206-8.
44. Torda G, Schmidt-Roach S, Peplow LM, Lundgren P, van Oppen MJH. A rapid genetic assay for the identification of the most common *Pocillopora damicornis* genetic lineages on the great barrier reef. *PLoS One*. 2013;8(3):e58447.
45. Arnaud-Haond S, Belkhir K. GENCLONE: a computer program to analyse genotypic data, test for clonality and describe spatial clonal organization. *Mol Ecol Notes*. 2007;7:15-7.
46. Arnaud-Haond S, Duarte CM, Alberto F, Serrao EA. Standardizing methods to address clonality in population studies. *Mol Ecol*. 2007;16(24):5115-39.
47. van Oppen MJ, Peplow LM, Kininmonth S, Berkelmans R. Historical and contemporary factors shape the population genetic structure of the broadcast spawning coral, *Acropora millepora*, on the Great Barrier Reef. *Mol Ecol*. 2011;20(23):4899-914.
48. Meirmans PG, Van Tienderen PH. Genotype and Genodive: two programs for the analysis of genetic diversity of asexual organisms. *Mol Ecol Notes*. 2004;4(4):792-4.

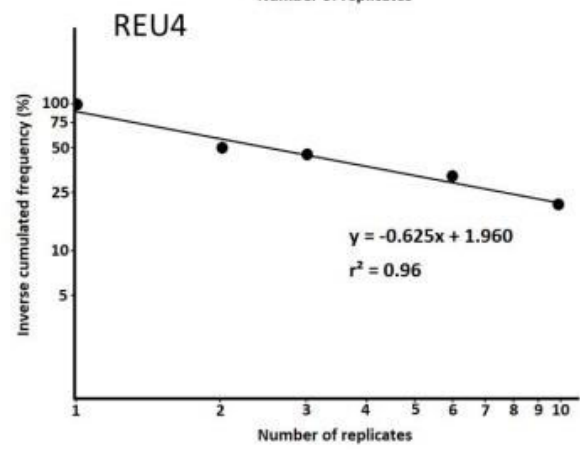
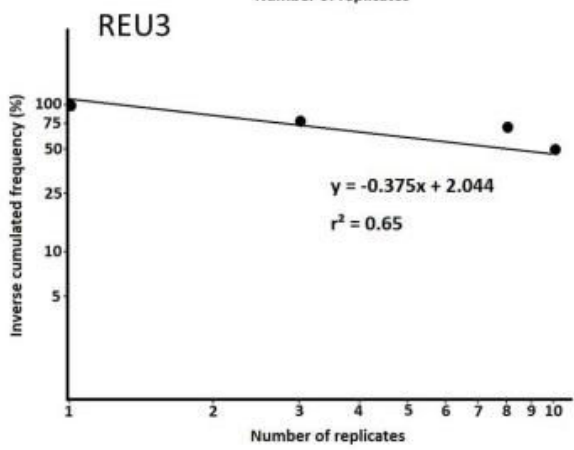
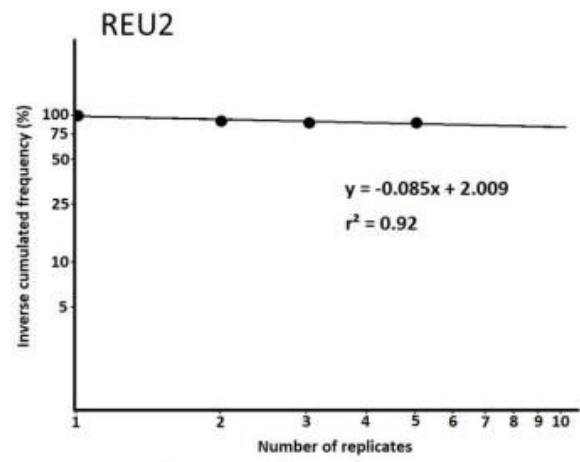
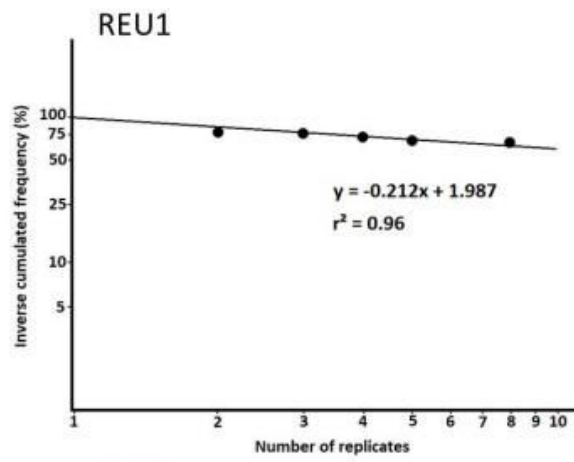
49. Dorken ME, Eckert CG. Severely reduced sexual reproduction in northern populations of a clonal plant, *Decodon verticillatus* (Lythraceae). *J Ecol.* 2001;89(3):339-50.
50. Stoddart JA, Taylor JF. Genotypic Diversity: Estimation and Prediction in Samples. *Genetics* 1988;118:705-11.
51. Baums IB, Miller MW, Hellberg ME. Geographic variation in clonal structure in a reef-building caribbean coral, *Acropora palmata*. *Ecol Monogr.* 2006;76(4):503-19.
52. Goudet J. FSTAT, a program to estimate and test gene diversities and fixation indices version 2.9. 3.2, updated from Goudet 1995. FSTAT; 2001.
53. Waits LP, Luikart G, Taberlet P. Estimating the probability of identity among genotypes in natural populations: cautions and guidelines. *Mol Ecol.* 2001;10(1):249-56.
54. Valière N. gimlet: a computer program for analysing genetic individual identification data. *Mol Ecol Notes.* 2002;2(3):377-9.
55. Hardy OJ, Vekemans X. SPAGeDi: a versatile computer program to analyse spatial genetic structure at the individual or population levels. *Mol Ecol Notes.* 2002;2:618-20.
56. van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P. Micro-checker: software for identifying and correcting genotyping errors in microsatellite data. *Mol Ecol Notes.* 2004;4(3):535-8.
57. Belkhir K, Borsa P, Chikhi L, Raufaste N, Bonhomme F. GENETIX 4.05, logiciel sous Windows TM pour la génétique des populations. Montpellier (France): Laboratoire Génome, Populations, Interactions, CNRS UMR 5000, Université de Montpellier II 2004.
58. Weir B, Cockerham C. Estimating F-statistics for the analysis of population structure. *Evolution.* 1984;38:1358-70.
59. Jost L. G_{ST} and its relatives do not measure differentiation. *Mol Ecol.* 2008;17:4015-26.
60. Alberto F, Gouveia L, Arnaud-Haond S, Perez-Llorens JL, Duarte CM, Serrao EA. Within-population spatial genetic structure, neighbourhood size and clonal subrange in the seagrass *Cymodocea nodosa*. *Mol Ecol.* 2005;14(9):2669-81.
61. Gao H, Williamson S, Bustamante CD. A Markov chain Monte Carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. *Genetics.* 2007;176(3):1635-51.
62. Gao H, Bryc K, Bustamante CD. On identifying the optimal number of population clusters via the Deviance Information Criterion. *PLoS One.* 2011;6(6):e21014.
63. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol.* 2005;14(8):2611-20.
64. Rosenberg NA. DISTRUCT: a program for the graphical display of population structure. *Mol Ecol Notes.* 2004;4(1):137-8.
65. Jakobsson M, Rosenberg NA. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics.* 2007;23(14):1801-6.
66. Perrier X, Jacquemoud-Collet J. DARwin software. 2006.
67. Kivelä M, Arnaud-Haond S, Saramäki J. EDENetworks: A user-friendly software to build and analyse networks in biogeography, ecology and population genetics. *Mol Ecol Resour.* 2015;15(1):117-22.
68. Gitzendanner MA, Weekley CW, Germain-Aubrey CC, Soltis DE, Soltis PS. Microsatellite evidence for high clonality and limited genetic diversity in *Ziziphus celata* (Rhamnaceae), an endangered, self-incompatible shrub endemic to the Lake Wales Ridge, Florida, USA. *Conserv Genet.* 2011;13(1):223-34.
69. Torda G, Lundgren P, Willis BL, van Oppen MJH. Genetic assignment of recruits reveals short and long distance larval dispersal in *Pocillopora damicornis* on the Great Barrier Reef. *Mol Ecol.* 2013;22:5821-34.

70. Caron V, Ede FJ, Sunnucks P. Unravelling the paradox of loss of genetic variation during invasion: superclones may explain the success of a clonal invader. *PLoS One*. 2014;9(6):e97744.
71. Japaud A, Bouchon C, Manceau J-L, Fauvelot C. High clonality in *Acropora palmata* and *Acropora cervicornis* populations of Guadeloupe, French Lesser Antilles. *Mar Freshw Res*. 2015;66:847-51.
72. Williams DE, Miller MW, Baums IB. Cryptic changes in the genetic structure of a highly clonal coral population and the relationship with ecological performance. *Coral Reefs*. 2014;33(3):595-606.
73. Hughes AR, Stachowicz JJ. Genetic diversity enhances the resistance of a seagrass ecosystem to disturbance. *Proc Natl Acad Sci USA*. 2004;101(24):8998-9002.
74. Hunter CL. Genotypic variation and clonal structure in coral populations with different disturbance histories. *Evolution*. 1993;47(4):1213-28.
75. Schmidt-Roach S, Miller KJ, Woolsey E, Gerlach G, Baird AH. Broadcast spawning by *Pocillopora* species on the Great Barrier Reef. *PLoS One*. 2012;7(12):e50847.
76. Combosch DJ, Vollmer SV. Population genetics of an ecosystem-defining reef coral *Pocillopora damicornis* in the Tropical Eastern Pacific. *PLoS One*. 2011;6(8):e21200.
77. Richmond RH. Energetics, competency, and long-distance dispersal of planula larvae of the coral. *Mar Biol*. 1987;93:527-33.
78. Harii S, Kayanne H, Takigawa H, Hayashibara T, Yamamoto M. Larval survivorship, competency periods and settlement of two brooding corals, *Heliopora coerulea* and *Pocillopora damicornis*. *Mar Biol*. 2002;141(1):39-46.
79. Bohonak AJ. Dispersal, gene flow and population structure. *Q Rev Biol*. 1999;74(1):21-44.
80. Shanks AL. Pelagic larval duration and dispersal distance revisited. *Biol Bull*. 2009;216:373-85.
81. Hohenhole PA. Limits to gene flow in marine animals with planktonic larvae: models of *Littorina* species around Point Conception, California. *Biol J Linn Soc*. 2004;82:169-87.
82. Schott FA, Xie S-P, McCreary JP. Indian Ocean circulation and climate variability. *Rev Geophys*. 2009;47(1):n/a-n/a.
83. Rosser NL. Asynchronous spawning in sympatric populations of a hard coral reveals cryptic species and ancient genetic lineages. *Mol Ecol*. 2015;24(19):5006-19.
84. Schmidt-Roach S, Johnston E, Fontana S, Jury CP, Forsman Z. Daytime spawning of *Pocillopora* species in Kaneohe Bay, Hawai'i. *Galaxea*. 2014;16:11-2.
85. Tanner JE. Seasonality and lunar periodicity in the reproduction of Pocilloporid corals. *Coral Reefs*. 1996;15(1):59-66.
86. Kojis BL, Quinn NJ. Reproductive Ecology of Two Faviid Corals (Coelenterata: Scleractinia). *Mar Ecol Prog Ser*. 1982;8(3):251-5.
87. Galbraith HS, Vaughn CC. Temperature and food interact to influence gamete development in freshwater mussels. *Hydrobiologia*. 2009;636(1):35-47.

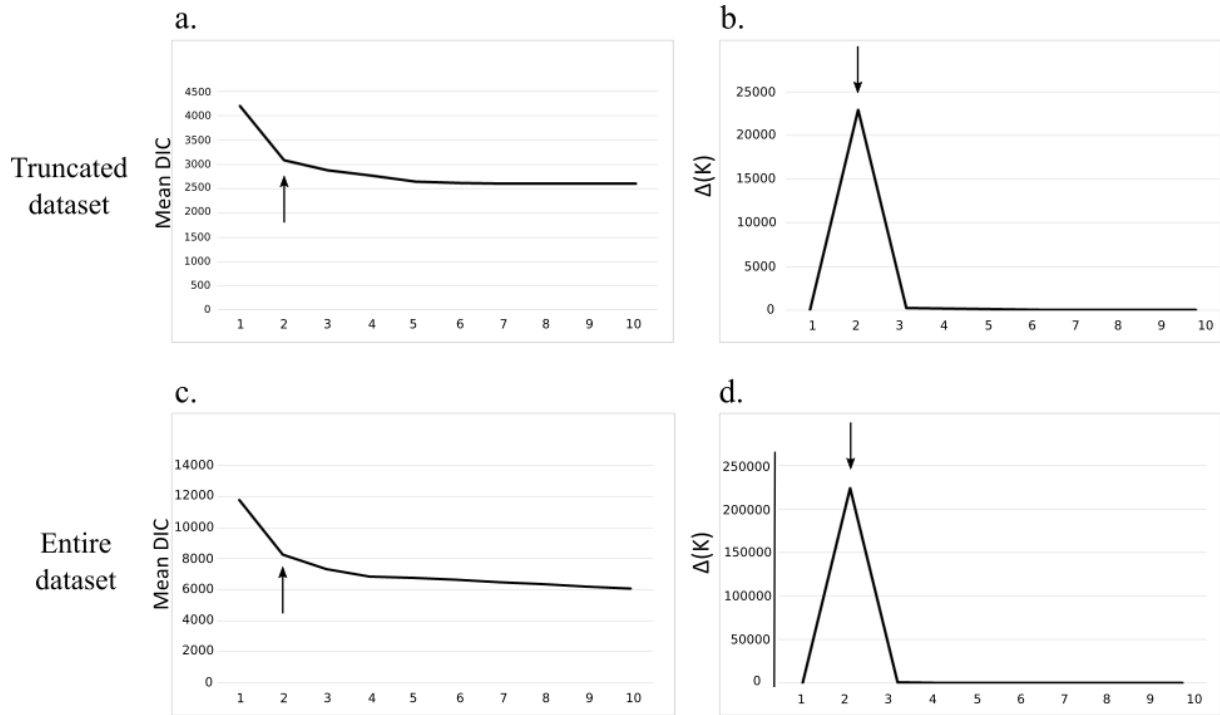
Appendix S1

Panel	Locus Name	\hat{A}	Dye	References
1	Pd3-004	2.46	6-FAM	[41]
1	Pd3-005	4.65	NED	[41]
1	Poc40	6.07	6-FAM	[42]
1	PV2	3.02	VIC	[43]
1	PV7	2.52	VIC	[43]
2	Pd2-001	4.31	VIC	[41]
2	Pd2-006	3.1	NED	[41]
2	Pd3-008	3.65	6-FAM	[41]
2	Pd3-009	5.35	6-FAM	[41]
3	Pd3-EF65	4.21	PET	[28]
3	Pd4	2.54	6-FAM	[44]
3	Pd11	3.85	VIC	[44]
3	Pd13	4.21	NED	[44]

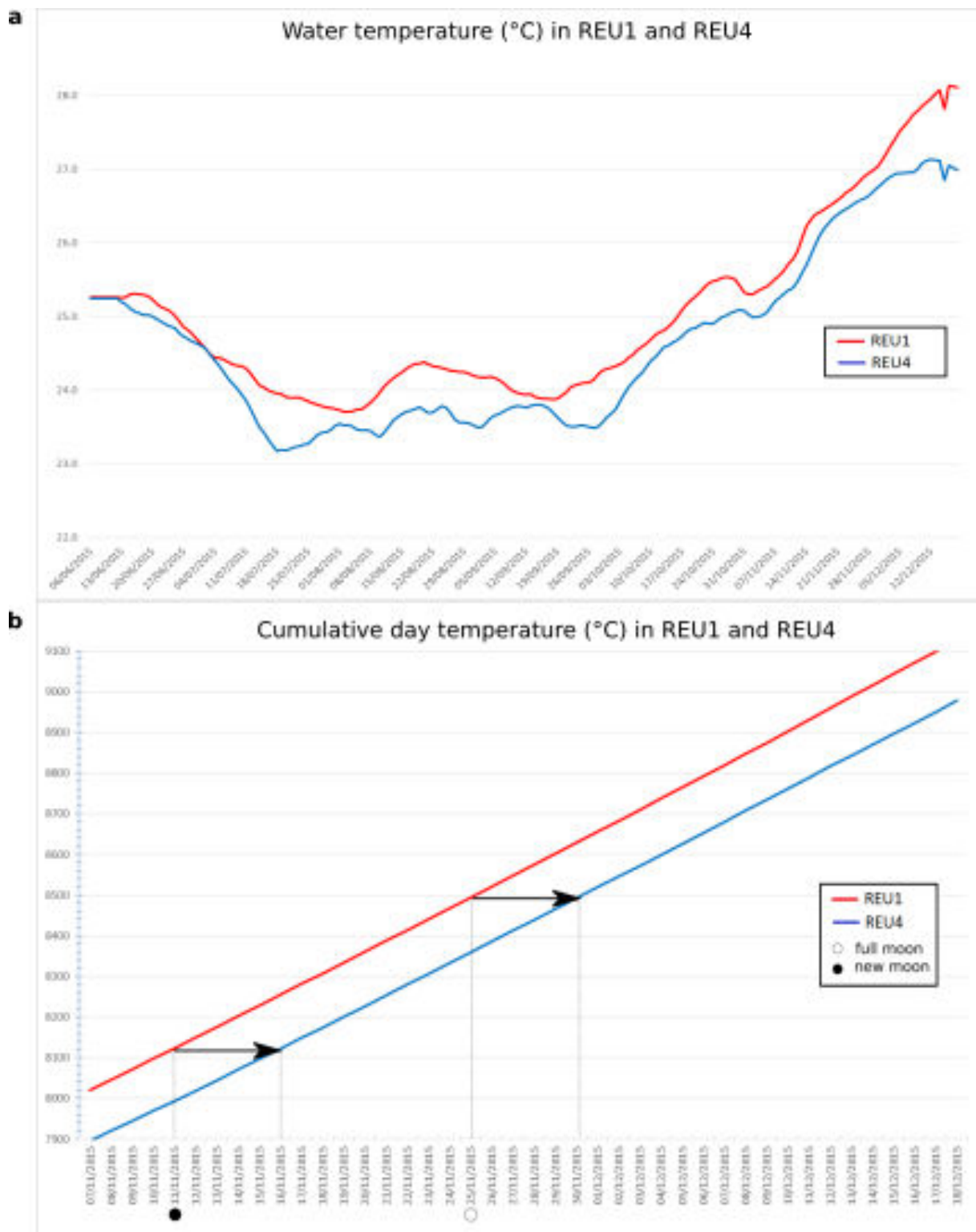
Appendix S2



Appendix S3



Appendix S4



CHAPITRE 3. Etude de la structure génétique et de la connectivité à différentes échelles spatiales chez le corail *Pocillopora damicornis* type β dans le Sud-Ouest de l'océan Indien et le Pacifique Sud-Ouest

Résumé

Ce chapitre s'intéresse à une des lignées génétique du complexe d'espèces *Pocillopora damicornis*, récemment révisé, *P. damicornis* type β . Le but de cette étude est d'estimer la structure génétique et la connectivité de 19 populations de ce corail, grâce à l'utilisation de 13 locus microsatellites. Ainsi, 1162 colonies ont été échantillonnées dans trois écorégions du sud de l'aire de répartition de cette espèce, ces dernières étant peu étudiées : l'Ouest et le Nord de Madagascar, l'archipel des Mascareignes et la Nouvelle-Calédonie. Les résultats révèlent l'existence de propagation clonale dans toutes les populations sur tous les sites étudiés. Cependant, la propagation clonale est restreinte à l'échelle des populations (sauf à La Réunion, où des colonies présentant le même génotype multilocus ont été retrouvées dans des populations séparées par 40 km). Par ailleurs, il existe une différenciation génétique plus importante entre des populations échantillonnées autour d'îles différentes qu'au sein d'une même île, suggérant un patron d'isolement par la distance. Cependant, les résultats des tests d'assignement révèlent que les populations de Nouvelle-Calédonie sont étrangement groupées avec les populations du Canal du Mozambique (particulièrement Mayotte et Juan de Nova). Ceci suggère qu'il pourrait exister des flux de gènes entre ces deux zones. Ces flux seraient unidirectionnels, allant de l'Est vers l'Ouest en lien avec la circulation océanique globale de l'océan Indien.

Par ailleurs, il a été précédemment supposé que *P. damicornis* type β , correspondant à l'hypothèse d'espèce primaire PSH05 (ORF18 et ORF19; Gélina et al. soumis) pourraient se scinder en quatre lignées génétiques (nommées SSH05a, SSH05b, SSH05c et SSH05d). Les résultats de ce chapitre suggèrent que PSH05 pourrait ne représenter qu'une seule hypothèse d'espèce avec un modèle en métapopulation et des flux de gènes limités entre les différentes populations, que l'on peut nommer *P. damicornis* β ou *P. acuta* comme cela a été proposé par Schmidt-Roach et al. (2014).

Population structure and connectivity at different spatial scales for the coral *Pocillopora damicornis* type β in the Southwestern Indian Ocean and the Tropical Southwestern Pacific

Gélin P¹, Pirog A¹, Fauvelot C², Magalon H¹

¹ UMR ENTROPIE (IRD, Université de La Réunion, CNRS), Laboratoire d'excellence-CORAIL; Université de La Réunion, St Denis, La Réunion

² UMR ENTROPIE (IRD, Université de La Réunion, CNRS), Laboratoire d'excellence-CORAIL; centre IRD de Nouméa, New Caledonia

Corresponding author: pauline.gelin87@gmail.com; Université de La Réunion, UMR ENTROPIE Faculté des Sciences et Technologies, 15 Bd René Cassin, CS 92003, 97744 St Denis Cedex 09, La Réunion

Abstract

Seascape genetics aim in understanding how movements of an organism through the seascape impact dispersal and gene flow, *i.e.* connectivity among marine populations. Here, we focused on one lineage of the recently revised *Pocillopora damicornis* complex, *P. damicornis* type β . The aim of our study was to infer the genetic structuring and connectivity among 19 populations for this coral. A total of 1162 colonies were sampled from three ecoregions of the south part of the distribution range of this lineage, which are largely understudied: the Western and Northern Madagascar, the Mascarene Islands (Western Indian Ocean) and New Caledonia (Tropical Southwestern Pacific). Our results indicate the presence of clonal propagation in each studied site. However, clonal propagation was restricted at the population level within each island (except in Reunion Island, where clones were found in several populations 40 km apart). Moreover, populations sampled in different islands appeared more genetically differentiated than populations located on the same island and the genetic differentiation appeared weaker for two estimators (F_{ST} and D_{est}) when estimations were performed on the truncated dataset (*i.e.*, restricting our analysis to colonies of putative sexual origin). Additionally, the result of the assignment tests revealed an odd structuring pattern grouping together populations from New Caledonia with populations of the Mozambique Channel (Mayotte and Juan de Nova Island) revealing potential gene flow that could be unidirectional from East to West considering the global oceanic circulation in the Indian Ocean. Moreover, it had been previously hypothesized that *Pocillopora damicornis* type β (ORF18 and ORF19; G lin et al., submitted) could be separated in four distinct lineages, representing Secondary Species Hypotheses (SSH, *sensu* Pante et al., 2015) and named SSH05a, SSH05b, SSH05c and SSH05d. Our results suggest that PSH05 may present only one putative species with metapopulation pattern, with restricted gene flow, that could be named *P. damicornis* β or *P. acuta* as proposed by Schmidt-Roach et al. (2014).

Keywords: population genetics, microsatellites, assignment tests, *Pocillopora*

Introduction

Seascape genetics, as reflected by the name of the discipline, aim in understanding how movements of an organism through the seascape impact dispersal and gene flow, i.e. connectivity among marine populations. This latter could be supposed to present no restrictions as the open sea should be free from barriers. More precisely, genetic connectivity tracks the dispersal of genes, which only accounts for the individuals that successfully reproduce after dispersing [reviewed in Selkoe et al., 2016)]. Indeed, two populations that are able to exchange genes are considered connected. On the contrary, if these populations do not exchange genes, they should be considered genetically isolated. Commonly, isolation among populations could be the result of geographical barriers or geographic distance. Nevertheless, this isolation might also result from pre- or post-zygotic reproductive isolation. In marine species, post-zygotic barriers are poorly studied because of the difficulty to follow the development of offsprings through complex life cycles and through long generation times (Palumbi, 1994).

For a majority of marine species, dispersal is undertaken during the larval phase, and trajectories of larvae, as well as long-term patterns of dispersal, are difficult to predict (Siegel et al., 2008). Nevertheless, a relationship between the duration of the planktonic phase and the dispersal potential has been consistently found [reviewed in Bohonak, (1999)]. Among marine organisms, scleractinians represent a good example for this relationship since larval abilities to move are limited. Actually, the dispersal of these species over large scales (more than hundreds of kilometers) results from regional oceanic vectors: supposedly the longer the planktonic phase duration, the higher the dispersal abilities. Moreover, the duration of the planktonic phase is difficult to estimate because it is not possible to track the larvae *in vivo*. Nevertheless, some species have been studied *in vitro* and the larvae lifetime was estimated between 30 days for *Heliopora coerulea* (Harii et al., 2002), 100 days for *Pocillopora damicornis* (Richmond, 1987) and more than 200 days for five other species (from genera *Acropora*, *Favia*, *Pectinia*, *Goniastrea*, *Montastraea*, Graham et al., 2008). Then, to approximate dispersal abilities, the use of genetic tools and assessment of population structure have been widely used and different patterns of dispersal have been identified. Among scleractinians, different patterns of population structuring were revealed. Indeed, focusing on nine species, Ayre and Hughes (2000) highlighted restricted dispersal abilities of larvae in the Great Barrier Reef (GBR) and Torda et al. (2013b) suggested that larvae originating from spawned gametes were locally retained, while brooded larvae were mainly vagabonds. Considering the origin of individuals

and the studied species, high levels of genetic differentiation were found in *Acropora cervicornis* populations from the Caribbean separated from two to 500 km (Vollmer and Palumbi, 2007), and, in *Acropora tenuis* populations from the African east coast, no differentiation was found among some populations separated by 1 000 km while some populations 300 km apart were genetically differentiated (van der Ven et al., 2016). However, the structuring pattern in this latter study was complex, revealing that central populations (approximately 460 km south from the northernmost one) appeared more differentiated than the two populations furthest apart (van der Ven et al., 2016). On the opposite, populations of *Acropora austera* studied in the same zone were not differentiated for distances smaller than 1 150 km while over 2 500 km they were (Montoya-Maya et al., 2016). The same result was observed for *Platygyra daedalea*, suggesting a pattern of isolation-by-distance on the African east coast (Montoya-Maya et al., 2016). Another study led in this zone revealed that populations of *Pocillopora verrucosa* from Mozambique were genetically differentiated from the populations of South Africa (Ridgway et al., 2008).

These last few years, the taxonomy of *Pocillopora damicornis* has been revised mainly with the contribution of phylogeny and population genetics, showing that it represents a species complex distinguishing five distinct lineages (α , β , γ , δ and ϵ ; Schmidt-Roach et al., 2013). These new insights lead to reconsider previous studies focusing on *P. damicornis* and the necessity to extract studies in which *P. damicornis* lineages have been clearly identified, either *a posteriori* with cross-checking information or *a priori* by identifying genetically the colonies studied. Thus, our study focused on *P. damicornis* type β , also designated as *P. acuta* (*sensu* Schmidt-Roach et al., 2014), *Pocillopora* type 5 (Pinzón et al., 2013), *P. damicornis* type F (Souter et al., 2009) and *Pocillopora* Clade Ia (Marti-Puig et al., 2014). Without judging these different designations, we chose to use the name *P. damicornis* type β (referred to *P. damicornis* β hereafter for lightening the writing) to refer to the lineage identity herein under study, as it seems to us the more explicit considering the knowledge on species delimitation in this genus to date *Pocillopora damicornis* β (Linnaeus, 1758) is quite common and found in lagoons from the Red Sea (Pinzón et al., 2013), the Indian (Pinzón et al., 2013) (Souter et al., 2009) (Gélin et al. submitted) and Pacific (Adjeroud et al., 2014; Flot et al., 2008; Gorospe and Karl, 2015; Hsu et al., 2014; Marti-Puig et al., 2014; Schmidt-Roach et al., 2013; Schmidt-Roach et al., 2014) oceans in shallow habitats (it has not been clearly reported yet in the Tropical Eastern Pacific). It is a brooder and is able to propagate asexually as clones were identified (Adjeroud et al., 2014; Souter et al., 2009; Torda et al., 2013b; Gélin et al., revised) through fragmentation (Gélin et al., revised; Highsmith, 1982), budding or polyp expulsion (Kvitt et al., 2015; Sammarco,

1982) in *P. damicornis sensu lato*, or release of asexual larvae (Torda et al., 2013b; G elin et al. revised), implying the necessity to take this into account when studying connectivity. Indeed, some multi-locus genotypes (MLG) are repeated, biasing allele frequencies.

In some studies, sampled colonies were clearly attributed to *P. damicornis* type β . Among them, Torda et al. (2013a) focused on population structure of *P. damicornis* type β on the east coast of Australia. Using nine microsatellites, they showed that genetic differentiation was higher between populations from distinct reefs (the two furthest apart sites were separated by more than 1 100 km) than from the same reef. Furthermore, in Reunion Island's reefs, no genetic differentiation was found within one reef while high genetic differentiation was found among reefs along a 40 km coast (Gelin et al., revised). Moreover, (Pinz on et al. (2013) assessed genetic connectivity of *P. damicornis* β , (named type 5 therein) over very large scales using seven microsatellites. From eight populations of the species distribution range (Central Pacific: Hawaii, Lizard Island and Heron Island; Indo-Pacific: Taiwan, Palau and the Gulf of Thailand; Western Indian Ocean: Mauritius and Madagascar), they found high levels of genetic differentiation with the exception of Taiwan and Lizard Island that were connected. Finally, Adjeroud et al. (2014) in French Polynesia found that populations from different islands separated by more than 250 km are differentiated, while populations from the same island separated by 13 km were not.

Here, focusing on one lineage from the *P. damicornis* complex, the aim of our study was to infer the genetic connectivity among 19 populations for the coral *Pocillopora damicornis* β . The studied colonies were sampled from three ecoregions of the south part of the distribution range of this lineage, which are largely understudied: the Western and Northern Madagascar, the Mascarene Islands (Western Indian Ocean) and New Caledonia (Tropical Southwestern Pacific). We identified genetically the colonies under study by sequencing and assignment tests, and then studied the genetic structuring of these populations, following a hierarchical sampling, using 13 microsatellite loci, allowing us to accurately define the distinct MLGs and allele frequencies across different spatial scales.

Material and methods

Sampling design

Pocillopora damicornis-like colonies were sampled (branch tip) in two marine provinces (Spalding et al., 2007): the Western Indian Ocean (WIO) and the Tropical Southwestern Pacific (TSP). In the WIO, colonies were sampled around five islands: Mayotte in May 2016, Juan de Nova Island in December 2013, South of Madagascar in October 2014, Reunion Island in April 2013 and March 2014 and Rodrigues in December 2014. In the TSP, populations were sampled in New Caledonia in November 2013 and February 2014 and in the Chesterfield/Bampton/Bellona in November 2015 (Figure 1, Table 1). The sampling design for REU1 to REU4 was detailed in Gélín et al. (revised). For all the other populations, colonies were sampled haphazardly (by snorkeling or scuba diving). It is noteworthy that two populations were added in Reunion Island to the ones used in Gélín et al. (revised). REU2b was sampled in the same location than REU2 but one year later in March 2014 after the cyclone

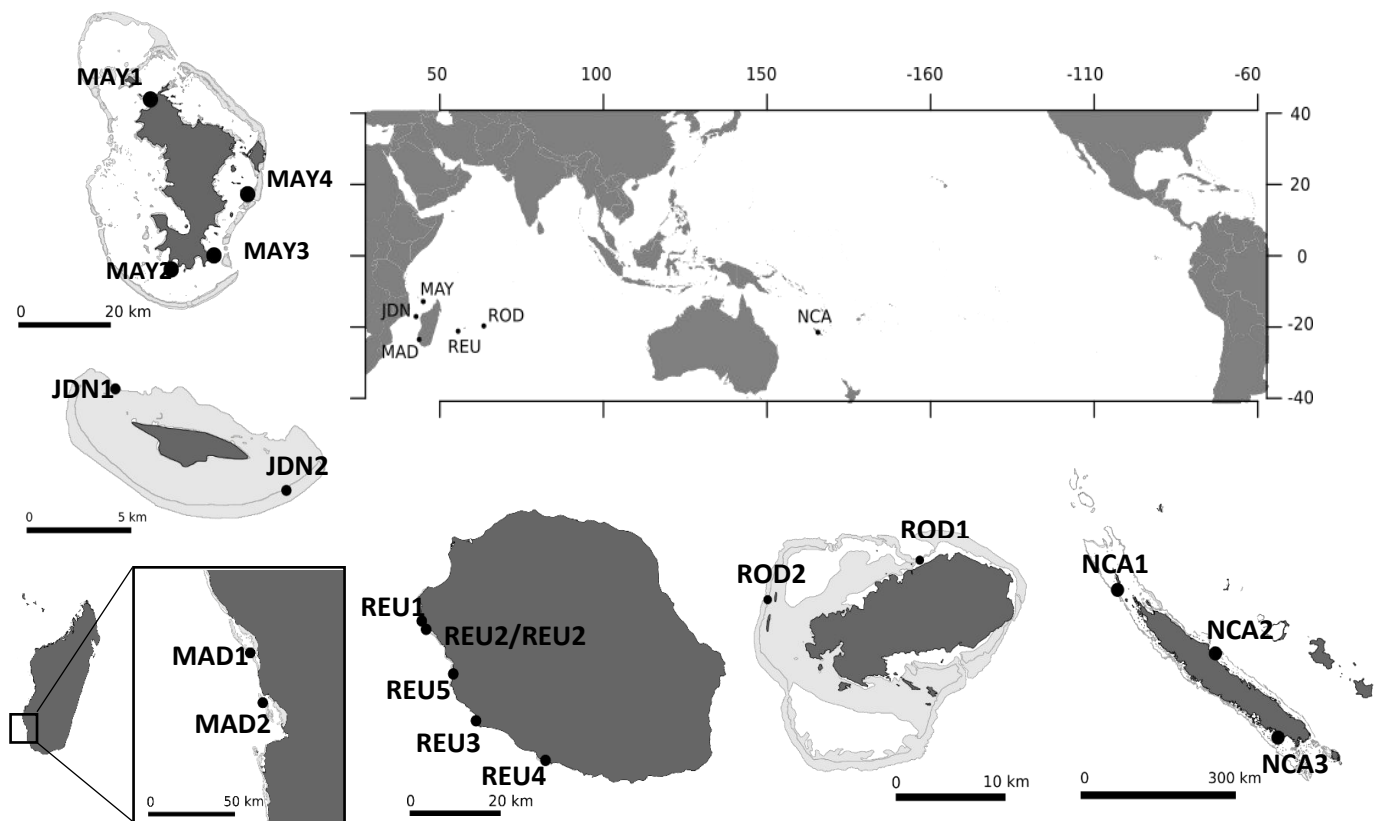


Figure 1. Sampling locations of *Pocillopora* colonies. Populations are numerically identified from the island code (JDN: Juan de Nova Island; MAD: Madagascar (Tulear); MAY: Mayotte; NCA: New Caledonia; REU: Reunion Island; ROD: Rodrigues Island). Constructed using © OpenStreetMap contributors CC BY-SA (<http://www.openstreetmap.org/copyright>) for landmasses and UNEP-WCMC, WorldFish Centre, WRI, TNC (2010). Global distribution of coral reefs, compiled from multiple sources including the Millennium Coral Reef Mapping Project. Version 1.3. Includes contributions from IMaRS-USF and IRD (2005), IMaRS-USF (2005) and Spalding et al. (2001). Cambridge (UK): UNEP World Conservation Monitoring Centre (<http://data.unep-wcmc.org/datasets/1>) for coral reefs.

Bejisa (January 2014) which went through close to the Reunion Island west coast. REU5 was also sampled on the west coast, located between REU2 and REU3 (Figure 1).

DNA extraction, sequencing and microsatellite genotyping

From the sampled colonies, DNA was extracted using DNeasy Blood & Tissue kit (Qiagen™). Colonies were genotyped using 13 microsatellite loci (Table 2). Forward primers were indirectly fluorochrome labelled (6-FAM, VIC, NED, PET) by adding a universal M13 tail at the 5'-end and were multiplexed post-PCR in three panels (see G  lin et al revised). Each amplification reaction was performed as in (Postaire et al., 2015). PCR products were genotyped using an ABI 3730 genetic analyser (Applied Biosystems) and allelic sizes were determined with GeneMapper V4.0 (Applied Biosystems) using an internal size standard (Genescan LIZ-500, Applied Biosystems). Individuals showing peak profiles either faint or presenting more than two peaks were processed again and, when remaining ambiguous, designated as missing data to avoid dealing with more than two alleles, since keeping one or another allele should lead to misestimate alleles frequencies. Because colonies were sampled based on their *corallum* macromorphology, *P. damicornis* β lineage identity was verified before further analyses by sequencing a subset of the colonies for the Open Reading Frame [ORF; see Flot and Tillier (2007)] and using assignment tests combining all multi-locus genotypes (MLG) found in this study and others from G  lin et al. (submitted). Missing data were scattered in the dataset and did not interfere in MLG identification. A test was performed on MLG identification when including or excluding the missing data. In both cases, the same MLGs were identified but the dataset including missing data identified more single MLG which could lead to a slight over estimation of the clonal richness in the populations.

Nevertheless, the total number of missing data represent less than 5% of the whole dataset and were considered negligible. Then, missing data were treated as a new allele and the 1162 colonies were kept for further analyses.

MLG identification and clonal structure

The occurrence of identical MLGs among the sampled colonies was assessed using the R (R Development Core Team 2016) package *RClone* (Bailleul et al., 2016). The probability of obtaining the same MLG more than once from distinct random reproductive events was

Table 1. Summary statistics for the 19 *Pocillopora damicornis* type β populations. N : number of colonies sampled; N_{MLG} : number of Multi-Locus Genotype (MLG); R : clonal richness; $N_{a_{pop}}$: the mean number of allele per population across 13 locus; $F_{IS\ entire}$: values are estimated on the entire dataset; $F_{IS\ truncated}$: values are estimated on the truncated dataset (one representative per MLG and per population). ^{NS}: non-significant; ** $P < 0.01$; *** $P < 0.001$.

Province	Ecoregion	Island	Code	Latitude	Longitude	N	N_{MLG}	R	$N_{a_{pop}}$	$N_{p_{pop}}$	$F_{IS\ entire}$	$F_{IS\ truncated}$	
Western Indian Ocean	Western and Northern Madagascar	Mayotte	MAY1	-12.68064	45.06917	48	42	0.872	5.31 \pm 0.54	0.15 \pm 0.15	0.287***	0.283***	
			MAY2	-13.00121	45.11096	41	40	0.975	5.92 \pm 0.51	0.23 \pm 0.12	0.173***	0.176***	
			MAY3	-12.98269	45.19897	67	62	0.924	5.39 \pm 0.68	0.15 \pm 0.10	0.168***	0.180***	
			MAY4	-12.85977	45.27265	55	45	0.815	5.46 \pm 0.63	0.15 \pm 0.10	0.188***	0.205***	
		Juan de Nova	JDN1	-17.03195	42.69186	29	20	0.679	4.08 \pm 0.47	0.08 \pm 0.08	0.310***	0.323***	
			JDN2	-17.07375	42.76714	48	39	0.809	4.46 \pm 0.58	0.38 \pm 0.24	0.188***	0.197***	
		Madagascar	MAD1	-23.15016	43.56768	24	22	0.913	4.92 \pm 0.53	0.38 \pm 0.14	0.100**	0.111**	
			MAD2	-23.40106	43.63468	7	6	0.833	3.00 \pm 0.25	0.08 \pm 0.08	0.309***	0.337***	
		Mascarene Islands	Réunion	REU1	-21.08236	55.22283	179	62	0.343	4.39 \pm 0.42	0.00 \pm 0.00	-0.079***	0.036 ^{NS}
				REU2	-21.09665	55.23380	313	30	0.093	4.23 \pm 0.52	0.08 \pm 0.08	-0.231***	-0.012 ^{NS}
	REU2b			-21.09665	55.23380	45	16	0.341	1.77 \pm 0.23	0.00 \pm 0.00	-0.701***	-0.233*	
	REU3			-21.27088	55.33231	50	21	0.408	3.85 \pm 0.39	0.00 \pm 0.00	0.090**	0.201***	
	REU4			-21.34574	55.47268	48	29	0.596	4.62 \pm 0.42	0.00 \pm 0.00	0.122***	0.199***	
	REU5			-21.18236	55.28639	14	12	0.846	3.08 \pm 0.54	0.00 \pm 0.00	-0.045 ^{NS}	0.001 ^{NS}	
	Rodrigues		ROD1	-19.66080	63.44196	59	34	0.569	4.00 \pm 0.39	0.08 \pm 0.08	0.046 ^{NS}	0.121**	
			ROD2	-19.70394	63.29929	56	37	0.655	3.77 \pm 0.43	0.08 \pm 0.08	0.051 ^{NS}	0.135***	
	Tropical Southwestern Pacific	New Caledonia	New Caledonia	NCA1	-19.88333	163.65528	32	23	0.710	4.77 \pm 0.53	0.23 \pm 0.12	0.319***	0.297***
				NCA2	-20.92847	165.37446	20	20	1.000	4.54 \pm 0.40	0.38 \pm 0.14	0.210***	0.210***
NCA3				-22.31276	166.47162	27	20	0.731	2.92 \pm 0.46	0.00 \pm 0.00	-0.513***	-0.397***	
			Total			1162	580		4.23 \pm 0.03	0.13 \pm 0.02			

further estimated using P_{SEX} (F_{IS}), which considers possible departure from Hardy-Weinberg equilibrium in order to obtain a less biased estimator (Arnaud-Haond et al., 2007). Colonies sharing the same MLG resulting from a unique reproductive event (one unique zygote) were considered as belonging to the same clone. The clonal richness R was estimated for each population with the formula: $R = (G-1)/(N-1)$ (Dorken and Eckert, 2001). The number of allele (N_a) and the number of private allele (N_p) were estimated for each locus and the mean number of alleles and the mean number of private alleles were estimated per population across 13 loci ($N_{a_{pop}}$ and $N_{p_{pop}}$, respectively). The percentage of missing data was estimated for each locus with the entire dataset (*i.e.*, keeping all colonies) and the truncated dataset (*i.e.*, keeping one representative per MLG per population). The inbreeding coefficient (F_{IS}) was estimated with FSTAT V2.9.3 (Goudet, 2001) for each population with both datasets (*i.e.*, entire and truncated).

Analysis of genetic connectivity

The genotypic linkage disequilibrium was assessed with FSTAT v 2.9.3 (Goudet, 2001) using both datasets.

Population differentiation indices, F_{ST} (Weir and Cockerham, 1984) and D_{est} (Jost, 2008) were estimated with GenAlex V6.5 (Peakall and Smouse, 2006, 2012) both on the entire and on the truncated datasets. Indeed, as some genotypes might present a higher fitness than others, keeping all the colonies sharing these genotypes might bias population differentiation estimation, as well as keeping only one representative per genotype, as described in Alberto et al. (2005).

Then, in the aim to test whether spatial patterns of population structuring could be identified, a Mantel test was performed with the R package *vegan* (Oksanen et al., 2013). This test is based on the comparison of the matrices of genetic distances [linearized F_{ST} ; Slatkin, (1993)] and of geographic distances (\log_{10}) built from latitude and longitude data from population locations.

Additionally, a Bayesian analysis to determine the most likely number of genetically homogenous clusters (K) was performed using STRUCTURE V2.3.4 (Pritchard et al., 2000). Conditions were set to 2×10^6 chain length after a burn-in of 5×10^5 assuming admixture. Simulations were run between $K = 1$ and $K = 20$ with 10 replicates for each K value. The optimal K was detected with Structure Harvester (Earl, 2012). Because STRUCTURE has strong hypotheses [Hardy-Weinberg equilibrium (HWE) of populations and no linkage disequilibrium among loci], two other methods were used. First, INSTRUCT (Gao et al., 2007),

which considers partial self-fertilization or inbreeding, was run to compare the results from the two Markov Chain Monte Carlo (MCMC) methods. For the latter, we ran five chains for each K from 2 to 10. Second, a Discriminant Analysis of Principal Components [DAPC; Jombart et al. (2010)] was performed in order to test whether the structuration observed with STRUCTURE and INSTRUCT can be found with DAPC, an analysis that does not make any assumption about HWE or linkage equilibrium and transforms genotypes using PCA as a prior step to a discriminant analysis. DAPC was applied using the *adegenet* package (Jombart, 2008) for R (R Development Core Team 2016).

Finally, network analyses were performed centered on individuals and populations. The global pattern of genetic relationship among individuals was illustrated by networks built with two different measures. These measures integrate genetic information in terms of time and divergence history: the Rozenfeld Distance index (RD) and the Shared Allele Distance index (SAD). RD has been developed from Goldstein distance index for use between individuals. It provides a parsimonious (*i.e.*, minimal) representation of the genetic distance based on the difference of the microsatellites allele lengths between individuals (Rozenfeld et al., 2007). On the other hand, SAD provides the genetic distance between individuals based on the proportion of shared alleles (Chakraborty and Jin, 1993). These genetic distances help to resolve the relationship between individuals at different time-scale: RD helps to resolve ancestral polymorphism through allele lengths impinged on slow evolutionary processes, while SAD helps to understand recent gene flow characterized by direct allelic exchange.

The global pattern of genetic relationship among populations was illustrated by networks built with two different measures: the Goldstein distance index (GD) and F_{ST} fixation index (F_{ST}). GD groups populations considering their historical origin while F_{ST} takes into account the population structure. Once the matrices of genetic distances between individuals or populations were estimated, different networks were built considering individuals/populations and genetic distances as nodes and links between them, respectively. For the network construction, links were included for all distances and were removed in decreasing order until the percolation threshold (D_{pe}) was reached (Rozenfeld et al., 2007), threshold below which the network fragmented into small clusters. The average clustering coefficient $\langle C \rangle$ of the whole network was estimated for each of the four built networks. These analyses were performed using EDENetwork software (Kivelä et al., 2015).

Results

Sampling

A total of 1162 colonies grouped together with colonies assigned in SSH05a to SSH05d from G elin et al. (submitted) and corresponding to *P. damicornis* β . They were extracted to constitute our dataset. This latter is composed, for the WIO, of four populations in Mayotte (MAY), two populations in Juan de Nova Island (JDN), two populations in South Madagascar (MAD), six populations in Reunion Island (REU) and two populations in Rodrigues (ROD), and for the TSP, of three populations in New Caledonia (NCA) (Figure 1, Table 1). It is surprising that no populations of *P. damicornis* β (except one colony) were found in Chesterfield/Bellona/Bampton Plateau (located mid-way from New Caledonia and the Great Barrier Reef, Australia) while an oceanographic campaign of 14 days and 35 stations explored by scuba diving or snorkeling.

MLG identification and clonal richness

The number of alleles per locus varied from five for Pd3-004 to 15 for Pd4 and Pd11 and the number of private alleles varied from zero for Pd3-004 to seven for PV2 and Pd4 (Table 2). All loci were considered polymorphic, but some loci were found monomorphic in some populations: PV7 in REU2b and NCA2, Poc40 in REU2b and ROD2, Pd3-EF65 in REU5 and Pd2-001 and Pd2-006 in REU2b. The mean number of alleles per population $N_{a_{pop}}$ (mean \pm s.e.) varied from 1.77 ± 0.23 for REU2b to 5.92 ± 0.51 for MAY2 and the mean number of private alleles per population $N_{p_{pop}}$ (mean \pm s.e.) varied from 0.00 ± 0.00 in REU1, REU2b, REU3, REU4, REU5 and NCA3 to 0.38 ± 0.14 and 0.38 ± 0.24 in NCA2 and JDN2, respectively (Table 1).

Over the 1162 colonies of *P. damicornis* β , 590 distinct MLGs were detected. For each MLG, $P_{SEX}(FIS)$ was very low ($< 2.21 \times 10^{-25}$), indicating that two colonies presenting the same MLG were the result of one unique sexual reproductive event and could be considered as members of the same clone. Colonies presenting a same MLG could be found in more than one population only in Reunion Island (see G elin et al., revised. Elsewhere, clones were not shared between populations).

Overall colonies, genotypic linkage disequilibrium (LD) was revealed with the entire dataset for all pairwise tests (100%) and for 15 tests over 78 (19%) with the truncated one.

Nevertheless, when considering LD for each population separately, the pairs of linked loci were not always the same, for both datasets. Thus, we chose to keep all loci for further analyses.

Table 2. List of the loci used in this study. %_{NA entire} and %_{NA truncated} are the percentage of missing data per locus overall populations estimated on the entire and the truncated dataset, respectively; *Na*, the number of alleles per locus and *Np*, the number of private allele per locus.

Mix	Locus	% _{NA entire}	% _{NA truncated}	<i>Na</i>	<i>Np</i>	References
1	Pd3-004	1.807	2.796	5	0	Starger et al. (2007)
1	Pd3-005	1.893	3.454	13	3	Starger et al. (2007)
1	Poc40	15.749	23.684	12	1	Pinzón and Lajeunesse (2011)
1	PV2	4.518	8.306	14	7	Magalon et al. (2004)
1	PV7	1.205	2.303	6	1	Magalon et al. (2004)
2	Pd2-001	2.108	3.701	8	1	Starger et al. (2007)
2	Pd2-006	1.721	3.289	9	1	Starger et al. (2007)
2	Pd3-008	1.033	1.809	8	2	Starger et al. (2007)
2	Pd3-009	3.27	5.592	13	2	Starger et al. (2007)
3	Pd11	7.874	13.405	15	1	Torda et al (2013)
3	Pd13	4.561	7.237	14	3	Torda et al (2013)
3	Pd3-EF65	6.713	9.868	13	3	Gorospe and Karl (2013)
3	Pd4	8.305	12.582	15	7	Torda et al (2013)
Total				11.15 ± 0.97	2.46 ± 0.62	

The clonal richness (*R*) varied across populations. The lowest value was found in REU2 (*R* = 0.093; Table 1), indicating the presence of numerous colonies sharing the same MLG due to its particular sampling scheme (Gélin et al. revised). Conversely, the highest value was found in NCA2 (*R* = 1.000; Table 1) where all the sampled colonies showed distinct MLGs.

F_{IS} values estimated on the entire dataset, varying from -0.701 (*P* < 0.001) for REU2b to 0.319 (*P* < 0.001) for NCA1, were significantly negative for four populations (REU1, REU2, REU2b and NCA3; Table 1), positive for 12 populations and not significantly different from 0 for the three others (Table 1). *F_{IS}* estimations performed on the truncated dataset revealed one value significantly negative for NCA3 (Table 1), two were not significantly different from 0 and the 16 other values were significantly positive (Table 1).

Analysis of population connectivity

Analyses were performed both on the entire and the truncated datasets as suggested in (Alberto et al., 2005). The truncated dataset circumvents the problem of clonality in samples; however, clonal colonies were represented by separate colonies which could take part to sexual reproduction and then be a component of the observed genetic structure of the populations.

Genetic differentiation

Considering the entire dataset, all populations were highly differentiated from each other (except one pair, with F_{ST} estimations varying from 0.006^{NS} for REU2b/REU3 pair to 0.393*** for REU3/NCA3 pair. The result was almost the same when F_{ST} were estimated on the truncated dataset (two pairs of populations revealed non-significant tests), varying from 0.018^{NS} between REU1 and REU2b to 0.283*** between REU3 and NCA3. For the D_{est} , lowest and highest values were observed for the same pairs as for the F_{ST} (Table S1). The non-significant tests found for both F_{ST} and D_{est} were all observed between populations from Reunion Island. Overall, F_{ST} values lower than 0.100 were observed between populations sampled around the same island, except between NCA2 and populations from MAY and JDN (Table S1).

The Mantel test was significant considering both the entire ($r = 0.39$ ***) and the truncated ($r = 0.50$ ***) datasets.

Clustering analysis

Considering the entire dataset, the Evanno's method revealed an optimal K of two or three. Moreover, there were no variations in the likelihood mean over the 10 runs for $K = 3$, indicating this one as "stable". For $K = 2$, the dataset was split into two groups grouping together REU1, REU2, REU2b and a part of colonies from REU3 in the first cluster, and all the other populations in the second cluster, meaning that colonies from MAY, JDN, MAD, ROD and NCA were grouped together (Figure S1). For $K = 3$, we expected that NCA would appear as a separate group, but examining the clustering plots for different K s (Figure S1a) revealed that populations from NCA were never assigned in a separate group and remained grouped with some populations from the Western and Northern Madagascar. Moreover, the DAPC analysis led to nearly identical results for $K = 2$ and $K = 3$ (Figure S2a). While the BIC distribution was fuzzy, a slow decrease in the curve could be observed for $K = 4$ (Figure S2a). The clustering result indicated a split of the first group (REU1, REU2, REU2b and a part of colonies from

REU3) into two new clusters and NCA remained grouped with colonies from the Western and Northern Madagascar (Figure 2a) and never appeared as a separated group among the different values of K . We hypothesized that this odd result might be due to the clonality particularly high in Reunion Island (see G elin et al., revised), and we supposed that we would not observe the same result with analyses performed on the truncated dataset. For the truncated dataset, Evanno’s method revealed an optimal K for $K = 3$ with no variations in the likelihood mean

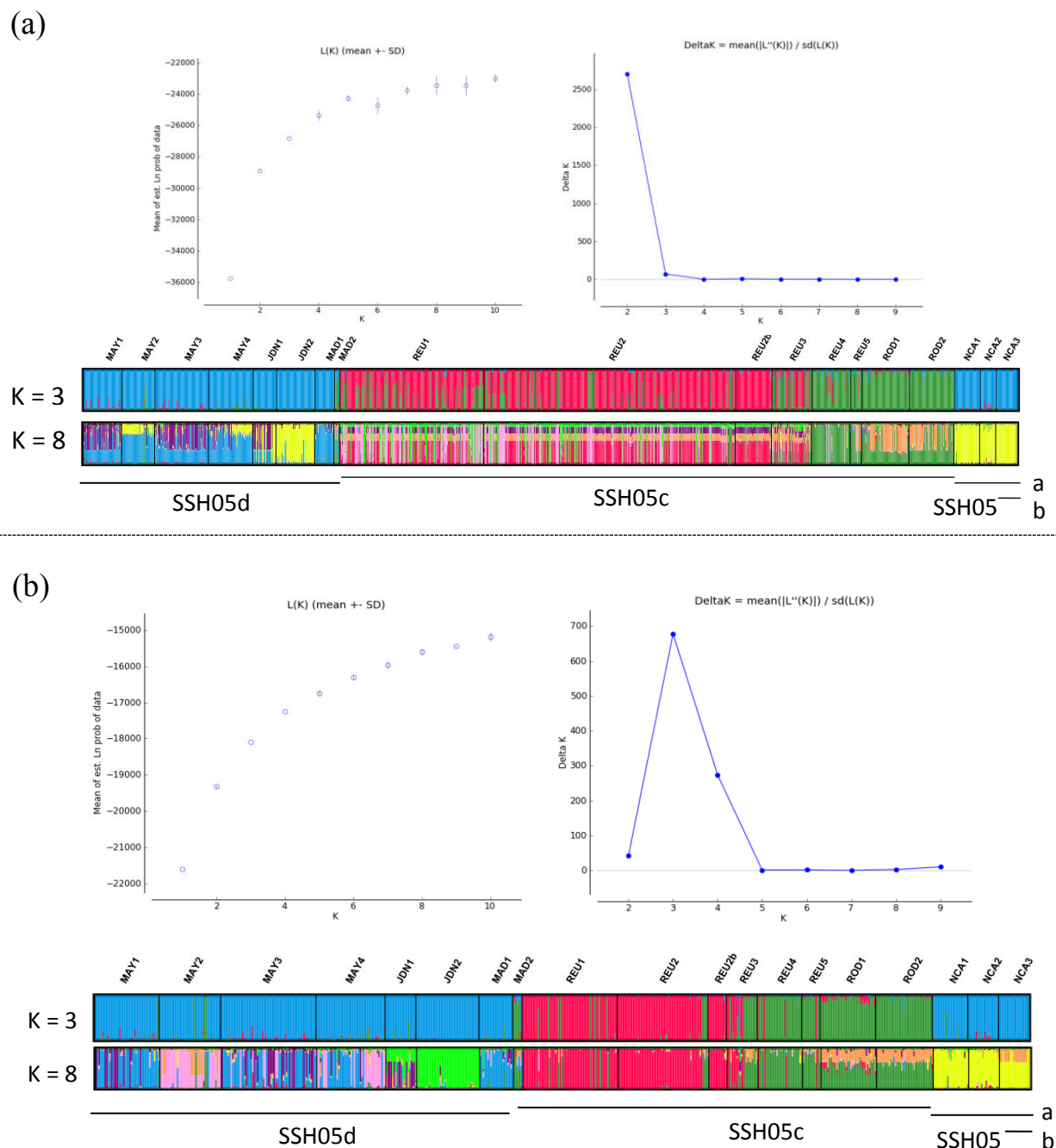


Figure 2. Results of the Bayesian clustering analysis performed on the entire dataset (a) and the truncated one (b). For each dataset are presented the mean likelihood distribution over 10 runs (upper left), the $\Delta(K)$ (upper right) and the result for $K=3$ and $K=8$.

over the 10 runs for this K . For this K , the same clustering pattern as the one observed with the entire dataset appeared (Figure 2a): colonies from NCA remained grouped with some colonies from Western and Northern Madagascar. However, NCA appeared as a clear separate group from $K = 8$. For this value, as said before, colonies from NCA were assigned in one genetic cluster, colonies from MAD2, REU and ROD were mixed in two clusters and colonies from MAY, JDN and MAD1 were mixed in four clusters (Figure 2b). Concerning the DAPC analysis, the BIC distribution did not reveal clear optimal K , but we could examine what happened for $K = 3$ (Figure S4a and S4b), as for the STRUCTURE result. The results were nearly identical. Moreover, once again, NCA did not appear as a separate group until $K = 8$ was reached. Moreover, considering the two populations from REU which were added to the four previously analysed in G elin et al. (revised), REU2b was grouped with REU1 and REU2 suggesting that no change in the genetic structure of the REU2 population was observable four months after the cyclone Bejisa hit Reunion Island. Moreover, colonies from REU5 were assigned in the southern cluster found in this latter study.

Trying to untangle these results, from the results of $K = 2$ obtained with the entire dataset, we performed new assignment tests. Then, 520 colonies corresponding to the first cluster (the green one) were extracted from the original 1162 colonies. This selection corresponded to a dataset where REU1, REU2, REU2b and a part of REU3 were removed (*i.e.* the red cluster). The results gave the same pattern as what was found with the entire dataset, except for $K = 2$ and $K = 3$ where individuals from NCA remained grouped with colonies from REU and ROD and with MAY and JDN, respectively (Figure S5).

Network analysis

The topology of the network built with the SAD index at the percolation threshold ($D_{pe} = 0.56$) showed that individuals from populations of one island remained linked and should be more closely related than individuals from populations of other islands (Figure 3a). The network built with the RD index ($D_{pe} = 1.33$) resulted in less linked MLGs (except those from Reunion Island) according to the geography suggesting that colonies from one island are not especially closer than colonies from other islands (Figure 3a). The average clustering coefficient was higher for the network built with the SAD index ($\langle C \rangle = 0.60$) than for the network built with the RD index ($\langle C \rangle = 0.47$), suggesting that the network was more structured using the SAD index. The topology of the network based on F_{ST} index ($\langle C \rangle = 0.65$, $P = 0.34$) at the percolation threshold ($D_{pe} = 0.23$) revealed strong relationships between

populations from the same ecoregion (Figure 3b). We could distinguish two groups, one composed of populations from REU and ROD and one composed of the other populations. Populations from NCA occupied a central position between the two groups. Moreover, the network built with the GD index ($\langle C \rangle = 0.39$, $P = 0.90$) at the percolation threshold ($D_{pe} = 14.11$) did not indicate clear groups (Figure 3b). Indeed, populations from one island are more scattered on this network compared to the network obtained with the F_{ST} index. The clustering coefficients suggest that the network built with the F_{ST} index is more structured than the network built with the GD index.

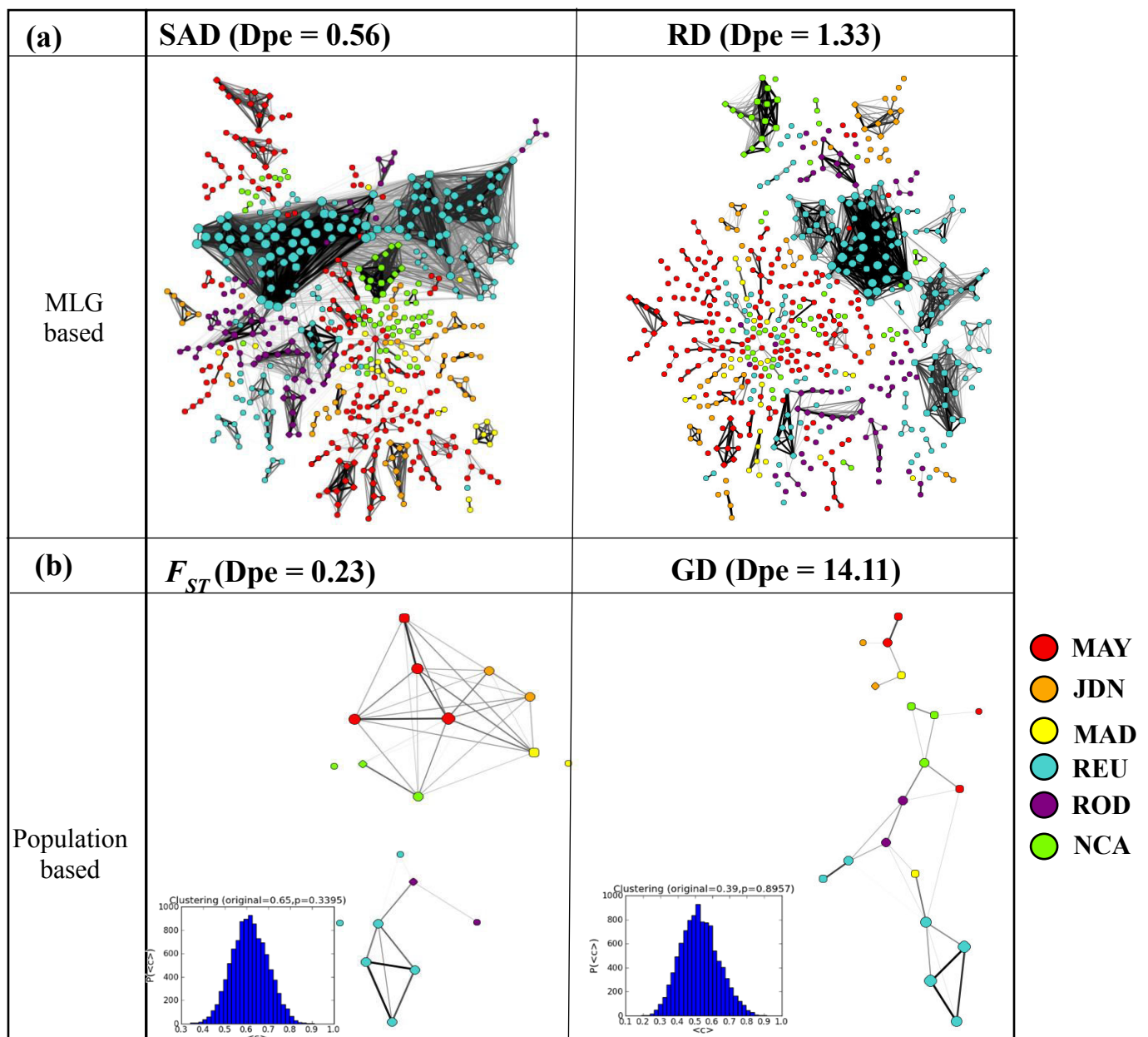


Figure 3. Network topology for *Pocillopora damicornis* β based on several genetic distances. (a) MLG-based network using the Shared Allele Distance (SAD) and the Rozenfeld distance (RD) and (b) populations-based networks based on the F_{ST} index (F_{ST}) and the Goldstein index (GD).

Discussion

Our results indicate the presence of clonal propagation in each studied sites. However, colonies from the same clone were always restricted at the population level within each island (except in Reunion Island, where clones were found in several populations 40 km apart). Moreover, populations sampled in different islands appeared more genetically differentiated than populations located on the same island. Additionally, the result of the assignment tests revealed an odd structuring pattern grouping together populations from New Caledonia with populations of the Mozambique Channel (Mayotte and Juan de Nova Island).

Clonal propagation, a common way to colonize?

Overall our sampling (Western Indian Ocean and New Caledonia) focusing on *P. damicornis* β , clones were found in all sampled sites (excepted in NCA2, New Caledonia). Indeed, even if we did not set up the more appropriate sampling scheme (it was, actually, not the purpose here) to study clonal propagation in the sampled sites, we were able to identify colonies that share the same MLG. Furthermore, clonal propagation among different sites has already been described for this species, in different localities such as the GBR in Australia (Torda et al., 2013a), French Polynesia (Adjeroud et al., 2014) or the east coast of Africa (Souter et al., 2009). Thus clonal propagation seems to be inherent to *P. damicornis* β . Additionally, clones were extended at the scale of the sampling site (< 100 m), except in Reunion Island where colonies sharing the same MLG were retrieved in different sampling sites 40 km apart (detailed in G lin et al., revised), suggesting a production of parthenogenetic larvae able to disperse clones over large distances. Nevertheless, clonal propagation was rarely documented over distances higher than 40 km (see G lin et al., revised; Adjeroud et al., 2014; Souter et al., 2009).

Our results along with previous studies on *P. damicornis* β suggested that clonality appeared as an important characteristic for this species. The matter with clonality is to find the way to treat clones for assessing population genetic structure and connectivity between populations. Indeed, removing repeated MLGs leads to estimate less biased estimations of “conventional” estimators which are not designed for clonal populations [*e.g.*, Halkett et al. (2005)]. But, if clonal propagation is inherent to that species, it should not be ignored, as colonies presenting the same MLG may participate to sexual reproduction and gene flow. Then, in order to faithfully approximate the population structure of *P. damicornis* β , it is crucial to analyze in parallel both

datasets (*i.e.*, the entire dataset where clonality is taken into account and the one with one representative per MLG which accounts for sexual reproduction only). Indeed, the “true” values may be in between both extremes.

Genetic differentiation and connectivity in a clonal species in the Western Indian Ocean and New Caledonia

Our results indicated that genetic differentiation appeared weaker for both estimators (F_{ST} and D_{est}) when estimations were performed on the truncated dataset (*i.e.*, restricting our analysis to colonies of putative sexual origin). Similar results were also found in French Polynesia (Adjeroud et al., 2014): considering their entire dataset, the populations located on distinct islands were genetically differentiated, while, considering their truncated dataset, the populations from distinct islands became genetically undifferentiated. It is noteworthy, as they identified lots of clones in their sampling, that the truncated dataset resulted in a low number of MLGs for the different locations, that might “not offer sufficient statistical power to detect a moderate departure from panmixia”, as the authors recognized. As for us, only one population in Madagascar (MAD2) presented a small number of colonies and MLGs and was not considered in the estimations of differentiation indices.

Concerning connectivity distances, whatever the dataset used, high levels of genetic differentiation were found among all our studied populations. However, the lowest F_{ST} and D_{est} values were obtained for pairs of populations separated by less than 400 km suggesting that gene flow is restricted over high distances (> 400 km). Results of the Mantel test indicated that the more distant, the more genetically differentiated populations were. The same pattern was observed in populations of *P. damicornis* β in East Australia: the lowest values were found within regions, *i.e.* among populations from the Northern part of the GBR and among those from the Central GBR (Torda et al., 2013a) and high values between populations from both regions (ca. 500 km). Likewise, high genetic differentiation was found between populations covering the Pacific and the Indian oceans [except between Taiwan and Lizard Island; Pinzón et al. (2013)]. Nevertheless, these results are to be taken carefully as the number of colonies in some populations was very small (four populations with at most 12 colonies among eight populations studied).

This discrepancy between the structuring pattern observed when conserving all colonies assuming that colonies from the same clone participate equally to gene flow and the one observed when giving the same weight to each genotype (even if one MLG shows a frequency

of 99% in the population) might be the reflection of better dispersal abilities for larvae sexually produced. Nevertheless, the debate about larvae dispersal abilities in corals is open. Indeed, Ayre and Miller (2004) found that asexual larvae of *P. damicornis sensu lato* tend to settle locally (near the broodparents). Indeed, previous models of dispersal and population maintenance in species with mixed modes of reproduction (Williams, 1975) postulate that sexually derived larvae provide long-distance colonists and gene flow between distant populations, whereas asexual reproduction plays a primary role in the maintenance of local populations (Adjeroud and Tsuchiya, 1999; Ayre and Willis, 1988). Conversely, it has been suggested that sexually produced larvae of *P. damicornis* β were retained locally while asexually produced ones were mainly vagabonds (Torda et al., 2013b). Furthermore, Isomura and Nishihira (2001) highlighted the size heterogeneity of planulae among colonies in *P. damicornis sensu lato* and the fact that the bigger the colony, the longer the planula lifetime, and the higher the dispersal ability. Actually, it might be possible that larvae, whatever their mode of production, are long-colonists and insure connectivity between distant populations while maintenance of local populations would be mainly insured by clonal propagation via fragmentation (Highsmith, 1982).

Another interesting result of the Bayesian clustering and DAPC analyses is the odd grouping of populations from NCA with populations from the Mozambique Channel (till $K = 8$) and the central position of NCA between REU/ROD and MAY/JDN/MAD in the network (even if the network topology did not appear robust). In a biological way, the gene flow, even if limited, is unidirectional from East to West considering the global oceanic circulation in the Indian Ocean [e.g., Schott et al. (2009)] explaining the central position of New Caledonia in the network. The West Pacific could feed in alleles (via larvae and recruits on floating objects) the Indian Ocean, at least its Southwestern part. Indeed, individuals (whichever their growth state) might have reach the Mozambique Channel step by step, colonizing reefs located in between New Caledonia to Mozambique Channel.

The fact that populations from Western Indian Ocean showed two groups, Madagascar continent acting like a barrier, may be more probably linked to the sampling disequilibrium due to the sampling design set up in Reunion Island (for searching clones). Indeed, this disequilibrium resulted in a greater number of samples for REU, and consequently links between MLGs from Reunion Island and populations (revealed by the network whatever it was MLG- or population-based) were strong. Nevertheless, when tests were performed on each cluster (from $K = 2$) for both datasets, NCA remained grouped with populations from MAY. Furthermore, the assignment tests grouped REU5 in the southernmost cluster of Reunion Island

(Gélin et al., revised) for both the entire and the truncated datasets, while it is located between REU2 and REU3. According to the hypothesis proposed in this latter study, we expected that it would be composed of a mix of the northern and the southern clusters, as in REU3. Nevertheless, one individual of REU5 was assigned in the northernmost cluster. This discrepancy could be due to the few number of colonies sampled in this population ($n = 14$ and $n = 12$), for the entire and the truncated datasets, respectively) that did not allow to sample enough representatives for each of the clusters.

Alternatively, the observed structuring pattern in Reunion Island could be due to a random sampling of the genotypes during colonization processes such as Initial Seedling Recruitment (ISR). Actually, it had been shown that chance effects during range expansions dramatically alter the gene pool in microbial populations (Excoffier and Ray, 2008; Hallatschek et al., 2007). Indeed, well mixed populations of two fluorescently labeled strains of neutral markers in *Escherichia coli* develop well defined sector-like regions with fractal boundaries in expanding colonies and resulting in a complete segregation of two neutral markers.

Moreover, even if it was unclear, the results of the network could indicate a geographical isolation among historically related populations (Barrandeguy and García, 2015). Then, the observed structure might be the result of random genetic drift that would have stabilized alleles with similar frequencies in NCA and in populations from the Mozambique Channel. Moreover, it had been shown that when migration is limited, the linkage disequilibrium between pairs of loci becomes large (Ohta, 1982a, b). As we found strong genotypic linkage disequilibrium among our loci, this might represent an additional proof of geographical isolation of the populations of *P. damicornis* β .

Species delimitation

In a previous study (Gélin et al. submitted), it had been hypothesized that *Pocillopora damicornis* β (ORF18 and ORF19 therein) could be separated in four distinct lineages, representing Secondary Species Hypotheses (SSH, *sensu* Pante et al., 2015) and named SSH05a, SSH05b, SSH05c and SSH05d. SSH05a and SSH05b were found exclusively in New Caledonia and in sympatry in one site (NCA3 herein and Nouméa therein) while SSH05c and SSH05d were found in the Western Indian Ocean with SSH05d exclusively in the Mozambique Channel and SSH05c almost exclusively in the Mascarene Islands (only one colony in sympatry with SSH05d in the South West of Madagascar). Colonies used in this study were extracted from a clustering analysis combining all the *Pocillopora* colonies we have sampled and all

sequenced colonies used in G elin et al. (submitted). Then, all the colonies assigned in the same clusters than colonies assigned to SSH05a to SSH05d were extracted and analyses were performed on this dataset. It is noteworthy that the result of the Bayesian clustering analysis did not allow to retrieve the previously four identified species hypotheses at small K (they did appear clearly at $K = 8$ when New Caledonia populations were assigned to their own cluster). This could be explained by the clonality harbored by *P. damicornis* β . Indeed, in G elin et al. (submitted), the assignment tests were performed using colonies from the whole genus without taking into account clonality and with intuitively very distant MLGs (all the genus diversity explored). In the opposite, in the present study, were included only colonies attributed to the same Primary Species Hypothesis [PSH05 in G elin et al. (submitted)] representing only a portion of the whole genus diversity and belonging to clones, themselves clustered in clonal lineages (*i.e.*, all the colonies harboring MLGs differing by few alleles due to somatic mutations or scoring errors), leading to the over-representation of close MLGs. This may have introduced a bias in the allele frequencies and departure from the hypotheses required by the Bayesian tests, leading to different clustering patterns. Thus, as (1) the number of individuals found in sympatry for SSH05a and SSH05b on one hand and for SSH05c and SSH05d on the other hand remained low and (2) the clonality of this species leads to a highly differentiated populations at local scale, the observed pattern is more likely to represent a geographical structuring pattern rather than four distinct putative species. Thus, PSH05 may present only one putative species with metapopulation pattern, with restricted gene flow, that could be named *P. damicornis* β or *P. acuta* as proposed by Schmidt-Roach et al. (2014).

Conclusion

This study offers an overview of *Pocillopora damicornis* β population structure in the south part of its distribution area. Since the recent revision of the *Pocillopora damicornis* species complex, this study is one of the few studies that have focused on one of its lineages, and in under-studied regions. Results indicated that clonality for this species appeared very common in all the localities where it had been studied and gene flow appeared restricted beyond hundreds of kilometres. Then, these two aspects are fundamental for properly setting up managing strategies for this species.

Acknowledgments

Corals sampling in New Caledonia (HM) was carried out during COBELO (doi: <http://dx.doi.org/10.17600/14003700>), BIBELOT (doi: <http://dx.doi.org/10.17600/13100100>) and CHEST (doi: <http://dx.doi.org/10.17600/15004500>) oceanographic campaigns on board of RV Alis (IRD). Sampling in Reunion Island (HM, PG) was supported by program CONPOCINPA (LabEx CORAIL fund); in Madagascar and (HM) supported by project Biodiversity (POCT FEDER fund); in Juan de Nova Island (HM) by program BIORECIE (financial supports from INEE, INSU, IRD, AAMP, FRB, TAAF, and the foundation Veolia Environnement), in Mayotte (HM) by program SIREME (FED); in Rodrigues Island (HM) supported by project Biodiversity (POCT FEDER fund). We acknowledge the Plateforme Gentyane of the Institut National de la Recherche Agronomique (INRA, Clermont-Ferrand, France) for guidance and technical support. The first author was financially supported by a PhD contract from the LabEx CORAIL.

Supplementary material

Figure S1: Results of the Bayesian assignments. Analyses were performed on (a) the entire dataset and (b) the truncated dataset from $K = 2$ to $K = 10$.

Figure S2: DAPC analysis performed with the entire dataset. (a) BIC distribution and (b) plots are presented from $K = 2$ to $K = 10$.

Figure S3: DAPC analysis performed with the truncated dataset (a) BIC distribution and (b) plots are presented from $K = 2$ to $K = 10$.

Table S1: Genetic differentiation between *Pocillopora damicornis* type β populations estimated with Weir and Cockerham's F_{ST} (Weir & Cockerham 1984) and with Jost's D_{est} (in parentheses, Jost, 2008). Values in the lower-left matrix are estimated on the entire data set; values in the upper-right matrix are estimated on the truncated data set. Values in bold are not significant ($P > 0.05$). Grayed out values were estimated between population separated by less than 400 km

Reference list

- Adjeroud, M., Guérécheau, A., Vidal-Dupiol, J., Flot, J.-F., Arnaud-Haond, S., Bonhomme, F., 2014. Genetic diversity, clonality and connectivity in the scleractinian coral *Pocillopora damicornis*: a multi-scale analysis in an insular, fragmented reef system. *Mar Biol* 161, 531-541.
- Adjeroud, M., Tsuchiya, M., 1999. Genetic variation and clonal structure in the scleractinian coral *Pocillopora damicornis* in the Ryukyu Archipelago, southern Japan. *Marine Biology* 134, 753-760.
- Alberto, F., Gouveia, L., Arnaud-Haond, S., Perez-Llorens, J.L., Duarte, C.M., Serrao, E.A., 2005. Within-population spatial genetic structure, neighbourhood size and clonal subrange in the seagrass *Cymodocea nodosa*. *Mol Ecol* 14, 2669-2681.

- Arnaud-Haond, S., Duarte, C.M., Alberto, F., Serrao, E.A., 2007. Standardizing methods to address clonality in population studies. *Mol Ecol* 16, 5115-5139.
- Ayre, D.J., Hughes, T.P., 2000. Genotypic diversity and gene flow in the brooding and spawning corals along the Great Barrier Reef, Australia. *Evolution* 54, 1590-1605.
- Ayre, D.J., Miller, K.J., 2004. Where do clonal coral larvae go? Adult genotypic diversity with reproductive effort in the brooding coral *Pocillopora damicornis*. *Mar Ecol Progr Ser* 277, 95-105.
- Ayre, D.J., Willis, B., 1988. Population structure in the coral *Pavona cactus*: clonal genotypes show little phenotypic plasticity. *Mar Biol* 99, 495-505.
- Bailleul, D., Stoeckel, S., Arnaud-Haond, S., 2016. RClone: a package to identify MultiLocus Clonal Lineages and handle clonal datasets in R. *Methods Ecol Evol*.
- Barrandeguy, M.E., García, M.V., 2015. Back to the future: Evidence of ancestral polymorphism in current populations of *Anadenanthera colubrina* var. *cebil* by means of Population Graphs. *Plant Sci* 2, 4.
- Bohonak, A.J., 1999. Dispersal, gene flow and population structure. *Q Rev Biol* 74, 21-44.
- Chakraborty, R., Jin, L., 1993. A unified approach to study hypervariable polymorphisms: Statistical considerations of determining relatedness and population distances. In: Pena, S.D.J., Chakraborty, R., Eppel, J.T., Jeffreys, A.J. (Eds.), *DNA Fingerprinting: State of the Science*. Birkhäuser Basel, Basel, pp. 153-175.
- Dorken, M.E., Eckert, C.G., 2001. Severely reduced sexual reproduction in northern populations of a clonal plant, *Decodon verticillatus* (Lythraceae). *J Ecol* 89, 339-350.
- Earl, D.A., 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conserv Genet Resour* 4, 359-361.
- Excoffier, L., Ray, N., 2008. Surfing during population expansions promotes genetic revolutions and structuration. *Trends Ecol Evol* 23, 347-351.
- Flot, J.F., Magalon, H., Cruaud, C., Couloux, A., Tillier, S., 2008. Patterns of genetic structure among Hawaiian corals of the genus *Pocillopora* yield clusters of individuals that are compatible with morphology. *CR Biol* 331, 239-247.
- 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.
- Gao, H., Williamson, S., Bustamante, C.D., 2007. A Markov chain Monte Carlo approach for joint inference of population structure and inbreeding rates from multilocus genotype data. *Genetics* 176, 1635-1651.
- Gélin, P., Fauvelot, C., Mehn, V., Bureau, S., Rouzé, H., Magalon, H., revised. Superclone expansion, long-distance clonal dispersal and local genetic structuring in the coral *Pocillopora damicornis* type β in Reunion Island, South Western Indian Ocean.
- Gélin, P., Postaire, B., Fauvelot, C., Hélène, M., submitted. Reevaluating species number, distribution and endemism of the coral genus *Pocillopora* Lamarck, 1816 using species delimitation methods and microsatellites.
- Gorospe, K.D., Karl, S.A., 2015. Depth as an organizing force in *Pocillopora damicornis*: intra-reef genetic architecture. *PLoS ONE* 10, e0122127.
- Goudet, J., 2001. FSTAT, a program to estimate and test gene diversities and fixation indices version 2.9. 3.2, updated from Goudet 1995. FSTAT.
- Graham, E.M., Baird, A.H., Connolly, S.R., 2008. Survival dynamics of scleractinian coral larvae and implications for dispersal. *Coral Reefs* 27, 529-539.
- Halkett, F., Simon, J.-C., Balloux, F., 2005. Tackling the population genetics of clonal and partially clonal organisms. *Trends Ecol Evol* 20, 194-201.

- Hallatschek, O., Hersen, P., Ramanathan, S., Nelson, D.R., 2007. Genetic drift at expanding frontiers promotes gene segregation. *Proc Natl Acad Sci U S A* 104, 19926-19930.
- 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.
- Highsmith, R.C., 1982. Reproduction by fragmentation in corals. *Mar Ecol Progr Ser* 7, 207-226.
- Hsu, C.-M., de Palmas, S., Kuo, C.-Y., Denis, V., Chen, C.A., 2014. Identification of scleractinian coral recruits using fluorescent censusing and DNA barcoding techniques. *PLoS ONE* 9, e107366.
- Isomura, N., Nishihira, M., 2001. Size variation of planulae and its effect on the lifetime of planulae in three pocilloporid corals. *Coral Reefs* 20, 309-315.
- Jombart, T., 2008. *adeigenet*: 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.
- Jost, L., 2008. G_{ST} and its relatives do not measure differentiation. *Mol Ecol* 17, 4015-4026.
- 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.
- Kvitt, H., Kramarsky-Winter, E., Maor-Landaw, K., Zandbank, K., Kushmaro, A., Rosenfeld, H., Fine, M., Tchernov, D., 2015. Breakdown of coral colonial form under reduced pH conditions is initiated in polyps and mediated through apoptosis. *Proc Natl Acad Sci U S A* 112, 2082-2086.
- Marti-Puig, P., Forsman, Z.H., Haverkort-Yeh, R.D., Knapp, I.S.S., Maragos, J.E., Toonen, R.J., 2014. Extreme phenotypic polymorphism in the coral genus *Pocillopora*; micro-morphology corresponds to mitochondrial groups, while colony morphology does not. *Bull Mar Sci* 90, 211-231.
- Montoya-Maya, P.H., Schleyer, M.H., Macdonald, A.H.H., 2016. Limited ecologically relevant genetic connectivity in the south-east African coral populations calls for reef-level management. *Mar Biol* 163, 1-16.
- Ohta, T., 1982a. Linkage disequilibrium due to random genetic drift in finite subdivided populations. *Proc Natl Acad Sci U S A* 79, 1940-1944.
- Ohta, T., 1982b. Linkage disequilibrium with the island model. *Genetics* 101, 139-155.
- Oksanen, J., Blanchet, F.G., Kindt, R., Legendre, P., Minchin, P.R., O'Hara, R.B., Simpson, G.L., Solymos, P.M., Stevens, H.H., Wagner, H., 2013. *vegan*: community ecology package.
- Palumbi, S.R., 1994. Genetic divergence, reproductive isolation, and marine speciation. *An Rev Ecol Syst* 25, 547-572.
- 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.
- Peakall, R., Smouse, P.E., 2006. GENALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Mol Ecol Notes* 6, 288-295.
- Peakall, R., Smouse, P.E., 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update *Bioinformatics* 28, 2537-2539.
- Pinzón, J.H., Sampayo, E., Cox, E., Chauka, L.J., Chen, C.A., Voolstra, C.R., LaJeunesse, T.C., Maggs, 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.

- Postaire, B., Aurelle, D., Bourmaud, C.A.F., Bruggemann, J.H., Magalon, H., 2015. Isolation and characterisation of 26 microsatellite loci from a widespread tropical hydrozoan, *Macrorhynchia phoenicea* (Leptothecata, Aglaopheniidae), and cross-amplification in closely related species. *Biochem Syst Ecol* 62, 137-141.
- Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 954-959.
- Richmond, R.H., 1987. Energetics, competency, and long-distance dispersal of planula larvae of the coral. *Mar Biol* 93, 527-533.
- Ridgway, T., Riginos, C., Davis, J., Hoegh-Guldberg, O., 2008. Genetic connectivity patterns of *Pocillopora verrucosa* in southern African Marine Protected Areas. *Mar Ecol Progr Ser* 354, 161-168.
- Rozenfeld, A.F., Arnaud-Haond, S., Hernández-García, E., Eguíluz, V.M., Matías, M.A., Serrao, E., Duarte, C.M., 2007. Spectrum of genetic diversity and networks of clonal organisms. *J Roy Soc Interface* 4, 1093-1102.
- Sammarco, P.W., 1982. Polyp bail-out: an escape response to environmental stress and new means of reproduction in corals. *Mar Ecol Progr Ser* 10, 57-65.
- Schmidt-Roach, S., Lundgren, P., Miller, K.J., Gerlach, G., Noreen, A.M.E., Andreakis, N., 2013. Assessing hidden species diversity in the coral *Pocillopora damicornis* from Eastern Australia. *Coral Reefs* 32, 161-172.
- Schmidt-Roach, S., Miller, K.J., 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 Lin Soc* 170, 1-33.
- Schott, F.A., Xie, S.-P., McCreary, J.P., 2009. Indian Ocean circulation and climate variability. *Rev Geophys* 47.
- Selkoe, K.A., D'Aloia, C.C., Crandall, E.D., Iacchei, M., Liggins, L., Puritz, J.B., von der Heyden, S., Toonen, R.J., 2016. A decade of seascape genetics: contributions to basic and applied marine connectivity. *Mar Ecol Progr Ser* 554, 1-19.
- Siegel, D.A., Mitarai, S., Costello, C.J., Gaines, S.D., Kendall, B.E., Warner, R.R., Winters, K.B., 2008. The stochastic nature of larval connectivity among nearshore marine populations. *Proc Natl Acad Sci* 105, 8974-8979.
- Slatkin, M., 1993. Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* 47, 264-279.
- Souter, P., Henriksson, O., Olsson, N., Grahn, M., 2009. Patterns of genetic structuring in the coral *Pocillopora damicornis* on reefs in East Africa. *BMC Ecol* 9, 19.
- Spalding, M.D., Fox, H.E., Allen, G.R., Davidson, N., Ferdaña, Z.A., Finlayson, M., Halpern, B.S., Jorge, M.A., Lombana, A., Lourie, S.A., Martin, K.D., McManus, E., Molnar, J., Recchia, C.A., Robertson, J., 2007. Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *BioScience* 57, 573-583.
- Torda, G., Lundgren, P., Willis, B.L., van Oppen, M.J., 2013a. Revisiting the connectivity puzzle of the common coral *Pocillopora damicornis*. *Mol Ecol* 22, 5805-5820.
- Torda, G., Lundgren, P., Willis, B.L., van Oppen, M.J.H., 2013b. Genetic assignment of recruits reveals short and long distance larval dispersal in *Pocillopora damicornis* on the Great Barrier Reef. *Mol Ecol* 22, 5821-5834.
- van der Ven, R.M., Triest, L., De Ryck, D.J.R., Mwaura, J.M., Mohammed, M.S., Kochzius, M., 2016. Population genetic structure of the stony coral *Acropora tenuis* shows high but variable connectivity in East Africa. *J Biogeogr* 43, 510-519.
- Vollmer, S.V., Palumbi, S.R., 2007. Restricted Gene Flow in the Caribbean Staghorn Coral *Acropora cervicornis*: Implications for the Recovery of Endangered Reefs. *J Hered* 98, 40-50.

- Weir, B., Cockerham, C., 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358-1370.
- Williams, G.C., 1975. *Sex and evolution*. Princeton University Press

Figure S1.

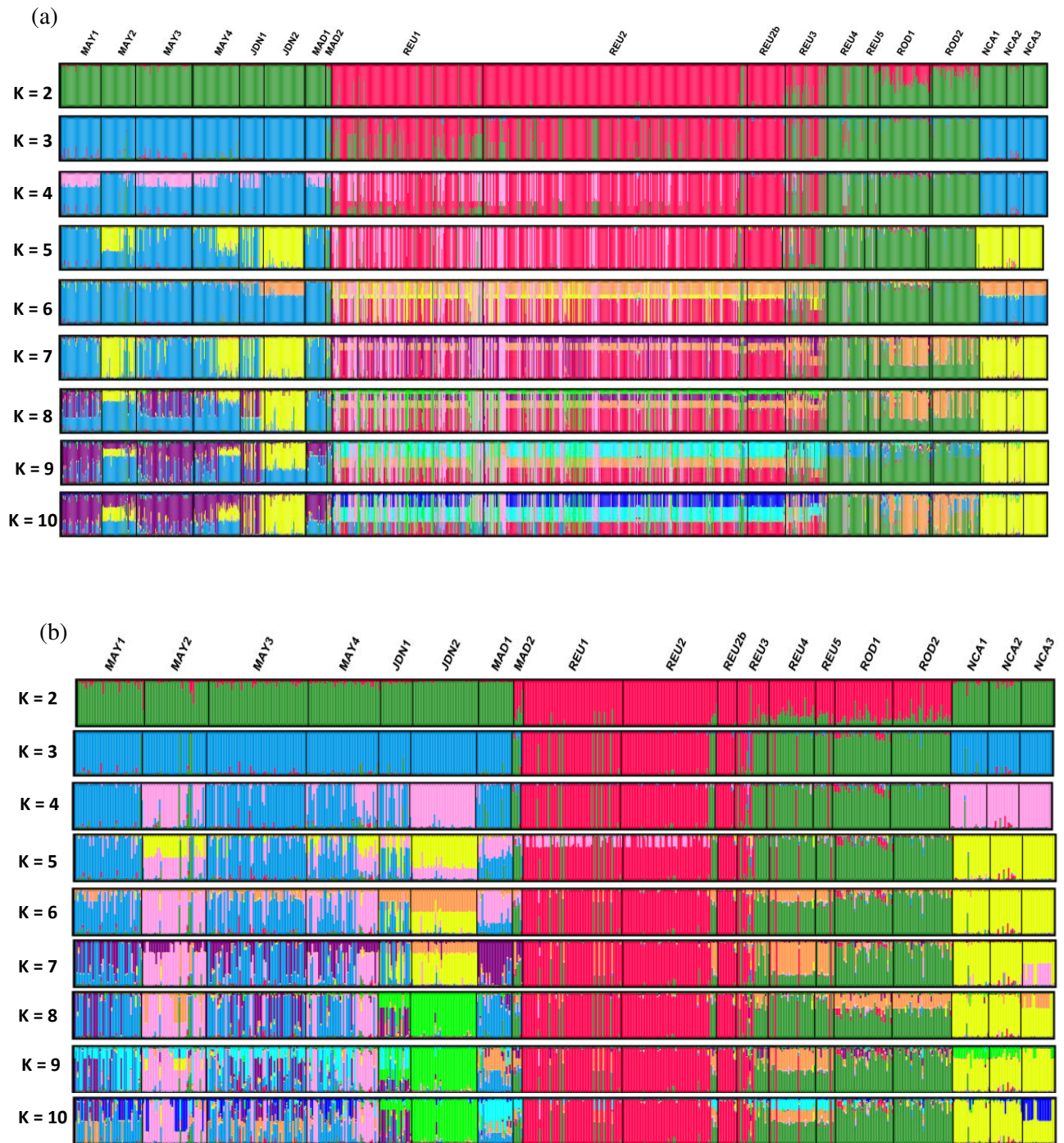
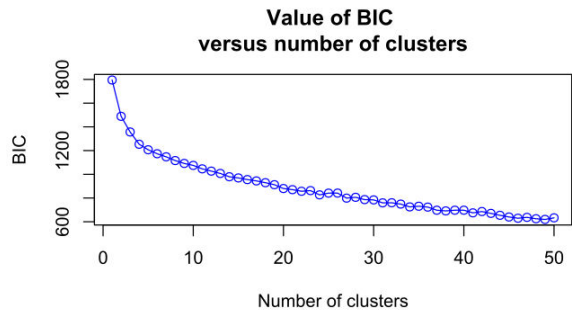


Figure S2.

(a)



(b)

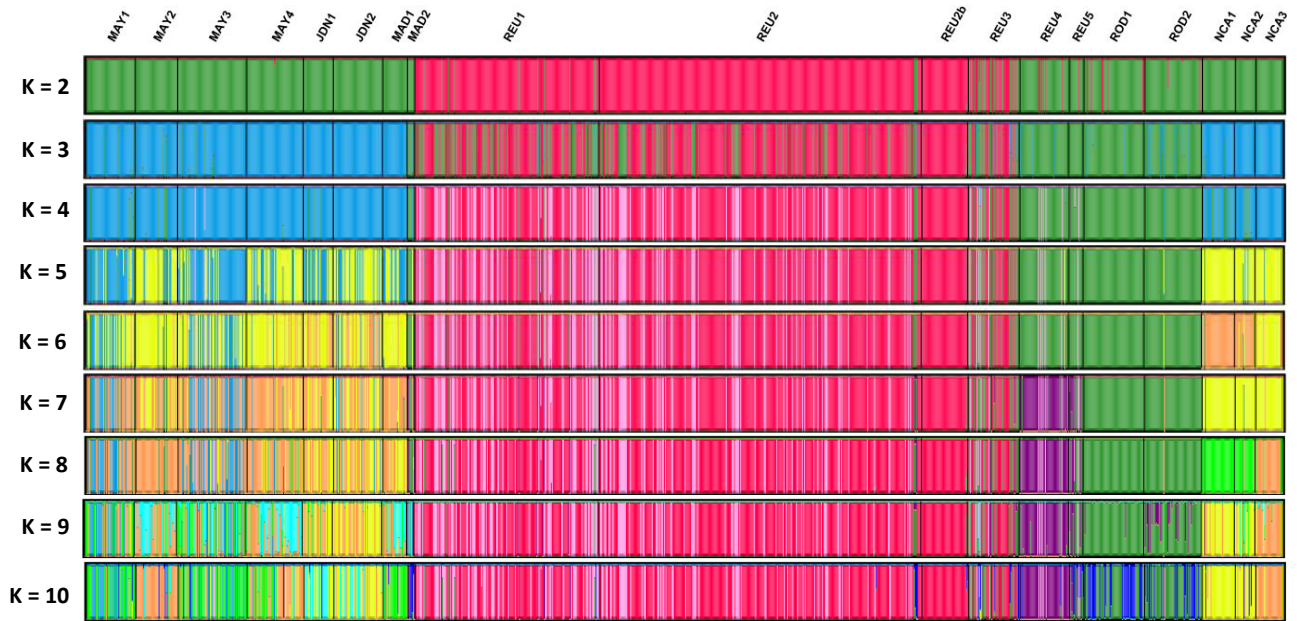
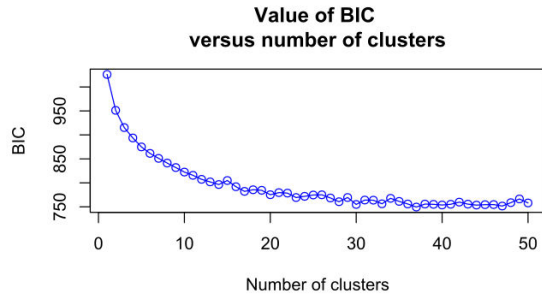


Figure S3.

(a)



(b)



Table S1.

	MAY1	MAY2	MAY3	MAY4	JDN1	JDN2	MAD1	MAD2	REU1	REU2	REU2b	REU3	REU4	REU5	ROD1	ROD2	NCA1	NCA2	NCA4
MAY1	-	0.069 (0.236)	0.029 (0.079)	0.049 (0.16)	0.092 (0.275)	0.081 (0.260)	0.062 (0.181)	0.149 (0.448)	0.155 (0.430)	0.112 (0.376)	0.174 (0.433)	0.182 (0.472)	0.119 (0.379)	0.147 (0.461)	0.131 (0.399)	0.147 (0.429)	0.116 (0.389)	0.090 (0.288)	0.128 (0.370)
MAY2	0.077 (0.260)	-	0.060 (0.205)	0.030 (0.089)	0.084 (0.256)	0.074 (0.244)	0.074 (0.240)	0.131 (0.392)	0.138 (0.388)	0.098 (0.337)	0.169 (0.437)	0.167 (0.441)	0.101 (0.327)	0.120 (0.371)	0.114 (0.353)	0.134 (0.396)	0.106 (0.360)	0.078 (0.250)	0.109 (0.315)
MAY3	0.031 (0.085)	0.062 (0.210)	-	0.037 (0.118)	0.076 (0.224)	0.083 (0.273)	0.059 (0.176)	0.135 (0.395)	0.140 (0.385)	0.104 (0.351)	0.163 (0.404)	0.164 (0.421)	0.116 (0.375)	0.153 (0.495)	0.129 (0.399)	0.150 (0.446)	0.118 (0.405)	0.089 (0.289)	0.136 (0.405)
MAY4	0.066 (0.213)	0.032 (0.094)	0.045 (0.143)	-	0.059 (0.169)	0.065 (0.215)	0.054 (0.163)	0.130 (0.397)	0.122 (0.344)	0.092 (0.319)	0.148 (0.377)	0.150 (0.396)	0.097 (0.320)	0.123 (0.393)	0.108 (0.338)	0.133 (0.404)	0.087 (0.291)	0.069 (0.221)	0.110 (0.327)
JDN1	0.102 (0.298)	0.087 (0.265)	0.083 (0.241)	0.063 (0.177)	-	0.069 (0.182)	0.096 (0.275)	0.165 (0.458)	0.140 (0.337)	0.127 (0.387)	0.161 (0.350)	0.165 (0.372)	0.127 (0.363)	0.155 (0.438)	0.144 (0.399)	0.174 (0.472)	0.118 (0.347)	0.108 (0.315)	0.162 (0.440)
JDN2	0.087 (0.270)	0.073 (0.241)	0.085 (0.278)	0.068 (0.218)	0.074 (0.198)	-	0.082 (0.249)	0.154 (0.451)	0.172 (0.466)	0.137 (0.458)	0.198 (0.490)	0.203 (0.518)	0.128 (0.397)	0.132 (0.384)	0.151 (0.455)	0.157 (0.441)	0.098 (0.302)	0.087 (0.264)	0.140 (0.394)
MAD1	0.069 (0.204)	0.074 (0.241)	0.060 (0.179)	0.056 (0.168)	0.102 (0.292)	0.082 (0.246)	-	0.121 (0.312)	0.152 (0.402)	0.129 (0.428)	0.185 (0.455)	0.179 (0.447)	0.122 (0.377)	0.155 (0.475)	0.136 (0.403)	0.169 (0.490)	0.113 (0.356)	0.095 (0.291)	0.154 (0.448)
MAD2	0.160 (0.474)	0.135 (0.408)	0.141 (0.415)	0.139 (0.420)	0.174 (0.479)	0.159 (0.464)	0.128 (0.342)	-	0.145 (0.310)	0.121 (0.308)	0.176 (0.351)	0.161 (0.318)	0.113 (0.257)	0.143 (0.34)	0.124 (0.279)	0.161 (0.378)	0.180 (0.551)	0.163 (0.488)	0.218 (0.610)
REU1	0.210 (0.495)	0.186 (0.458)	0.184 (0.436)	0.174 (0.417)	0.190 (0.396)	0.229 (0.547)	0.201 (0.468)	0.191 (0.370)	-	0.050 (0.107)	0.018 (0.007)	0.019 (0.017)	0.046 (0.089)	0.140 (0.316)	0.098 (0.205)	0.166 (0.362)	0.184 (0.498)	0.145 (0.377)	0.248 (0.625)
REU2	0.196 (0.484)	0.179 (0.466)	0.174 (0.432)	0.170 (0.430)	0.190 (0.420)	0.223 (0.560)	0.198 (0.489)	0.181 (0.365)	0.011 (0.012)	-	0.068 (0.118)	0.063 (0.119)	0.040 (0.094)	0.097 (0.254)	0.063 (0.158)	0.097 (0.244)	0.143 (0.473)	0.105 (0.328)	0.179 (0.529)
REU2b	0.243 (0.524)	0.226 (0.514)	0.218 (0.470)	0.214 (0.474)	0.229 (0.440)	0.271 (0.597)	0.241 (0.520)	0.224 (0.399)	0.010 (0.007)	0.009 (0.006)	-	0.024 (0.009)	0.066 (0.105)	0.168 (0.347)	0.120 (0.222)	0.187 (0.368)	0.214 (0.539)	0.173 (0.413)	0.281 (0.668)
REU3	0.268 (0.544)	0.250 (0.534)	0.242 (0.488)	0.241 (0.499)	0.255 (0.459)	0.299 (0.622)	0.268 (0.542)	0.248 (0.418)	0.023 (0.019)	0.016 (0.013)	0.006 (0.002)	-	0.068 (0.119)	0.173 (0.369)	0.091 (0.165)	0.166 (0.325)	0.221 (0.574)	0.177 (0.437)	0.283 (0.684)
REU4	0.145 (0.427)	0.117 (0.356)	0.139 (0.421)	0.119 (0.358)	0.151 (0.405)	0.146 (0.423)	0.144 (0.420)	0.125 (0.281)	0.094 (0.160)	0.080 (0.141)	0.116 (0.175)	0.136 (0.192)	-	0.045 (0.074)	0.063 (0.143)	0.101 (0.239)	0.131 (0.399)	0.096 (0.272)	0.198 (0.569)
REU5	0.171 (0.493)	0.134 (0.393)	0.174 (0.526)	0.141 (0.413)	0.182 (0.485)	0.157 (0.434)	0.170 (0.489)	0.143 (0.316)	0.190 (0.350)	0.174 (0.332)	0.225 (0.378)	0.253 (0.400)	0.029 (0.036)	-	0.112 (0.267)	0.119 (0.270)	0.130 (0.368)	0.110 (0.300)	0.201 (0.551)
ROD1	0.150 (0.428)	0.124 (0.369)	0.144 (0.423)	0.118 (0.341)	0.164 (0.433)	0.166 (0.477)	0.149 (0.422)	0.140 (0.316)	0.132 (0.228)	0.111 (0.199)	0.156 (0.241)	0.181 (0.261)	0.076 (0.162)	0.113 (0.238)	-	0.057 (0.112)	0.153 (0.453)	0.112 (0.310)	0.177 (0.458)
ROD2	0.171 (0.468)	0.146 (0.415)	0.170 (0.482)	0.151 (0.426)	0.200 (0.516)	0.171 (0.459)	0.187 (0.523)	0.171 (0.386)	0.217 (0.391)	0.186 (0.341)	0.244 (0.397)	0.271 (0.411)	0.111 (0.235)	0.132 (0.268)	0.085 (0.164)	-	0.172 (0.488)	0.135 (0.366)	0.161 (0.378)
NCA1	0.127 (0.412)	0.118 (0.408)	0.127 (0.427)	0.102 (0.338)	0.132 (0.387)	0.112 (0.346)	0.117 (0.368)	0.186 (0.558)	0.236 (0.553)	0.227 (0.557)	0.274 (0.591)	0.298 (0.602)	0.148 (0.416)	0.152 (0.403)	0.172 (0.483)	0.195 (0.527)	-	0.050 (0.119)	0.129 (0.348)
NCA2	0.095 (0.296)	0.078 (0.253)	0.091 (0.297)	0.071 (0.224)	0.112 (0.324)	0.089 (0.269)	0.093 (0.285)	0.163 (0.483)	0.190 (0.435)	0.178 (0.425)	0.225 (0.471)	0.249 (0.488)	0.106 (0.287)	0.120 (0.308)	0.120 (0.321)	0.147 (0.383)	0.056 (0.139)	-	0.112 (0.295)
NCA3	0.143 (0.389)	0.116 (0.326)	0.153 (0.436)	0.123 (0.344)	0.171 (0.439)	0.148 (0.397)	0.162 (0.452)	0.228 (0.606)	0.321 (0.697)	0.306 (0.690)	0.365 (0.738)	0.393 (0.749)	0.226 (0.585)	0.229 (0.567)	0.209 (0.506)	0.188 (0.412)	0.159 (0.421)	0.120 (0.308)	-

CHAPITRE 4. Étude de la structure génétique chez le corail *Pocillopora eydouxi* : un nouveau puzzle à résoudre

Résumé

Ce chapitre s'intéresse à la structure génétique et à la connectivité génétique du corail *Pocillopora eydouxi*. Dans le chapitre 1, nous avons montré que deux haplotypes mitochondriaux (ORF27 et ORF28) ont été identifiés comme correspondant à cette morpho-espèce et formant une hypothèse d'espèce primaire [PSH, *sensu* Pante et al. (2015)] que les tests d'assignement (basés sur 13 locus microsatellites) scindent en trois hypothèses d'espèces secondaires (SSH). Ainsi, à partir d'un échantillonnage de 2253 colonies réalisé dans trois provinces marines (Sud-Ouest de l'océan Indien, Pacifique Tropical Sud-Ouest et Polynésie Sud-Est), cette étude confirme l'existence des trois SSHs précédemment identifiées (SSH09a, SSH09b et SSH09c) grâce à des tests d'assignements (STRUCTURE et DAPC) réalisés à partir des génotypes multi-locus (13 locus microsatellites). La SSH09a est restreinte à l'océan Indien et les SSH09b et SSH09c sont presque exclusivement restreintes à l'océan Pacifique indiquant qu'il n'existe pas (ou presque pas) de flux de gènes entre ces deux océans pour ces lignées. Par ailleurs, tous les génotypes multi-locus identifiés au cours de cette étude sont uniques, suggérant qu'à l'échelle de cet échantillonnage, la reproduction clonale est inexistante ou extrêmement rare. En outre, l'analyse de la structure génétique révèle un découpage fractal des différentes SSHs. En effet, chaque SSH se divise en cluster qui eux-mêmes se divisent en sous-clusters et ce, avec les deux méthodes d'assignement utilisées (STRUCTURE and DAPC). Cependant, la comparaison des deux différentes méthodes révèle que les assignements diffèrent pour les niveaux de structuration les plus bas. Ainsi, il a été choisi de ne considérer que les niveaux de structuration congruents entre les différentes méthodes en suivant un principe de parcimonie. Cette approche permet d'identifier huit différents clusters pour l'ensemble des SSHs (trois clusters pour SSH09a et SSH09c et deux pour SSH09b). Les analyses de connectivité génétique menées pour chacun des huit clusters identifiés ne permettent pas de révéler de patron de structuration clair. Le partitionnement fractal révélé dans le jeu de données soulève la question de savoir où placer les limites pour définir les unités de base pour les analyses de connectivité.

Examining *Pocillopora eydouxi* population structure: a new beautiful puzzle to solve

Gélin P¹, Fauvelot C², Bigot L.¹, Baly J², Magalon H¹

¹ UMR ENTROPIE (IRD, Université de La Réunion, CNRS), Laboratoire d'excellence-CORAIL; Université de La Réunion, St Denis, La Réunion

² UMR ENTROPIE (IRD, Université de La Réunion, CNRS), Laboratoire d'excellence-CORAIL; centre IRD de Nouméa, New Caledonia

Corresponding author: helene.magalon@univ-reunion.fr; Université de La Réunion, UMR ENTROPIE Faculté des Sciences et Technologies, 15 Bd René Cassin, CS 92003, 97744 St Denis Cedex 09, La Réunion

Abstract

Seascape genetics aim in understanding how movements of an organism through the seascape impact dispersal and gene flow, *i.e.* connectivity among marine populations. This paper reports the genetic structuring in the coral *Pocillopora eydouxi*. Gélin et al. (submitted) found that this species could correspond to two ORF haplotypes (ORF27 and ORF28) forming a primary species hypothesis [PSH, *sensu* Pante et al. (2015)] that split into three secondary species hypotheses [SSH, *sensu* Pante et al. (2015)] with assignment tests performed on multi-locus genotypes based on 13 microsatellite loci. Thus, from a large sampling (2253 colonies) achieved in three marine provinces [Western Indian Ocean (WIO), Tropical Southwestern Pacific (TSP) and Southeast Polynesia (SEP)], this study confirms the existence of the three SSHs previously identified (SSH09a, SSH09b and SSH09c) using 13 microsatellites loci and assignment tests (STRUCTURE and DAPC). SSH09a is restricted to the WIO and SSH09b and SSH09c are almost exclusively in the TSP, suggesting that gene flow is extremely rare between the Pacific Ocean and the Western Indian Ocean. Additionally, all the multi-locus genotypes were unique, suggesting that clonal reproduction might be extremely rare at the scale of our sampling. Moreover, genetic structuring analysis revealed a fractal partitioning for each SSH separately. Indeed, each SSH split into clusters and the clusters into sub-clusters with the two clustering analyses used (STRUCTURE and DAPC). The comparison of this two clustering methods showed that the assignments were not always congruent in lower clustering levels. Thus, we chose arbitrarily to consider only the clustering that was retrieved by different methods of analyses following a principle of parsimony. This approach revealed eight different clusters in PSH09 (three clusters in SSH09a and SSH09c and two in SSH09b). Analyses of genetic connectivity were performed for each cluster separately but no clear pattern was revealed.

Considering the fractal partitioning of our data, we cannot be sure where to place the limits to assess genetic connectivity and further investigations are needed to fully conclude.

Keywords: Bayesian assignments, DAPC, microsatellites, *Pocillopora*, population genetics

Introduction

Seascape genetics, as reflected by the name of the discipline, aim in understanding how movements of an organism through the seascape impact dispersal and gene flow, *i.e.* connectivity among marine populations. This latter could be supposed to present no restrictions as the open sea should be free from barriers. More precisely, genetic connectivity tracks the dispersal of genes, which only accounts for the individuals that successfully reproduce after dispersing [reviewed in Selkoe et al., (2016)]. Indeed, two populations are considered connected when they are able to exchange genes. On the contrary, if these populations do not exchange genes, they should be considered genetically isolated. Commonly, isolation among populations could be the result of geographical barriers or geographic distance. Nevertheless, this isolation might also result from pre- or post-zygotic reproductive isolation. In marine species, post-zygotic barriers are poorly studied because of the difficulty to follow the development of offspring through complex life cycles and through long generation times (Palumbi, 1994).

For a majority of marine species, dispersal is undertaken during the larval phase, and trajectories of larvae, as well as long-term patterns of dispersal, are difficult to predict (Siegel et al., 2008). Nevertheless, a relationship between the duration of the planktonic phase and the dispersal potential has been consistently found [reviewed in Bohonak, (1999)]. Among marine organisms, scleractinians represent a good example for this relationship since larval abilities to move are limited. Actually, the dispersal of these species over large scales (more than hundreds of kilometers) results from regional oceanic vectors: supposedly the longer the planktonic phase duration, the higher the dispersal abilities. Moreover, the duration of the planktonic phase is difficult to estimate because it is not possible to track the larvae *in vivo*. Nevertheless, some species have been studied *in vitro* and the larvae lifetime was estimated between 30 days for *Heliopora coerulea* (Harii et al., 2002), 100 days for *Pocillopora damicornis* (Richmond, 1987) and more than 200 days for five other species [from genera *Acropora*, *Favia*, *Pectinia*, *Goniastrea*, *Montastraea*, Graham et al., (2008)].

Then, to approximate dispersal abilities, the use of genetic tools and assessment of population structure have been widely used and different patterns of dispersal have been identified. Among scleractinians, different patterns of population structuring were revealed. Indeed, focusing on nine species, Ayre and Hughes (2000) highlighted restricted dispersal abilities of larvae in the Great Barrier Reef (GBR) and Torda et al. (2013a), studying *P. damicornis* types α and β [*sensu* Schmidt-Roach et al., (2014)], suggested that type α recruits exhibited predominantly philopatric

recruitment, while the majority of type β recruits were either migrants from identified putative source populations or assumed migrants based on genetic exclusion from all known populations. Torda et al. (2013a) also suggested that locally retained larvae originated predominantly from spawned gametes, while brooded larvae are mainly vagabonds. Nevertheless, with the exception of *P. damicornis sensu lato* (probably mainly from types α and β), species from the genus *Pocillopora* remained underexplored. This particularly true for the morpho-species *P. eydouxi* Milne Edwards, 1860 accepted as *Pocillopora grandis* Dana, 1846 for which almost no data are available.

Like all the species of the genus, it has been described widely distributed in the Pacific Ocean, the Indian Ocean and the Red Sea, and absent from the Atlantic Ocean. Some recent studies have revisited *Pocillopora* taxonomy in the light of molecular data. Indeed, using species delimitation methods based on mitochondrial markers such as the Open Reading Frame (ORF), G lin et al. (submitted) found that this species could correspond to two ORF haplotypes (ORF27 and ORF28, therein) forming a primary species hypothesis [PSH, *sensu* Pante et al. (2015)], named PSH09 therein. The two haplotypes corresponding to this putative species were retrieved in several studies from colonies harboring various morphs: this PSH corresponds to type A in Flot et al. (2008), Clade 3 in [Schmidt-Roach, 2014 #2047@@author-year], type 1 in Pinz n et al. (2013), Clade IIb in Marti-Puig et al. (2014), colonies corresponding to *P. eydouxi*, *P. meandrina*, *P. capitata*, *P. verrucosa*, *P. elegans*, *P. damicornis*, *P. zelli*, *P. woodjonesi* and *P. molokensis* morphs, according to these authors and ourselves. Moreover, in colonies belonging to PSH09 were genotyped using 13 microsatellites loci and assignment tests performed on multi-locus genotypes revealed three secondary species hypotheses [SSH, *sensu* Pante et al. (2015)]: SSH09a restricted in the Western Indian Ocean, and SSH09b and SSH09c restricted to the Pacific Ocean and found in sympatry at the site level. In addition, in Pinz n et al. (2013), using seven microsatellite loci, they found two different genetic clusters within type 1, which were found in sympatry at some locations (Hawaii and Lizard Island). Would they correspond to the SSHs we found remains to be clarified.

Generally, whichever the names given, colonies belonging to PSH09 in G lin et al. (submitted) present common characteristics: large colonies presenting robust erected branches, rounded or flattened, with more or less pronounced verrucae that are uniform in shape and spacing, irrespective to the general *corallum* shape. Its three-dimensional structure is an element of reef structuring and its broad interbranch width provisions habitats for a huge variety of species (notably fish such as damselfish from the Pomacentridae family and *Paracirrhites arcatus*, and arthropods from the genera *Alpheus* and *Trapezia*), these interactions playing a major role in reef

growth and health. It is present on all reef slopes and less frequently in lagoons from surface to 40 m (HM pers. obs.). Nevertheless, *P. eydouxi* morpho-species has been described as a broadcast-spawner (Hirose et al., 2001), as the authors reported from colonies in northern Okinawa, Japan that both male and female gametes were spawned in the water column. This latter study revealed that this morpho-species spawned its gamete sequentially around the sunrise between two and four days after the full moon in June. Moreover, they evidenced that zooxanthellae are maternally inherited in oocytes, suggesting that larvae are autotrophic and intuitively meaning that their dispersal abilities are not limited by intrinsic resources as lecithotrophic larvae could be. Apart from studies of taxonomy, no one has focused yet on population connectivity (strictly speaking) in this species (or complex of).

In front of such a lack of knowledge on *P. eydouxi*, it seems urgent to collect and reinforce data for this too long ignored scleractinian species despite its undeniable role in reef architecture and maintenance over the Indian and Pacific Oceans. Thus, this study aims to (1) confirm or infirm the status of putative species for the three SSHs revealed within PSH09 in (Gélin et al., submitted) and, then, (2) to assess the genetic structuring and connectivity for each putative species found at different spatial scales using 13 microsatellite loci. Indeed, we achieved a hierarchical sampling focusing on three under-studied provinces (from a genetic connectivity point of view): the Western Indian Ocean, the Tropical Southwestern Pacific and Southeast Polynesia.

Material and methods

Sampling

Colonies of *Pocillopora* presenting robust branches were sampled (tip of branches) in three marine provinces extended over six ecoregions (Spalding et al., 2007): the Western Indian Ocean (WIO), the Tropical Southwestern Pacific (TSP) and the Southeast Polynesia (SEP). The sampling followed a hierarchical scheme with several islands within a province and several sites within an island (Figure 1, Table 1). It represented a total of 12 islands (included South Madagascar and New Caledonia) and 60 sites. The colonies were collected from April 2011 to May 2016. All colonies were photographed [except for those from Tromelin Island (Scattered Islands) and Polynesia]. As *corallum* morphology is not a diagnostic character in *Pocillopora* genus, species identification was realized molecularly *a posteriori* of sampling and *a priori* of analyses.

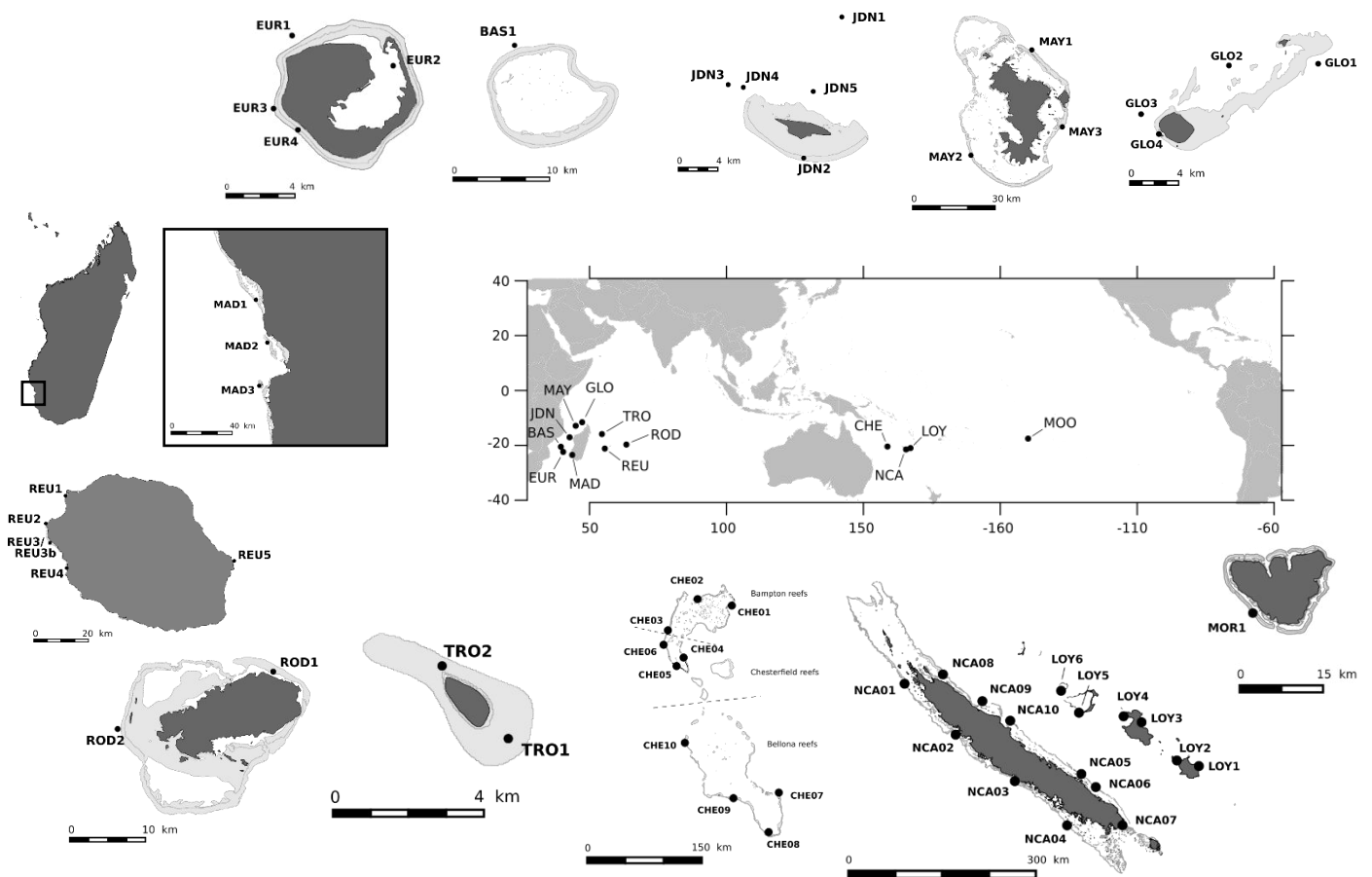


Figure 1. Sampling locations of *Pocillopora* colonies. Populations are numerically identified from the island code [GLO: Glorioso Islands, MAY: Mayotte, JDN: Juan de Nova Island; BAS: Bassas da India, EUR: Europa Island, MAD: Madagascar (Tular), REU: Reunion Island, ROD: Rodrigues Island, CHE: Chesterfield Islands, NCA: Grande Terre (New Caledonia), LOY: Loyalty Islands (New Caledonia) and MOR: Moorea (French Polynesia)]. Constructed using © OpenStreetMap contributors CC BY-SA (<http://www.openstreetmap.org/copyright>) for landmasses and UNEP-WCMC, WorldFish Centre, WRI, TNC (2010). Global distribution of coral reefs, compiled from multiple sources including the Millennium Coral Reef Mapping Project. Version 1.3. Includes contributions from IMaRS-USF and IRD (2005), IMaRS-USF (2005) and Spalding et al. (2001). Cambridge (UK): UNEP World Conservation Monitoring Centre (<http://data.unep-wcmc.org/datasets/1>) for coral reefs.

DNA extraction, sequencing and microsatellite genotyping

From the sampled colonies, DNA was extracted using DNeasy Blood & Tissue kit (Qiagen™). Colonies were genotyped using 13 microsatellite loci (Table 2). Forward primers were indirectly fluorochrome labelled (6-FAM, VIC, NED, PET) by adding a universal M13 tail at the 5'-end. Each amplification reaction was performed as in (Postaire et al., 2015). PCR products were genotyped using an ABI 3730 genetic analyser (Applied Biosystems) and allelic sizes were determined with GeneMapper V4.0 (Applied Biosystems) using an internal size standard (Genescan LIZ-500, Applied Biosystems). Because colonies were sampled based on

their *corallum* macromorphology, *P. eydouxi* lineage identity was verified *a priori* using assignment tests performed with STRUCTURE (Pritchard et al., 2000) combining all sampled colonies which included the 943 colonies from G elin et al. (submitted) that were sequenced for the Open Reading Frame (ORF). From that, 2253 colonies grouped together with colonies assigned in PSH09 (corresponding to haplotype ORF27) from G elin et al. (submitted) and were extracted, constituting the final dataset.

MLG identification and assignment to the SSHs

To check for clonal propagation among the sampled colonies, identical multi-locus genotypes (MLG) were identified using the R (R Development Core Team 2016) package *RClone* (Bailleul et al., 2016). Then, keeping one representative per MLG, we tried to retrieve the three SSHs (SSH09a, SSH09b and SSH09c) identified in G elin et al. (submitted). So, we performed assignment tests on the 2253 colonies using STRUCTURE (Pritchard et al., 2000) for $K = 1-10$, running five chains with 2×10^6 generation steps after a burn-in of 2×10^5 assuming admixture and correlated frequencies. This analysis assumes that, within a set of samples, there are K populations and individuals are assigned to each putative population under Hardy Weinberg equilibrium (HWE) and minimized linkage disequilibrium (LD). We used the statistic proposed by Evanno et al. (2005) to estimate the number of clusters K implemented in STRUCTURE HARVESTER (Earl and von Holdt, 2012). The STRUCTURE outputs were summarized with CLUMPP v. 1.0 (Jakobssen and Rosenberg 2007) and formatted with DISTRUCT v. 1.1 (Rosenberg 2004). As STRUCTURE demands strong assumptions on the genetic clusters, we performed in parallel a Discriminant Analysis of Principal Components [DAPC; Jombart et al., (2010)] in order to test whether the structuring observed with STRUCTURE v 2.3.4 can be found with DAPC. This analysis does not make any assumption about HWE or LD and transforms genotypes using PCA as a prior step to a discriminant analysis. DAPC was applied using the *adegenet* package (Jombart, 2008) for R (R Development Core Team 2016). We used the function `find.clusters()` to assess the optimal number of groups with the Bayesian information criterion (BIC) method (*i.e.* K with the lowest BIC value is ideally the optimal number of clusters). We tested values of K ranging from 1 to 30, but BIC values may keep decreasing after the true K value in case of genetic clines and hierarchical structure (Jombart et al., 2010). Therefore, the rate of decrease in BIC values was visually examined to identify values of K after which BIC values decreased only subtly (Jombart et al., 2010). The `dapc()` function was then executed using the best grouping, retaining axes of PCA sufficient to explain $\geq 90\%$ of total variance of data.

Table 1. Sampling of *Pocillopora eydouxi* colonies. For each site, are indicated the island, the ecoregion and the province. The total number of sampled colonies (N_{TOT}) and the number of sampled colonies are indicated per SSH (SSH09a, SSH09b and SSH09c) for each site.

Province	Ecoregion	Island	Population	Position GPS N'	Position GPS E'	N_{TOT}	SSH09a	SSH09b	SSH09c	
Western Indian Ocean	Western and Northern Madagascar	Mayotte	MAY1	-12.62755	45.1775	45	45			
			MAY2	-12.97175	44.97857	35	31		4	
			MAY3	-12.8791	45.27705	42	42			
		Glorioso Island	GLO1	-11.53364	47.40126	31	31			
			GLO2	-11.5349	47.33545	29	29			
			GLO3	-11.57074	47.27066	45	45			
			GLO4	-11.58543	47.28376	19	19			
		Juan de Nova Island	JDN1	-16.95337	42.76011	58	57			1
			JDN2	-17.08177	42.72536	44	43			1
	JDN3		-17.01493	42.65645	48	48				
	JDN4		-17.01737	42.67027	31	31				
	JDN5		-17.02115	42.73402	29	29				
	Bassas da India	BAS1	-21.43418	39.65295	28	28				
	Europa	EUR1	-22.33505	40.33439	60	60				
		EUR2	-22.35081	40.38706	39	39				
		EUR3	-22.373	40.32483	9	9				
		EUR4	-22.38403	40.3375	11	10			1	
	Madagascar (Tuléar)	MAD1	-23.15016	43.56768	48	48				
		MAD2	-23.40106	43.63468	49	48			1	
		MAD3	-23.65441	43.58764	68	68				
	Macarene Islands	Reunion Island	REU1	-20.94338	55.27963	50	44	3	3	
			REU2	-21.03401	55.21478	38	38			
			REU3	-21.09812	55.22826	39	39			
			REU3b	-21.09778	55.23239	15	15			
			REU4	-21.18029	55.28371	33	33			
		REU5	-21.15695	55.8362	47	47				
		Rodrigues	ROD1	-19.66726	63.46867	47	45			2
ROD2			-19.75397	63.46803	54	53			1	
ROD3			-19.65073	63.4002	35	34			1	
ROD4			-19.73523	63.28396	47	47				
Cargados Carajos/Tromelin Island	TRO1	-15.90072	54.53346	16	16					
	TRO2	-15.88354	54.51773	16	16					
Chesterfield Islands	Bampton Reefs	CHE01	-19.20365	158.93706	38		2	36		
		CHE02	-19.11688	158.6	39			39		
		CHE03	-19.48349	158.25203	30			30		
	Chesterfield Reefs	CHE04	-19.80292	158.43595	38		6	32		
		CHE05	-19.90287	158.3533	61		5	56		
		CHE06	-19.65317	158.20108	50		2	48		
	Bellona Reefs	CHE07	-21.38758	159.55353	48			48		
		CHE08	-21.85115	159.43348	35		4	31		
		CHE09	-21.45272	159.02081	56		7	49		
		CHE10	-20.80306	158.45104	53		5	48		
Tropical Southwestern Pacific	New Caledonia	NCA01	-20.30244	163.88345	26		14	11		
		NCA02	-21.0339	164.61812	42		28	8		
		NCA03	-21.70086	165.46918	31		10	21		
		NCA04	-22.33536	166.2219	25		11	14		
		Grande Terre	NCA05	-21.59848	166.42591	54		3	50	
			NCA06	-21.7857	166.63385	56		13	43	
		NCA07	-22.33438	167.01815	48		15	29		
		NCA08	-20.16931	164.43608	61		13	46		
		NCA09	-20.5531	165.00238	38		13	25		
		NCA10	-20.8337	165.40346	37		4	33		
Loyalty Islands	LOY1	-21.40448	167.8007	10		3	7			
	LOY2	-21.48428	168.11743	13			13			
	LOY3	-20.85568	167.29346	25		15	3			
	LOY4	-20.77063	167.03459	55		33	21			
	LOY5	-20.71891	166.39266	47		28	18			
	LOY6	-20.404	166.13475	49		38	7			
Southeast Polynesia	Society Islands	Moorea	MOR1	-17.57569	-149.87803	10		10		
TOTAL						2253	1187	285	781	

The genotypic linkage disequilibrium was assessed with FSTAT v 2.9.3 both on the whole dataset and for each SSH separately (Goudet, 2001) and differentiation between pairs of SSHs was estimated using the F_{ST} (Weir and Cockerham, 1984) with ARLEQUIN v 3.5 (Schneider et al., 2000). For each SSH, the allelic frequency distributions were plotted and the number of alleles (N_a) and the number of private alleles (N_p) were assessed with FSTAT v 2.9.3 (Goudet, 2001). Within each SSH, differentiation between pairs of populations (referring to the set of colonies sampled at the same site) was estimated using the F_{ST} (Weir and Cockerham, 1984) with ARLEQUIN v 3.5 (Schneider et al., 2000).

Structuring analyses within each SSH

Further analyses were performed on each SSH sub-dataset separately. We applied a hierarchical approach, once we identified our SSHs within the whole dataset, we tried to identify the clusters within each SSH, and then the sub-cluster within these latter clusters. First, to determine the most likely number of genetically homogenous clusters (K) within each SSH, a Bayesian analysis was performed using STRUCTURE v 2.3.4 (Pritchard et al., 2000) with the same conditions as above, except that 10 chains were run per K . For the same reasons exposed to identify the SSHs, we performed in parallel DAPC analyses. Then, for each cluster found within each SSH, the allelic frequencies for each locus were plotted and F_{ST} (Weir and Cockerham, 1984) among clusters for each locus were calculated using ARLEQUIN v 3.5 (Schneider et al., 2000). As STRUCTURE and DAPC may be sensitive to hierarchical sub-structuring, when appropriate, each cluster found within each SSH was rerun separately for both analyses. Additionally, Minimum Spanning Networks (MSN) based on the Shared Alleles Distance (DAS) were drawn with individuals coloured according to the clusters found in the different analyses. However, these methods used to detect the most likely number of clusters (Evanno and BIC) might differ between each other and as they are purely mathematical, they might not reflect the biological truth. Thus we decided to conserve the clusters that were retrieved by the different methods congruently (STRUCTURE /DAPC/MSN/ F_{ST} combined with hierarchical approach).

Structuring analyses within clusters of each SSH

Then, considering the clusters finally kept within each SSH, a hierarchical analysis of molecular variance (AMOVA) was performed using ARLEQUIN v 3.5 (Schneider et al., 2000) considering the whole dataset with SSH as group and clusters within SSH as populations. Then,

for each cluster found within each SSH, differentiation indices (F_{IS} and F_{ST} (Weir and Cockerham, 1984)) and departure from HWE were estimated among populations with FSTAT v 2.9.3 (Goudet, 2001). Populations are considered as the set of colonies sampled either at the same site or in the same island, depending on the analysis and the number of colonies sampled. Finally, an AMOVA was performed for each cluster with ecoregion as groups and island as population using ARLEQUIN v 3.5 (Schneider et al., 2000).

RESULTS

MLG identification and assignment to the SSHs

Among the 2253 colonies studied, each represented a unique MLG. So the final dataset comprised 2253 MLGs, which all presented haplotype ORF27 and were all were assigned to PSH09 when confronted to the dataset comprising all the *Pocillopora* colonies from G elin et al. (submitted). First, the Bayesian assignment tests allowed to assign the colonies in three clusters that did correspond to the three SSHs previously found (G elin et al., submitted): the colonies from this previous study included in the present analysis were assigned similarly than in (G elin et al., submitted), corroborating the SSHs (SSH09a, SSH09b and SSH09c; Figure 2). In addition, all the colonies from TSP and SEP ($n = 1075$) were assigned to both SSH09b and SSH09c (none to SSH09a), while all colonies from WIO ($n = 1205$) were assigned to SSH09a (except three and 15 that were assigned to SSH09b and SSH09c, respectively) (Table 1). This first partition obtained from the Bayesian analysis was considered as the first level of differentiation. In parallel, concerning DAPC, the colonies were assigned to the same three SSHs, but the pattern appeared at $K = 4$, where SSH09c was split into two clusters (observed from $K = 3$). Finally, a similar clustering pattern was found for both types of analyses for $K = 5$. Moreover, high values of F_{ST} were observed among the three SSHs (considering the partition from Structure for $K = 3$): 0.164^{***} for SSH09a/SSH09b, 0.252^{***} for SSH09a/SSH09c and 0.229^{***} for SSH09b/SSH09c. Then, considering the whole dataset irrespective to SSHs, 21 tests of genotypic disequilibrium were significant over 78, indicating 27% of genotypic disequilibrium among the 13 loci after Bonferroni correction, while, considering each SSH separately, none of the 78 tests revealed disequilibrium within both SSH09a and SSH09c and 10 out of 78 within SSH09b.

In addition, looking at the allelic frequencies for each SSH (Figure S1), the mean number of alleles per locus was high, varying from 13.00 ± 0.87 for SSH09b to 13.77 ± 1.18 for SSH09c (Table 2). For each SSH, between 10 and 14% of private alleles were identified, the mean number of private alleles varying from 1.46 ± 0.33 for SSH09c to 1.85 ± 0.41 for SSH09b (Table 2). In addition, loci Pd11 and Pd13 showed a weaker rate of amplifications on colonies from SSH09a than from the two others (data not shown). Considering all these results, this indicated a high genetic differentiation among the three SSHs, each with its own dynamics, indicating restricted gene flow between these SSHs. Thus, from now, we will consider these three SSHs as independent genetic lineages and will perform the subsequent analyses on each SSH separately ($N_{SSH09a} = 1187$, $N_{SSH09b} = 285$, $N_{SSH09c} = 781$; Table 1).

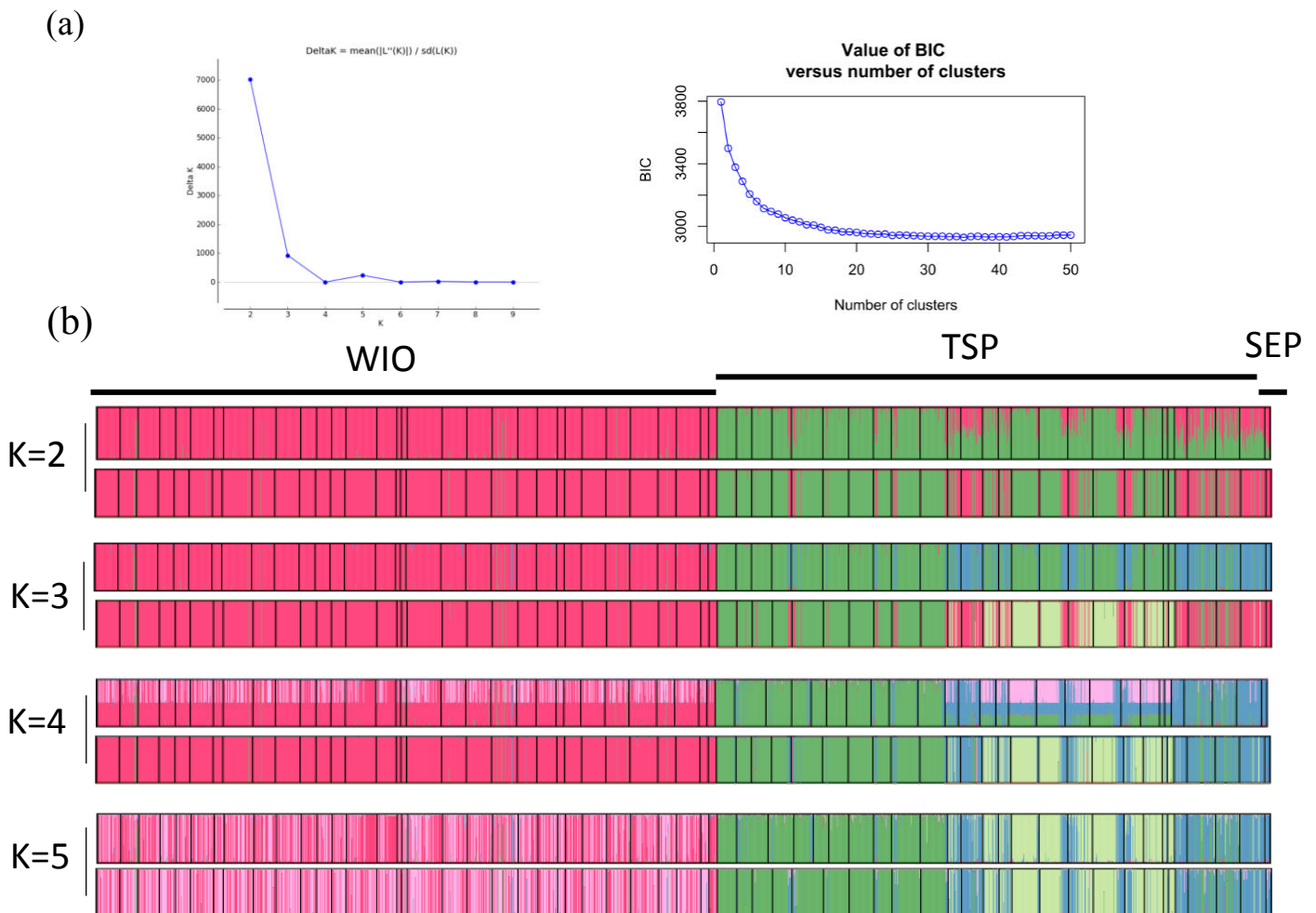


Figure 2. Results of the assignment tests including 2253 colonies. (a) $\Delta(K)$ and BIC (b) Structure (above) and DAPC (below) plots are presented from $K = 2$ to $K = 5$. The colonies used in the analyses were sampled in three ecoregions: the Western Indian Ocean (WIO), the Tropical Southwestern Pacific (TSP) and Southeast Polynesia (SEP).

Concerning SSH09a, the pairwise F_{ST} among populations (Table S1a) ranged from 0.000^{NS} to 0.129*** (TRO2/ROD4) without following a geographic pattern. Some populations from the same island were differentiated (MAY, JDN, REU) while populations from different islands showed no genetic differentiation (e.g. EUR with ROD). For SSH09b, the pairwise F_{ST} among populations (Table S1b) ranged from 0.000^{NS} to 0.120*** with no obvious geographic correlation and for SSH09c, from 0.000^{NS} to 0.349*** (Table S1c), with CHE populations highly differentiated from NCA and LOY populations and no obvious pattern with NCA populations nor LOY populations.

Table 2: Summary statistics of the microsatellite loci used to genotype *Pocillopora eydouxi* colonies, assigned to PHS09 in G elin et al. (submitted). The number of alleles per locus (N_a) and the number of private alleles per locus (N_p) for the three SSHs (SSH09a, SSH09b and SSH09c) are given. The total number of alleles (Total) and the mean number of allele per locus and standard error (Mean \pm se) are also indicated for each SSH.

Locus	N_a (SSH09a)	N_p (SSH09a)	N_a (SSH09b)	N_p (SSH09b)	N_a (SSH09c)	N_p (SSH09c)	Source
Pd2-001	10	1	9	0	13	3	Starger et al. (2007)
Pd3-004	11	1	14	1	17	3	
Pd3-005	19	0	18	4	20	2	
Pd2-006	13	0	16	5	11	1	
Pd3-008	9	1	8	2	6	0	
Pd3-009	12	2	12	2	10	0	
Poc40	7	0	13	3	10	1	Pinz�on and LaJeunesse (2011)
PV2	12	4	8	2	11	2	Magalon et al. (2004)
PV7	9	1	14	1	15	2	
Pd4	18	3	12	1	13	2	Torda et al. (2013b)
Pd11	19	3	14	1	16	0	
Pd13	21	4	15	2	21	3	
Pd3-EF65	15	1	16	0	16	0	Gorospe and Karl (2013)
Total	175	21	169	24	179	19	
Mean \pm se	13.46 \pm 1.25	1.62 \pm 0.40	13.00 \pm 0.87	1.85 \pm 0.41	13.77 \pm 1.18	1.46 \pm 0.33	

Structuring analyses within each SSH and within each cluster in each SSH

SSH09c

Assignment tests for SSH09c revealed an optimal value of K for $K = 2$ or $K = 3$, using the Evanno's method (Figure 3b). At $K = 2$, the structuring pattern nearly corresponded to a geographical pattern. Indeed, on one hand, populations from Chesterfield Islands together with LOY4 and LOY5 (Loyalty Islands) were grouped in a first cluster (SSH09c-1, Figure 3c), and on the other hand, all the other populations from New Caledonia (Grande Terre and Loyalty

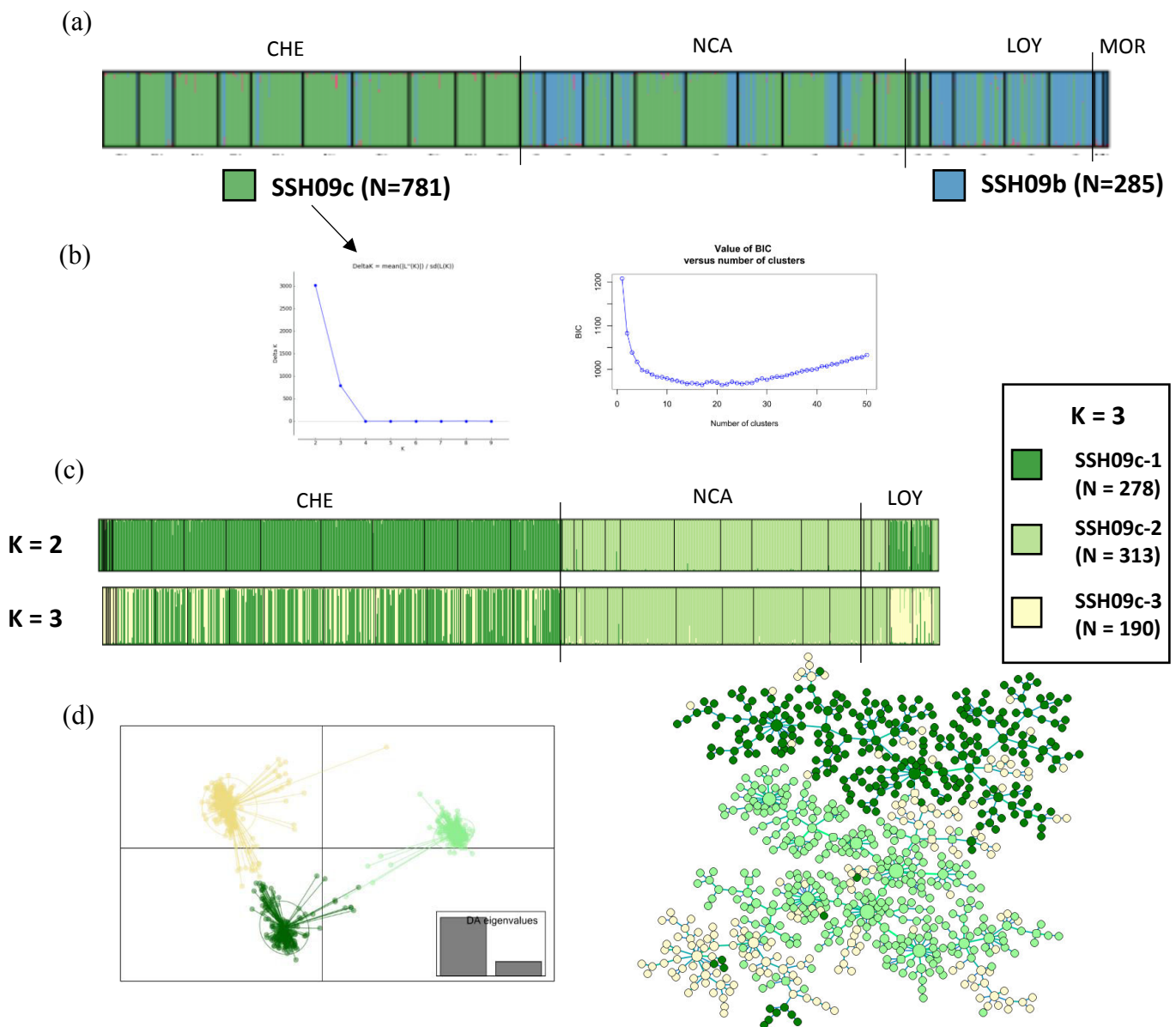


Figure 3. Results of the assignment tests for SSH09c. (a) The assignment result for colonies of PSH09 sampled in the Pacific Ocean. Considering only colonies from SSH09c, (b) $\Delta(K)$ and BIC values and (c) the STRUCTURE plots for $K = 2$ and $K = 3$ are presented. (d) Results of the DAPC assignment for $K = 3$ and (d) Minimum Spanning Network both colored according to the three clusters identified by STRUCTURE.

Islands) together with the 15 individuals from WIO composed the second cluster (Figure 3c, SSH09c-2) (Table S1c). At $K = 3$, the SSH09c-1 split into two clusters, individuals from Loyalty Islands grouped in a third cluster (SSH09c-3) and individuals from Chesterfield Islands in both clusters. Using DAPC without any assumption on population properties, individuals were assigned nearly exactly to the same clusters as in Structure, either for $K = 2$ or $K = 3$ (Figure 3d). The three genetic groups were retrieved in the minimum spanning network when individuals are colored by their cluster. The genetic differentiation among these three clusters was high and of the same order of magnitude with F_{ST} ranging from 0.339*** (SSH09c-1/SSH09c-2) and 0.163*** (SSH09c-1/SSH09c-3) (Table 3). So, in view of all these results, we chose to keep these three clusters (SSH09c-1, SSH09c-2 and SSH09c-3) and to consider them separately for further analyses.

Table 3. Pairwise F_{ST} values between the three clusters of SSH09c. *** $P < 0.001$

	SSH09c-1	SSH09c-2	SSH09c-3
SSH09c-1	-		
SSH09c-2	0.339***	-	
SSH09c-3	0.163***	0.221***	-

Considering SSH09c-1 which was found in Chesterfield Islands and in the WIO (but these 15 individuals were excluded from further analyses), pairwise F_{ST} among populations (sites) showed that populations from these reefs were barely differentiated together, except CHE02 (the northernmost population in Bampton reefs) that differentiated from all the others populations (mean $F_{ST} = 0.079$ ***; Table S1a).

Considering SSH09c-2 which was found in New Caledonia and few colonies in Loyalty Islands; some sites were pooled to get sufficient samples, *e.g.*, NCA01 with NCA02 forming NCA01-02), the structuring pattern was less obvious (Table S1b) with F_{ST} ranging from 0.000 to 0.067*** (NCA01-02 with NCA10). Considering SSH09c-3 which was found in Chesterfield Islands and Loyalty Islands, LOY4 in Lifou was found differentiated from all other populations (mean $F_{ST} = 0.085$ ***; Table S1c) and CHE02 was also found differentiated from all other populations except from CHE07 and CHE09 (mean $F_{ST} = 0.058$ ***; Table S1c).

SSH09b

Concerning SSH09b, the Evanno's method indicated optimal values of K for $K = 2$ or $K = 4$ (Figure 4b). At $K = 2$, colonies from Chesterfield Islands and colonies from French Polynesia were assigned in the first cluster (SSH09b-1; Figure 4b). Colonies from New Caledonia were assigned in both clusters (SSH09b-1 and SSH09b-2) with no geographic pattern. Moreover, colonies of each cluster were not evenly spread across populations. At $K = 3$ and $K = 4$, cluster SSH09b-1 split into two and three, respectively, colonies from Chesterfield Islands and French Polynesia remaining in the same cluster, and colonies from New Caledonia splitting into two and three genetic clusters with no geographical pattern, respectively (Figure S2). Using DAPC, the individual assignments were similar only for $K = 2$. When analyzing both clusters separately, cluster SSH09b-1 ($n = 222$) was sub-divided into four clusters, either with Structure or DAPC,

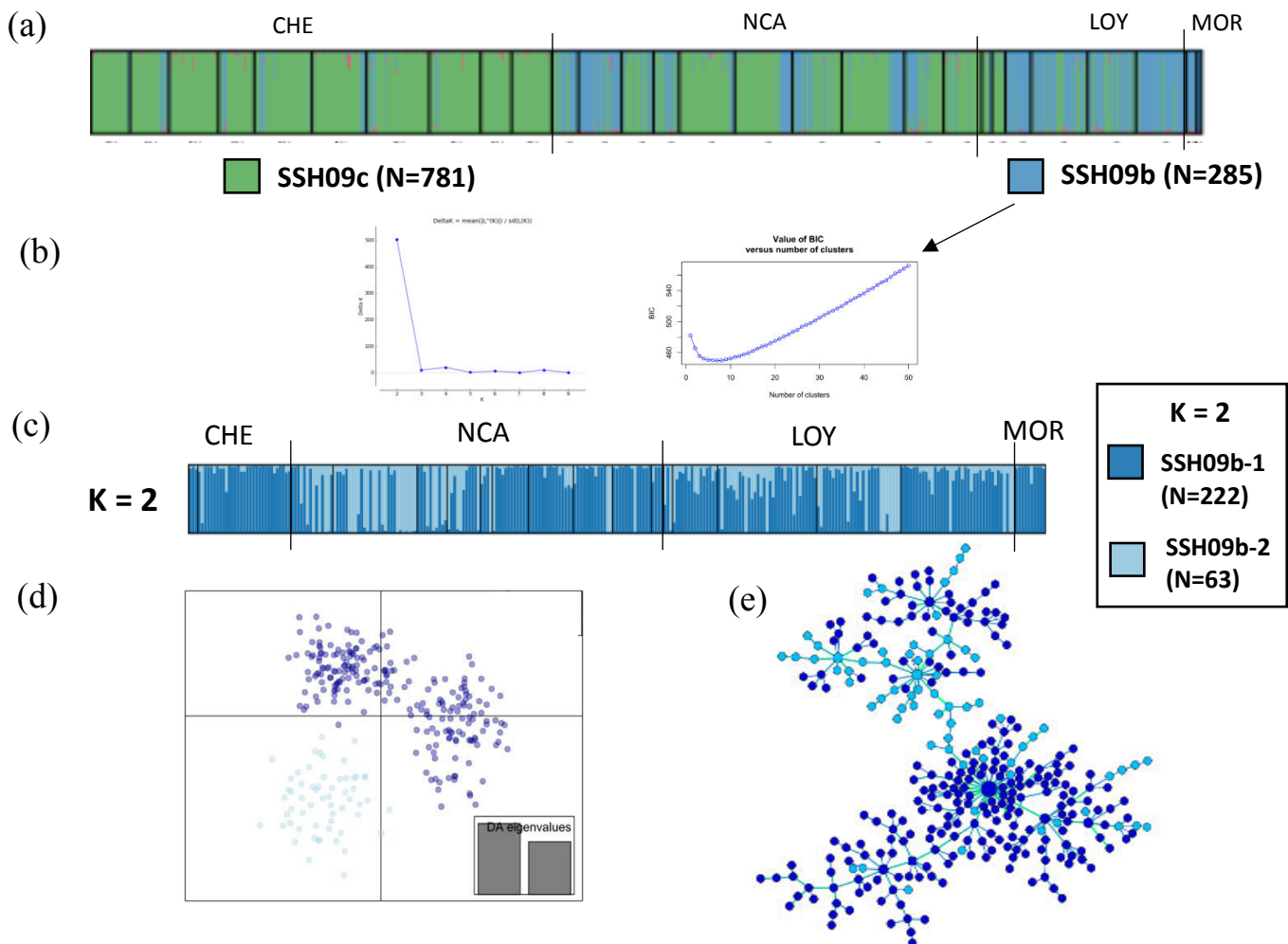


Figure 4. Results of the assignment test for SSH09b. (a) The assignment result for colonies of PSH09 sampled in the Pacific Ocean. Considering only colonies from SSH09b, (b) $\Delta(K)$ and BIC values and (c) the STRUCTURE plots for $K = 2$ are presented. (d) Results of the DAPC assignment for $K = 2$ and (e) the Minimum Spanning Network both colored according to the two clusters identified by STRUCTURE.

but individual assignments differed between both methods. Meanwhile, cluster SSH09b-2 ($n = 63$) was sub-divided into three clusters with Structure, while using DAPC, the individual assignments were congruent with Structure only for $K = 2$. Thus, we considered that SSH09b split into two primary clusters (SSH09b-1 and SSH09b-2; Figure 4c and 4d). The minimum spanning network revealed also these two clusters (Figure 4e). These two clusters were highly differentiated ($F_{ST} = 0.131^{***}$).

Then, when considering the first cluster (SSH09b-1), which was the more represented ($n = 222$ colonies over 285), the F_{ST} ranged from 0.000^{NS} to 0.072^{***} (due to small sample size, some populations were pooled: all CHE populations together, NCA01/02, NCA03/04, NCA05/06/07 and NCA08/09/10). Colonies from Chesterfield Islands were differentiated with the populations from the East Coast of Grande Terre and from Loyalty Islands, but not from populations for the West Coast of Grande Terre. Populations from the West Coast were differentiated with the South-East Coast of Grande Terre, and from Lifou (Loyalty Islands). As cluster SSH09b-2 presented few individuals per population, pairwise F_{ST} among populations were not assessed.

SSH09a

Considering all the colonies assigned to SSH09a and all the loci, the Evanno's method indicated an optimal K value for $K = 2$ followed by $K = 3$ (Figure 5a), DAPC showing similar individual assignments than STRUCTURE for both K : At $K = 2$, we observed two clusters SSH09a-1 and SSH09a-2, then, at $K = 3$, SSH09a-2 split into two (SSH09a-2 and SSH09-3). Considering $K = 2$, the first cluster (SSH09a-1) when reanalyzed alone did not separate in any more cluster using STRUCTURE and gave a compact cloud with DAPC. Conversely, the second cluster (SSH09a-2) when reanalyzed separately, did separate in three distinct clusters with STRUCTURE but with weak ΔK value (SSH09a-2, SSH09a-3 and SSH09a-4) (data not shown), but when compared with DAPC, SSH09a-3 and SSH09a-4 were overlapping. Then, keeping all the individuals together and looking at the assignment for $K = 4$, STRUCTURE did find four clusters, but they did not correspond to the ones found when treating the two first clusters independently (in this case, SSH09a-1 and SSH09a-2 both split into two). When using DAPC using all the individuals for $K = 4$, DAPC assignments diverged from those of STRUCTURE. So, we concluded that the individuals could be grouped into three genetic clusters (SSH09a-1, SSH09a-2 and SSH09-3 from Figure 5).

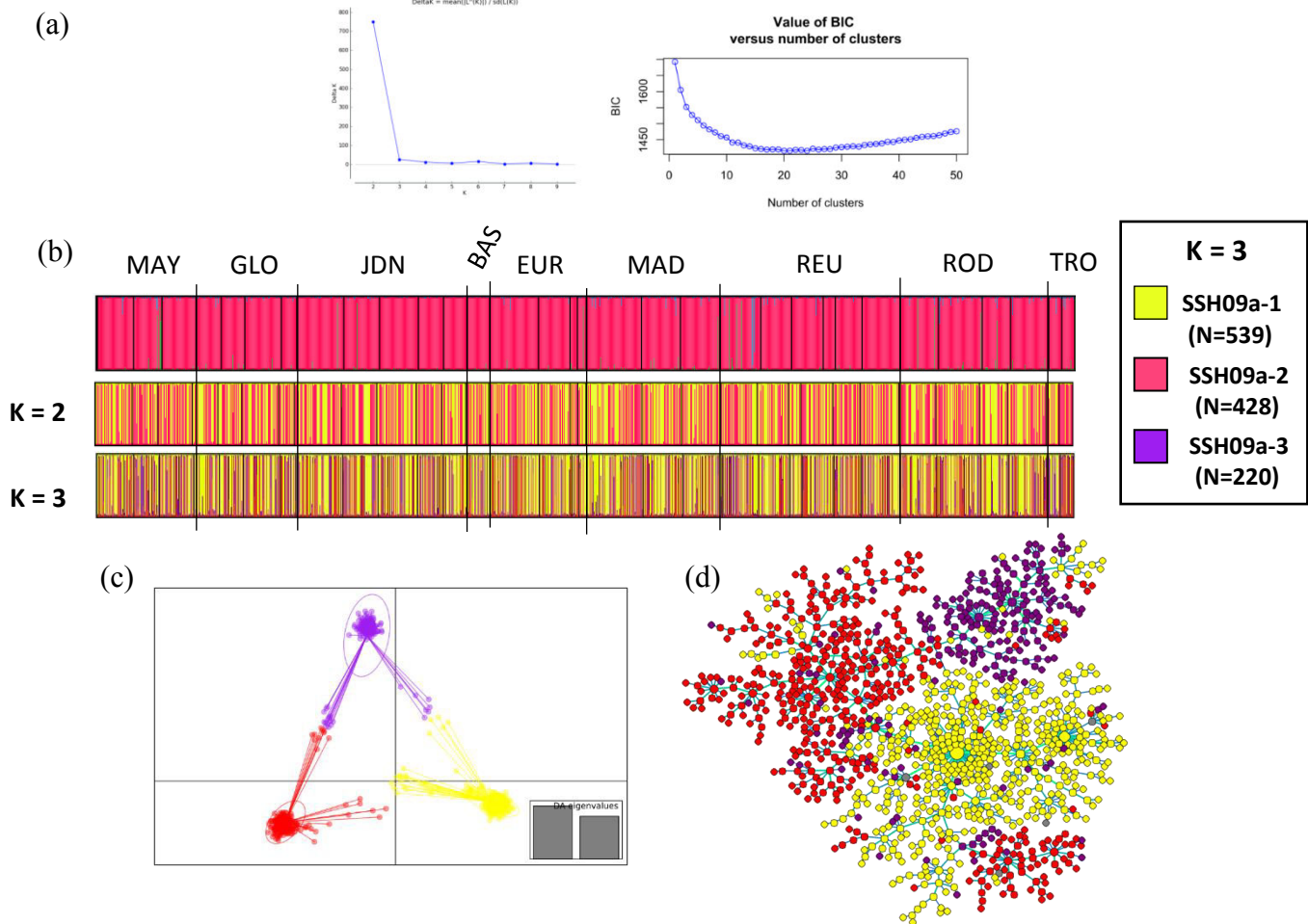


Figure 5. Results of the assignment test for SSH09a. (a) The assignment result for colonies of PSH09 sampled in the Indian Ocean. Considering only colonies from SSH09a, (b) $\Delta(K)$ and BIC values and (c) the STRUCTURE plots for $K = 2$ and $K = 3$ are presented. (d) Results of the DAPC assignment for $K = 2$ and (d) the Minimum Spanning Network both colored according to the three clusters identified by STRUCTURE.

Surprisingly, these latter, whatever the method used to assign the individuals (STRUCTURE/DAPC, all individuals/sub-structuring), were found nearly evenly distributed in all the populations with no geographical pattern, cluster SSH09a-1 being the most represented (Figure 5). Moreover, pairwise F_{ST} among the three clusters were of the same order from 0.128*** and 0.151*** (Table 4).

For SSH09a, considering the populations irrespective to the three clusters, F_{ST} values varied from 0.000^{NS} (for several population pairs, Table S2a) to 0.129*** for TRO2/ROD4, possibly expressing the differences in frequencies of each of the three clusters among each pair of populations.

Table 4. Pairwise F_{ST} values between the three clusters of SSH09a. (SSH09a-1, SSH09a-2 and SSH09a-3) *** $P < 0.001$

	SSH09a-1	SSH09a-2	SSH09a-3
SSH09a-1	-		
SSH09a-2	0.144***	-	
SSH09a-3	0.128***	0.152***	-

Now, considering only cluster SSH09a-1 with 539 colonies, F_{ST} values between pairs of populations ranged from 0.000^{NS} to 0.008* (GLO/MAD). The population from the South of Madagascar (MAD) was slightly differentiated from Glorioso Islands (GLO), Mayotte (MAY) and Juan de Nova Island (JDN) mean $F_{ST} = 0.007^*$, but not from Europa (EUR), Reunion Island (REU) and Rodrigues (ROD). Considering SSH09a-2, the population from Tromelin (TRO) was differentiated from all others (mean $F_{ST} = 0.064^{***}$) and MAD was differentiated from MAY and JDN ($F_{ST} = 0.026^{***}$ and 0.009^{**} , respectively). SSH09a-3 showed similar pattern of slight differentiation of MAD with all the populations, except EUR and ROD (mean $F_{ST} = 0.019^{***}$), the remaining populations being not differentiated from each other.

Additionally, looking at the allelic frequencies for the three clusters (Figure S3, the locus PV7 showed a high variance between the three clusters identified. So it was removed from this sub-dataset and the analyses were performed again. Then, STRUCTURE got difficulties in assigning individuals in different clusters while DAPC succeeded in assigning the colonies in three clusters but the individual assignments differed from the one with PV7. However, when keeping only this locus, STRUCTURE did manage to assign individuals into three clusters, but the assignments differed from those using all the loci.

Finally, considering this hierarchical clustering of the colonies, the AMOVA showed that the genetic variation was explained by the partition into three SSHs (14.9%) and then by the partition into clusters within SSH (12.1%) of the genetic variation.

Discussion

Species delimitation

Gélin et al. (submitted) previously found three Secondary Species Hypotheses within the Primary Species Hypothesis PSH09 (that was attributed to the *P. eydouxi* name). In this previous study, the authors coupled several species delimitation methods applied to two mitochondrial markers and clustering analyses applied on multi-locus genotypes obtained with 13 microsatellite loci (Gélin et al., submitted). Our results confirmed the existence of three Secondary Species Hypotheses in the *P. eydouxi*, while focusing on colonies from PSH09 (the former study used all the genetic variability within the genus *Pocillopora*). Noteworthy, all the colonies belonging to these three SSHs (SSH09a, SSH09b and SSH09c) that were sequenced harbored the same ORF haplotype [ORF27; Gélin et al., (submitted)]. Nevertheless, the structuring pattern for each SSH appeared complex: SSH09a, SSH09b and SSH09c might be composed of three, two and three genetic clusters, respectively. They were all highly differentiated between each other and not corresponding to a geographic pattern. Furthermore, all these clusters were found in sympatry in each population.

Here we used microsatellites to refine boundaries in the *P. eydouxi* complex [PSH09; Gélin et al., (submitted)]. Despite the fact that microsatellites are not commonly used for species identification, their high polymorphism could help in refining boundaries between closely related species. Indeed, they are known to be short tandem repeats (typically 5-50 times) ranging in length from 2-5 base pairs. While different evolution models have been described, no consensus raised among geneticists. Moreover, they were considered to be evolutionary neutral but some studies contradict with that idea. Indeed, in humans, microsatellites could be located in gene sequences and be responsible for neurodegenerative disorders according to the number of repetitions they exhibit reviewed in Sutherland and Richards (1995). Nevertheless, they present a high mutation rate compared to other parts of the genome (Brinkmann et al., 1998) leading to a high polymorphism particularly interesting in genetic structuring studies. Indeed, their high mutation rates promote allele size homoplasy describing the fact that two species presenting a similar character are not necessarily derived from a common ancestor. This character may have arisen in two distinct species due to factors such as convergence, parallelism or reversion (*e.g.* Estoup et al., 2002; Estoup et al., 1995; Viard et al., 1998), which could make more complex interpretations on species boundaries. Nevertheless, it has been evidenced that allele size homoplasy did not represent a significant problem for many types of population genetics analyses (Estoup et al., 2002). Moreover, they are generally species-specific and developed for one species, even if some of them cross-amplify in related species, generally within the genus (Moore et al., 1991). Nevertheless, while microsatellites exhibit a high polymorphism, they remain

conserved among closely related species (Moore et al., 1991), making them potentially useful for clarifying interspecific relationships between recently diverged taxa (*e.g.* Dawson et al., 2010). Actually, we did detect some genotypic linkage disequilibrium when performing analyses over the entire dataset (mix of the three SSHs) but very few when it was estimated for each SSH separately. This suggests that particular alleles were found to not be randomly distributed when looking the entire dataset whereas they were randomly distributed inside each SSH. This suggests that (1) the number of private allele found for each SSH and (2) the variation of the allele frequencies among SSHs (Figure S3) played a role in the genotypic linkage disequilibrium observed. These results argued in favor of the existence of three distinct SSHs. Even if we cannot fully conclude in an integrative approach (microstructure, micro-environment ecology,...), it is highly probable that these three SSHs, even so the eight clusters composing the three SSHs represent distinct independent genetic lineages engaged in a speciation process or real species following the unified concept of de Queiroz (2007).

Sexual reproduction

Overall the whole sampling, we did not find any colony sharing the same MLG. Considering this, it would appear that, among our sampling, colonies belonging to these three SSHs do not exhibit clonal propagation. This seems congruent with the morphology of the colonies because these colonies are generally massive with stout branches that limits fragmentation. However, it does not rule out the fact that these colonies could be able of producing asexual larvae, but this phenomenon was not observed at the scale of our sampling. Yet, type 1 colonies (that harbored the same ORF haplotype than the PSH09) have been studied in the Tropical Eastern Pacific (TEP) and clonal propagation was revealed (Baums et al., 2014; Pinzón et al., 2012). The confusing thing in this two latter studies was the fact that the colonies were described as *P. damicornis*-like colonies, a morph that we only anecdotally attributed to our colonies belonging to this PSH. Supposedly, the colonies from these studies should likely be hybrids type 1/type 3 as revealed in (Combosch and Vollmer, 2015) or type 1/type 5 or another unrevealed cluster or independent genetic lineage found in the margins of *Pocillopora* distribution area. Thus, biological traits (*e.g.* reproduction mode) could be inherited from the nuclear genome rather than from the mitochondrial one explaining that type 1 colonies can reproduce clonally in TEP while, it seems extremely rare in the WIO and the TSP.

Fractal partitioning in PSH09

All the methods used to help in revealing population structuring in PSH09 led to a repeating pattern of clustering at different scale. Indeed, the PSH split into SSHs, the SSH into cluster, the cluster into sub-cluster, and so on...as in a fractal. It seems that the dataset kept on going dividing till each MLG would represent a genetic lineage on its own if we want to caricature this over-partitioning. Usually, using STRUCTURE, when we are at the finest level of structuring, assignments are not concluding and individuals are generally assigned to all the clusters in the same proportions. Here, some individuals kept on going assigned to clusters when K increased. Similarly, using DAPC, individuals kept to be grouped in non-overlapping clusters. Concerning STRUCTURE, this model-based approach relies on assumptions such as the type of population subdivision, which are often difficult to verify and can restrict its applicability (Jombart, 2010). So we chose arbitrarily to consider only the clustering that was retrieved by different methods of analyses following a principle of parsimony. Thus, all SSHs were subdivided in several clusters (at least two), which were all found in sympatry at the island scale. Let assuming that it is not a mathematical artefact. So we could imagine that each SSH represent a complex of species (more or less cryptic, complementary studies on microstructure are needed to fully conclude) or a mixing of different lineages within the same species, or something half-way in the 'grey zone' between two populations highly differentiated and two species highly divergent. In whichever case we are, the processes leading to genetic differentiation should be the same.

First, all our provinces, each SSH partitioning might be the result of a systematic sampling of the sink populations and lack of sampling of the source populations. Colonies from source populations are supposed to be assigned in one unique cluster each and to broadcast their propagules through the ocean till sink populations, which could explain the pattern we observed. Moreover, this hypothesis suggests that, whatever the propagules origin is, colonies belonging to one cluster do not reproduce with colonies from other clusters, otherwise we should observe more genetically homogeneous populations. Thus, the observed pattern might result from a high migration abilities of the highly dispersive propagules of several source populations without reproduction among colonies of these different source populations that allowed us to identify a mix pattern of geographical origins in each sampled locality. This implies that some reproductive barriers must exist among the clusters.

Second, another hypothesis rejoins the previous one, but no longer considers source populations but clonal lineages. Reproductively isolated clones (pre- or post-zygotic barriers impeding reproduction with closely related individuals) would have spread and diversified resulting in the

different clusters observed and in the *ad libitum* partitioning. Nevertheless, as we did not reveal clonal reproduction in our sampling, it seems unlikely. However, clonal propagation has been evidenced in the Tropical Eastern Pacific (TEP) (Baums et al., 2014; Pinzón et al., 2012) for ORF type 1 colonies (*sensu* Pinzón et al., 2012). It could represent an ancestral character in TEP which should have been lost in other localities, a hypothesis suggested by previous studies hypothesizing character reacquisition in *Pocillopora* corals. These studies mentioned particularly the possibility that broadcast-spawning was derived from brooding in this genus (Kerr et al., 2010; Pinzón et al., 2013).

The reproductive isolation evoked in these two hypotheses would have resulted from an allopatric differentiation among populations. Moreover, it had been evidenced that in populations that experienced a reduction of gene flow (such as in allopatric speciation processes), the populations could exhibit heterogeneity in genetic differentiation along genomes (Tine et al., 2014). Indeed, studying two sea bass populations by screening their genomes, one inhabiting the north-eastern Atlantic Ocean and the other, the Mediterranean and the Black seas, Tine et al. (2014) revealed that, among the populations, individuals exhibited some parts of the genome that were very differentiated and other parts less differentiated. This heterogeneity could lead to consider these two populations as two distinct species when looking at the highly differentiated zones of the genome. Thus, if the loci used were located in these zones, it would lead to explain the observed differentiation by the existence of two cryptic species. To date, nothing is known about *Pocillopora* nuclear genomes, even the number of chromosomes [supposed to be between 14 and 40, as described for *Favia pallida* and *Acropora nasuta*, respectively; (Flot et al., 2006)]. Thus, we ignore where the microsatellite loci we used are located [except PV7 which is known to be in the ITS1 [HM pers. com. and Magalon et al., (2004)] and whether they could be located in highly differentiated zones of the genome, which could reveal a “false” genetic differentiation among the clusters we observed. Nevertheless, this idea suggests, as in Tine et al. (2014), that ancestral populations experienced a gene flow reduction in the past. However, looking at the allelic distributions of each locus among clusters, no particular locus seems an outlier, except PV7 in SSH09a. As already said, it is located in the ITS1, which is a region of the genome supposedly not under selection while flanked with two rDNA genes (18S and 5.8S) that code for ribosomal RNAs. But, when keeping only this locus for the clustering analyses, the individual assignments differed from keeping all the loci, thus it suggests that PV7 alone cannot explain the partitioning observed and that this latter should result from the combination of the evolutionary history of each locus.

Another way to explain the observed clusters could be the habitat selection by individuals that could have specialized in habitat types. Habitat selection have already been highlighted as a speciation factor. Indeed, for European anchovy populations, the habitat type (marine vs. coastal) accounts for most of the genetic structure (Le Moan et al., 2016). In this latter case, the genomic approach strongly supported a model of ecotypic divergence shaped by recent differential gene flow after a period of complete isolation (Le Moan et al., 2016). For corals, micro-habitat selection could be the result of variability in light, current exposure, salinity, food availability and/or temperature. Summarizing different habitat particularities, depth could be a factor of population structuring according to the habitat type. Depth structuring have been evidenced for *Seriatopora hystrix* in Australia (Van Oppen et al., 2011), where genetic differentiation among sites from different depths was greater than differentiation among sites within the same depths (over 50 km). However, colonies from one site were collected during a single dive and we tried to stay at the same depth (or from deeper to shallower in the range) while collecting, and this for all sites (*i.e.* sampling depth between 14 and 8 m), except the population from TRO1 where colonies were collected at 24 m. Moreover, when looking at the distribution of the different clusters among sites, particularly for SSH09a, they were not segregated in space within a site (linear trajectory while sampling), nor within island, nor within ecoregion, nor in time (as islands were visited at different dates). All in all, depth should not be the explanatory factor of genetic differentiation in our case, but micro-habitat might be, considering fine-scale variations of both abiotic and biotic factors that could be indiscernible to the human eye. This hypothesis deserves to be tested with full attention.

In addition to speciation and selection, hybridization could also explain the odd clustering pattern, each cluster being the result of different mixes between two sufficiently different genetic entities that were not sampled or ancestral. Indeed, hybridization is not uncommon in corals (*e.g.* Flot et al., 2011; Frade et al., 2010; Isomura et al., 2013; Márquez et al., 2002; Vollmer and Palumbi, 2002). As an example, *Acropora prolifera* has been evidenced to be a hybrid (and not a hybrid species, see Willis et al., 2006) between *Acropora cervicornis* and *Acropora palmata* (van Oppen et al., 2000). In this way, (Combosch and Vollmer, 2015) revealed hybridization between *Pocillopora* type 1 and type 3 using RADseq. Moreover, each time hybridization has been demonstrated in *Pocillopora* corals (Combosch et al., 2008; Combosch and Vollmer, 2015; Pinzón and LaJeunesse, 2011), it seemed dominated by type 1 maternal lineages since all hybrid colonies exhibited mitochondrial type 1 haplotypes (Combosch and Vollmer, 2015). Frequent event of hybridization could comfort Veron's suggestion about meta-species (or syngameon) existence in corals (Veron, 1995). Whichever the causes, PSH09 (even the whole *Pocillopora*

genus) may present a meta-species with some hybrids between entities (such as the individuals that were assigned to two different clusters herein).

To date, we cannot favor one or another hypothesis and more investigations are needed to fully conclude. Several hypotheses imply an ancestral reduction of gene flow that created reproductive barriers or genome incompatibilities among the different clusters for each SSH, which contemporary gene flows have not homogenized.

Is it possible to speak about connectivity?

Facing to this over-partitioning of our dataset, is it possible to speak about connectivity? The problem is not how to calculate connectivity but on what. Similarly, the problem is not how to define a species, but where to put the limits between two species. First let assume that each SSH represents a genetic species (with no gene flow by definition). The advantage (even if it could resemble to a *reductio ad absurdum*) is that it allowed to keep all the individuals of the sampling and to preserve the hierarchical sampling scheme. In this case, no gene flow (SSH09a restricted to the WIO) or extremely rare gene flow (for SSH09b and SSH09c almost exclusively in the TSP) was observed between the Pacific Ocean and the Western Indian Ocean. Indeed, no individuals from SSH09a were retrieved in the TSP and few individuals (less than 1% and 2% from SSH09b and SSH09c, respectively) found in the WIO. Obviously, the gene flow (if efficient as the presence of colonies does not imply gene exchange with conspecifics and if any as we cannot exclude scoring or assignment error) may be unidirectional from East to West and very weak. When looking at each SSH separately, populations showed pattern of differentiation at local scale (i.e., island scale, the more distant populations in the same island were 100 km apart in Reunion Island), and also at regional scale (within the province of the WIO). However, at both scales, some populations remain highly connected. The high F_{ST} variance for each SSH may reflect the different clusters found in different proportions at the inferior hierarchical level of structuring. So we might prefer to work on the lower level of structuring as referring to one homogenous genetic entity. So our choice criteria to define the “true” number of clusters are arbitrary but are intended to be objective (more or less) and repeatable over different methods. That being said, let assume that the species boundaries are among the different clusters found within each SSH and let consider the eight different clusters as our reference unit to assess connectivity. Consequently, this approach presents the strong drawback to fragment our sampling into pieces. Then the sample size of each population for a given cluster is drastically reduced, implying to pool together individuals from different populations within the same island,

removing the lowest spatial structuring level (local scale). Noteworthy, the pairwise F_{ST} values among clusters were of the same order of magnitude (even if a little bit less high) than among SSHs. Considering the WIO, the three clusters found there showed that the populations seem highly connected at the scale of the WIO, only Madagascar and Tromelin (but few individuals) seem to present a particular position in the region. Indeed, populations from Madagascar are weakly differentiated from Mayotte, Glorioso Islands (except for SSH09a-2), Juan de Nova Island (all situated in the north part of the Mozambique Channel) and Reunion Island (only for SSH09a-3), situated east of Madagascar. The populations from Madagascar are the southernmost of the sampling and may be differentiated with these populations just by geographic isolation and current eddies present in the Mozambique Channel. Moreover, these reefs are subjected to high anthropogenic pressures (fisheries, pollution) that may have impacted their genetic diversity. In the TSP, examining pattern of genetic differentiation within SSH09b-1 (unfortunately, SSH09b-2 presented too few individuals), the South-East Coast of New Caledonia was differentiated from all the other populations from Chesterfield Islands, New Caledonia and Loyalty Islands. Moreover, Chesterfield Islands were differentiated from all others populations, except the west of Caledonia and Lifou (situated east of New Caledonia). The currents should be studied more precisely or local conditions to explain this pattern. However, SSH09b tended to be cut in an infinity of clusters and it is the mixing of the different clusters that could lead to this pattern of differentiation, preventing to conclude about population connectivity in the genetic lineage.

Lastly, considering SSH09c-1 which is restricted to Chesterfield Islands, the pattern seems quite clear: population CHE02, the northernmost population situated in Bampton reefs is highly differentiated to all the other populations. This site presents deep vertical cliffs and high currents from east to west that may flush out any propagule to the west. Looking at the second cluster present in Chesterfield Islands and also Loyalty Islands (SSH09c-3), CHE02 is found differentiated to all the other populations, confirming its isolated position and LOY4 (Lifou) is differentiated from all the populations even from LOY5 (Ouvéa, Loyalty Island). Finally, examining the cluster found only in New Caledonia (allowing to not distort the sampling), no general trend can be revealed, some pairs of populations being differentiated. Noteworthy it is one of the only clusters that did not keep on dividing, suggesting that it represents a homogenous genetic lineage.

Conclusion and perspectives

Examining the population structuring of the *Pocillopora eydouxi* species hypothesis revealed a beautiful puzzle to solve, with a fractal partitioning of the different SSHs, obliging to think what is the unit on which connectivity should be assessed. Several hypotheses have been proposed to explain the observed pattern but more investigations are needed to understand the structuring pattern of this species. Knowing more about *Pocillopora* genome and microsatellites positions seem now to be a fundamental key to improve the understanding of this species hypothesis. Moreover, a genomic approach would help in understanding this species history of divergence and would offer clues to favor one or another exposed hypothesis.

Acknowledgments

Corals sampling in New Caledonia (HM) was carried out during COBELO (doi: <http://dx.doi.org/10.17600/14003700>), BIBELOT (doi: <http://dx.doi.org/10.17600/13100100>) and CHEST (<http://dx.doi.org/10.17600/15004500>) oceanographic campaigns on board of RV Alis (IRD). Sampling in Reunion Island (HM, PG) was supported by program CONPOCINPA (LabEx CORAIL fund); in Madagascar (HM) and in Rodrigues Island (HM) supported by project Biodiversity (POCT FEDER fund); in Europa, Juan de Nova and Glorioso Islands (HM) by program BIORECIE (financial supports from INEE, INSU, IRD, AAMP, FRB, TAAF, and the foundation Veolia Environnement); in Tromelin Island (HM) by program ORCIE (INEE), and in Mayotte (HM) by program SIREME (FED). HM thanks all the buddies who helped in photographs during diving (J. Butscher, S. Andréfouët, L. Bigot, M. Pinault). We acknowledge the Plateforme Gentyane of the Institut National de la Recherche Agronomique (INRA, Clermont-Ferrand, France) for guidance and technical support. PG was financially supported by a PhD contract from the LabEx CORAIL.

Supplementary material

Figure S1. Distribution of the allelic frequencies for each loci presented for each SSH separately

Figure S2. Results of the assignment tests for SSH09b. At $K = 2$ for SSH09b (a), each of the two clusters (SSH09b-1 and SSH09b-2) were run separately. (b) $\Delta(K)$ and BIC values and (c) STRUCTURE plots for $K = 2$ to $K = 4$ are presented both for SSH09b-1 and SSH09b-2.

Figure S3. Distribution of the allelic frequencies for SSH09a. (a) Allelic frequencies for each locus for the three identified clusters (SSH09a-1, SSH09a-2, SSH09a-3) and (b) Weir & Cockerham (1984) F_{ST} estimated per locus overall colonies of SSH09a.

Table S1. Genetic differentiation between *Pocillopora eydouxi* populations for each SSH estimated with Weir and Cockerham's F_{ST} (Weir and Cockerham, 1984). (a) SSH09a, (b) SSH09b) and (c) SSH09c. ***: $P < 0.001$

Table S2. Pairwise F_{ST} values estimated between populations within each cluster at $K = 3$ within SSH09c [(a) SSH09c-1, (b) SSH09c-2 and (c) SSH09c-3]. For each cluster, the number of colonies per population is indicated in parentheses. *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$

Table S3. Pairwise F_{ST} values estimated between populations within for SSH09b-1. The number of colonies per population is indicated in parentheses. *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$

Table S4. Pairwise F_{ST} values estimated between populations within each cluster at $K = 3$ within SSH09a [(a) SSH09a-1, (b) SSH09a-2 and (c) SSH09a-3]. For each cluster, the number of colonies per population is indicated in parentheses. *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$

References

- Ayre, D.J., Hughes, T.P., 2000. Genotypic diversity and gene flow in the brooding and spawning corals along the Great Barrier Reef, Australia. *Evolution* 54, 1590-1605.
- Bailleul, D., Stoeckel, S., Arnaud-Haond, S., 2016. RClone: a package to identify multiLocus clonal lineages and handle clonal datasets in R. *Methods in Ecology and Evolution*.
- Baums, I.B., Devlin-Durante, M., Laing, B.A.A., Feingold, J., Smith, T., Bruckner, A., Monteiro, J., 2014. Marginal coral populations: the densest known aggregation of *Pocillopora* in the Galápagos Archipelago is of asexual origin. *Frontiers in Marine Science*.
- Bohonak, A.J., 1999. Dispersal, gene flow and population structure. *Q Rev Biol* 74, 21-44.
- Brinkmann, B., Klitschar, M., Neuhuber, F., Hühne, J., Rolf, B., 1998. Mutation rate in human microsatellites: influence of the structure and length of the tandem repeat. *Am J Hum Genet* 62, 1408-1415.
- 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.
- Dawson, D.A., Horsburgh, G.J., KÜpper, C., Stewart, I.R.K., Ball, A.D., Durrant, K.L., Hansson, B., Bacon, I.D.A., Bird, S., Klein, Á., Krupa, A.P., Lee, J.-W., Martín-GÁLvez, D., Simeoni, M., Smith, G., Spurgin, L.G., Burke, T., 2010. New methods to identify conserved microsatellite loci and develop primer sets of high cross-species utility – as demonstrated for birds. *Mol Ecol Resour* 10, 475-494.
- Estoup, A., Jarne, P., Cornuet, J.-M., 2002. Homoplasy and mutation model at microsatellite loci and their consequences for population genetic analysis. *Mol Ecol* 11, 1591-1604.

- Estoup, A., Tailliez, C., Cornuet, J.M., Solignac, M., 1995. Size homoplasy and mutational processes of interrupted microsatellites in two bee species, *Apis mellifera* and *Bombus terrestris* (Apidae). *Mol Biol Evol* 12, 1074-1084.
- Flot, J.-F., Ozouf-Costaz, C., Tsuchiya, M., van Woesik, R., 2006. Comparative coral cytogenetics. *Proc. 10th Int. Coral Reef Symp*, pp. 4-8.
- Flot, J.F., Blanchot, J., Charpy, L., Cruaud, C., Licuanan, W.Y., Nakano, Y., Payri, C., Tillier, S., 2011. Incongruence between morphotypes and genetically delimited species in the coral genus *Stylophora*: phenotypic plasticity, morphological convergence, morphological stasis or interspecific hybridization? *BMC Ecol* 11, 22.
- Flot, J.F., Magalon, H., Cruaud, C., Couloux, A., Tillier, S., 2008. Patterns of genetic structure among Hawaiian corals of the genus *Pocillopora* yield clusters of individuals that are compatible with morphology. *CR Biol* 331, 239-247.
- Frade, P.R., Reyes-Nivia, M.C., Faria, J., Kaandorp, J.A., Luttikhuizen, P.C., Bak, R.P., 2010. Semi-permeable species boundaries in the coral genus *Madracis*: introgression in a brooding coral system. *Mol Phylogenet Evol* 57, 1072-1090.
- Gélin, P., Postaire, B., Fauvelot, C., Hélène, M., submitted. Reevaluating species number, distribution and endemism of the coral genus *Pocillopora* Lamarck, 1816 using species delimitation methods and microsatellites.
- Gorospe, K.D., Karl, S.A., 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.
- Goudet, J., 2001. FSTAT, a program to estimate and test gene diversities and fixation indices version 2.9. 3.2, updated from Goudet 1995. FSTAT.
- Graham, E.M., Baird, A.H., Connolly, S.R., 2008. Survival dynamics of scleractinian coral larvae and implications for dispersal. *Coral Reefs* 27, 529-539.
- 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.
- Hirose, M., Kinzie III, R.A., Hidaka, M., 2001. Timing and process of entry of zooxanthellae into oocytes of hermetic corals. *Coral Reef* 20, 273-280.
- Isomura, N., Iwao, K., Fukami, H., 2013. Possible natural hybridization of two morphologically distinct species of *Acropora* (Cnidaria, Scleractinia) in the Pacific: fertilization and larval survival rates. *PLoS ONE* 8, e56701.
- Jombart, T., 2008. *ade4*: 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.
- Kerr, A.M., Baird, A.H., Hughes, T.P., 2010. Correlated evolution of sex and reproductive mode in corals (Anthozoa: Scleractinia). *Proc Roy Soc B: Biol Sci* 278, 75.
- Le Moan, A., Gagnaire, P.A., Bonhomme, F., 2016. Parallel genetic divergence among coastal-marine ecotype pairs of European anchovy explained by differential introgression after secondary contact. *Mol Ecol* 25, 3187-3202.
- 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.
- Márquez, L.M., Van Oppen, M.J., Willis, B., Reyes, A., Miller, D.J., 2002. The highly cross-fertile coral species, *Acropora hyacinthus* and *Acropora cytherea*, constitute statistically distinguishable lineages. *Mol Ecol* 11, 1339-1349.
- Marti-Puig, P., Forsman, Z.H., Haverkort-Yeh, R.D., Knapp, I.S.S., Maragos, J.E., Toonen, R.J., 2014. Extreme phenotypic polymorphism in the coral genus *Pocillopora*; micro-morphology

- corresponds to mitochondrial groups, while colony morphology does not. *Bull Mar Sci* 90, 211-231.
- Moore, S.S., Sargeant, L.L., King, T.J., Mattick, J.S., Georges, M., Hetzel, D.J.S., 1991. The conservation of dinucleotide microsatellites among mammalian genomes allows the use of heterologous PCR primer pairs in closely related species. *Genomics* 10, 654-660.
- Palumbi, S.R., 1994. Genetic divergence, reproductive isolation, and marine speciation. *An Rev Ecol Syst* 25, 547-572.
- 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.
- 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., Reyes-Bonilla, H., Baums, I.B., LaJeunesse, T.C., 2012. Contrasting clonal structure among *Pocillopora* (Scleractinia) communities at two environmentally distinct sites in the Gulf of California. *Coral Reefs* 31, 765-777.
- Pinzón, J.H., Sampayo, E., Cox, E., Chauka, L.J., Chen, C.A., Voolstra, C.R., LaJeunesse, T.C., Maggs, 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.
- Postaire, B., Aurelle, D., Bourmaud, C.A.F., Bruggemann, J.H., Magalon, H., 2015. Isolation and characterisation of 26 microsatellite loci from a widespread tropical hydrozoan, *Macrorhynchia phoenicea* (Leptothecata, Aglaopheniidae), and cross-amplification in closely related species. *Biochem Syst Ecol* 62, 137-141.
- Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 954-959.
- Richmond, R.H., 1987. Energetics, competency, and long-distance dispersal of planula larvae of the coral. *Mar Biol* 93, 527-533.
- Schmidt-Roach, S., Miller, K.J., 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.
- Schneider, S., Roessli, D., Excoffier, L., 2000. Arlequin ver. 2.000: A software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.
- Selkoe, K.A., D'Aloia, C.C., Crandall, E.D., Iacchei, M., Liggins, L., Puritz, J.B., von der Heyden, S., Toonen, R.J., 2016. A decade of seascape genetics: contributions to basic and applied marine connectivity. *Mar Ecol Progr Ser* 554, 1-19.
- Siegel, D.A., Mitarai, S., Costello, C.J., Gaines, S.D., Kendall, B.E., Warner, R.R., Winters, K.B., 2008. The stochastic nature of larval connectivity among nearshore marine populations. *PNAS* 105, 8974-8979.
- Spalding, M.D., Fox, H.E., Allen, G.R., Davidson, N., Ferdaña, Z.A., Finlayson, M., Halpern, B.S., Jorge, M.A., Lombana, A., Lourie, S.A., Martin, K.D., McManus, E., Molnar, J., Recchia, C.A., Robertson, J., 2007. Marine ecoregions of the world: A bioregionalization of coastal and shelf areas. *BioScience* 57, 573-583.
- Starger, C.J., Yeoh, S.S., Dai, C.F., Baker, A.C., Desalle, R., 2007. Ten polymorphic STR loci in the cosmopolitan reef coral, *Pocillopora damicornis*. *Mol Ecol Resour* 8, 619-621.
- Sutherland, G.R., Richards, R.I., 1995. Simple tandem DNA repeats and human genetic disease. *PNAS* 92, 3636-3641.
- Tine, M., Kuhl, H., Gagnaire, P.-A., Louro, B., Desmarais, E., Martins, R.S.T., Hecht, J., Knaust, F., Belkhir, K., Klages, S., Dieterich, R., Stueber, K., Piferrer, F., Guinand, B., Bierne, N., Volckaert, F.A.M., Bargelloni, L., Power, D.M., Bonhomme, F., Canario, A.V.M., Reinhardt,

- R., 2014. European sea bass genome and its variation provide insights into adaptation to euryhalinity and speciation. *Nat Com* 5, 5770.
- Torda, G., Lundgren, P., Willis, B.L., van Oppen, M.J.H., 2013a. Genetic assignment of recruits reveals short and long distance larval dispersal in *Pocillopora damicornis* on the Great Barrier Reef. *Mol Ecol* 22, 5821–5834.
- Torda, G., Schmidt-Roach, S., Peplow, L.M., Lundgren, P., van Oppen, M.J.H., 2013b. A rapid genetic assay for the identification of the most common *Pocillopora damicornis* genetic lineages on the great barrier reef. *Plos One* 8, e58447.
- van Oppen, M.J., Willis, B.L., Vugt, H.W., Miller, D.J., 2000. Examination of species boundaries in the *Acropora cervicornis* group (Scleractinia, cnidaria) using nuclear DNA sequence analyses. *Mol Ecol* 9, 1363-1373.
- Van Oppen, M.J.H., Bongaerts, P.I.M., Underwood, J.N., Peplow, L.M., Cooper, T.F., 2011. The role of deep reefs in shallow reef recovery: an assessment of vertical connectivity in a brooding coral from west and east Australia. *Mol Ecol* 20, 1647-1660.
- Veron, J.E.N., 1995. Corals in space and time: the biogeography and evolution of the Scleractinia. Comstock/Cornell.
- Viard, F., Franck, P., Dubois, M.-P., Estoup, A., Jarne, P., 1998. eVariation of microsatellite size homoplasy across electromorphs, loci, and populations in three invertebrate species. *J Mol Evol* 47, 42-51.
- Vollmer, S.V., Palumbi, S.R., 2002. Hybridization and the evolution of reef coral diversity. *Science* 296, 2023-2025.
- Weir, B., Cockerham, C., 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358-1370.
- Willis, B.L., van Oppen, M.J.H., Miller, D.J., Vollmer, S.V., Ayre, D.J., 2006. The Role of Hybridization in the Evolution of Reef Corals. *An Rev Ecol Evol Syst* 37, 489-517.

Figure S1

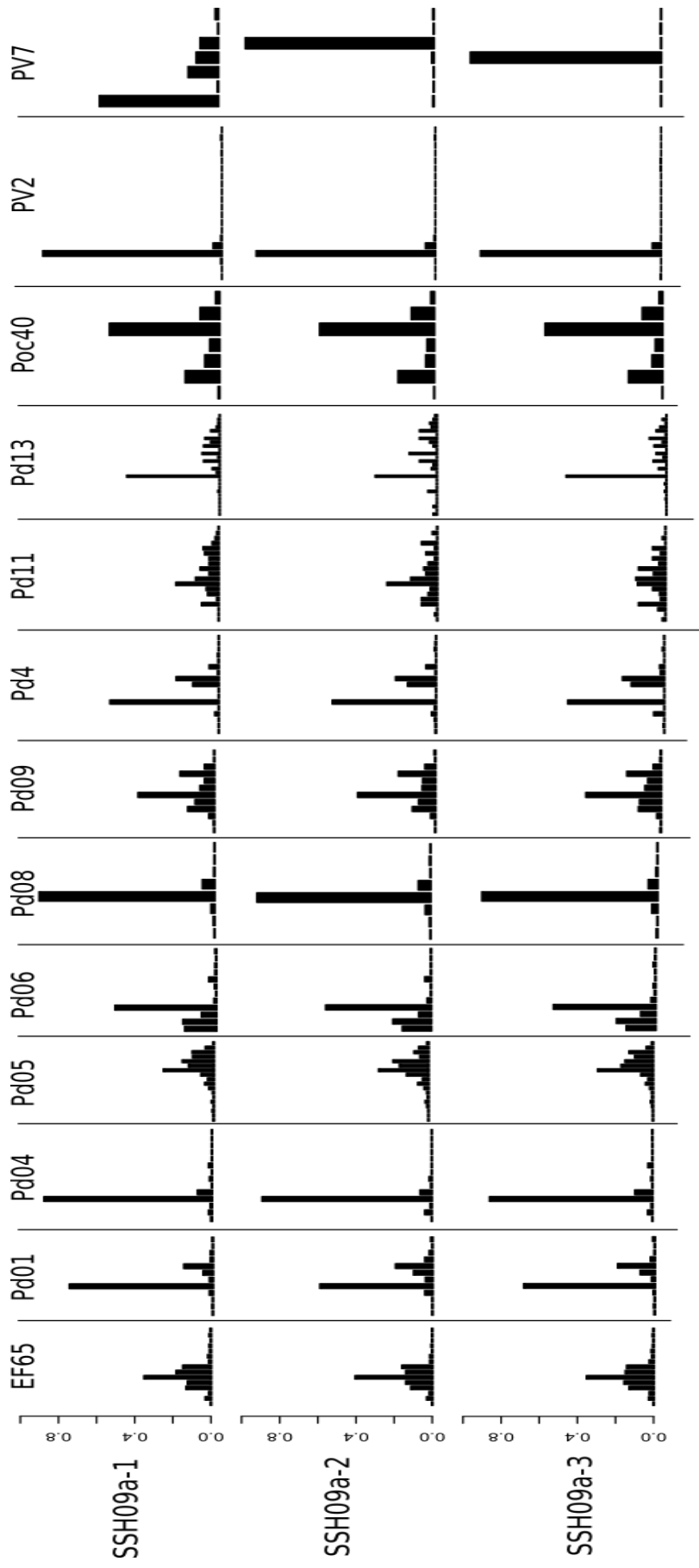


Figure S2

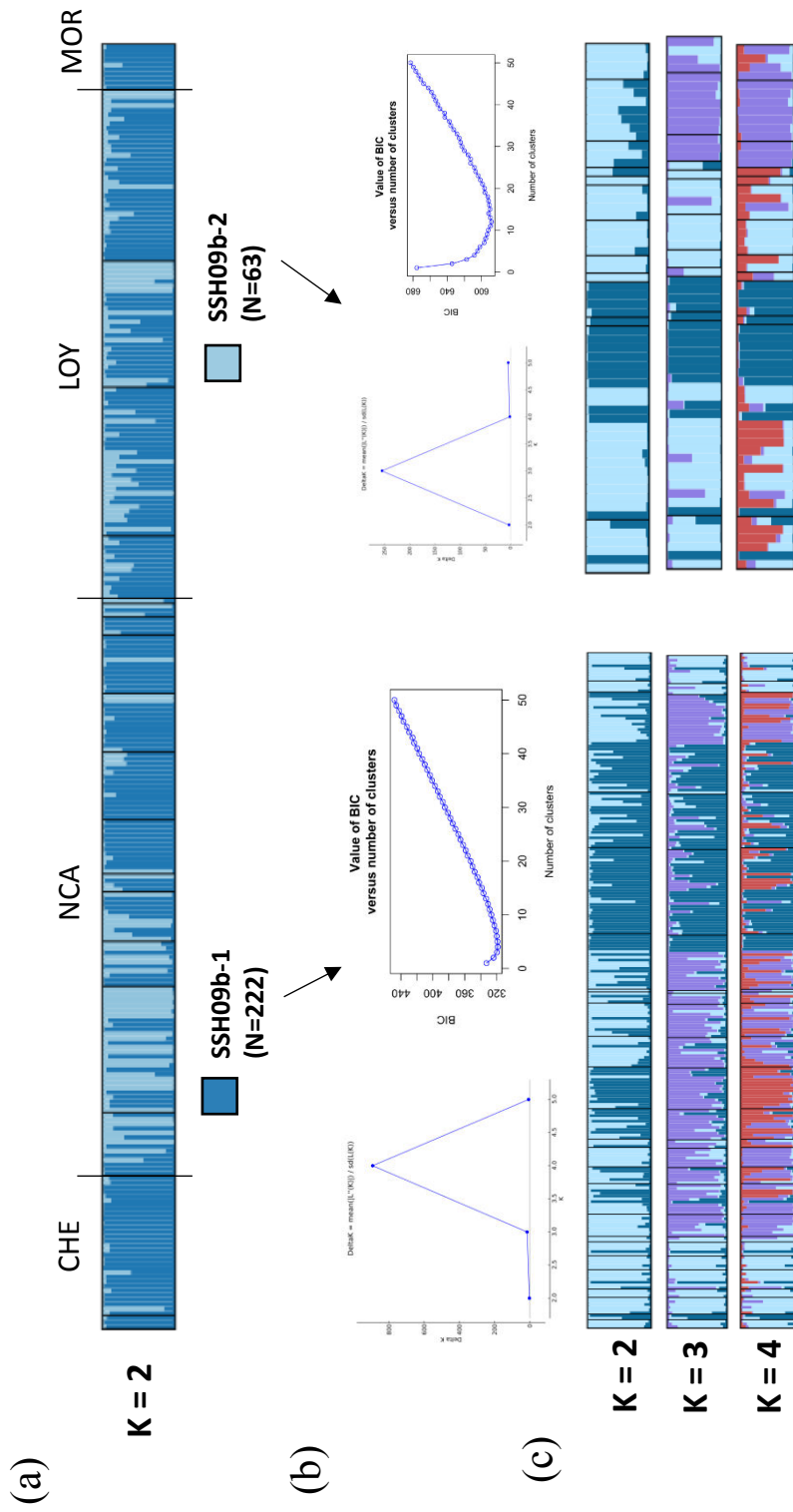


Figure S3

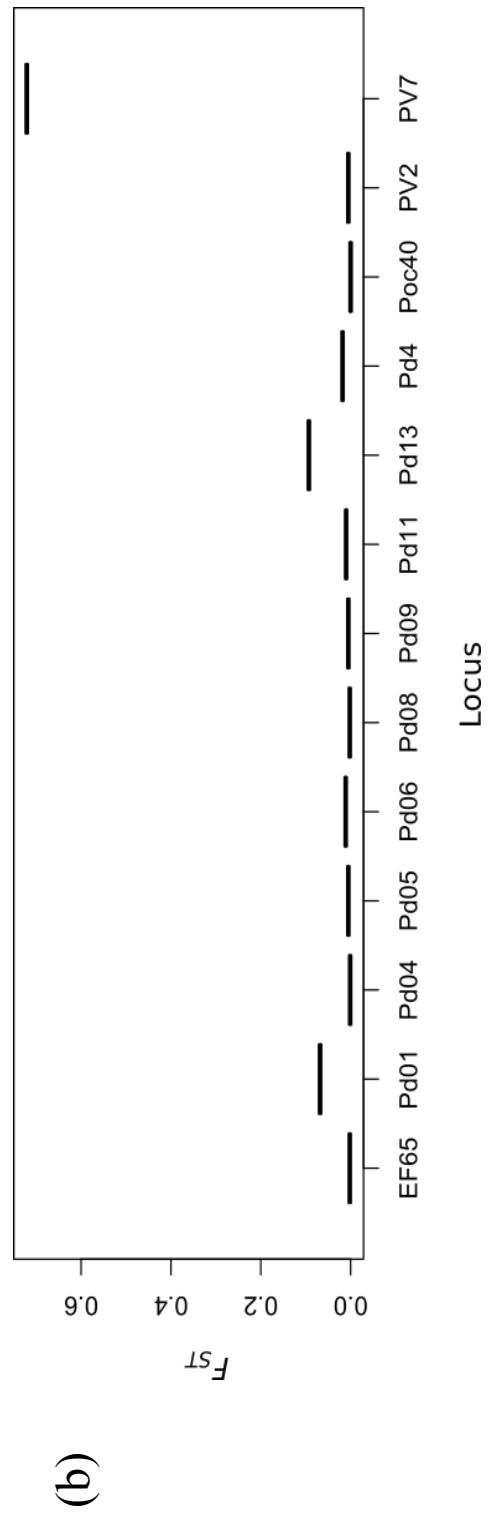
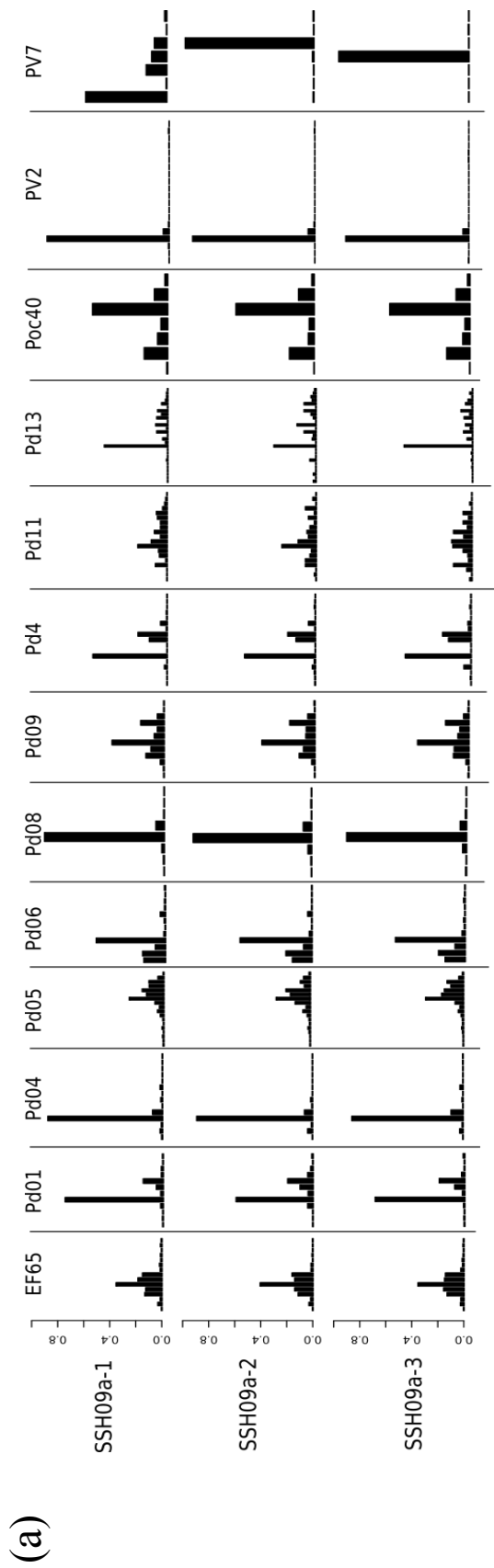


Table S1. *Pocillopora eydouxi* pairwise F_{ST} values for all pairs of populations for each SSH. (a) SSH09a, (b) SSH09b and (c) SSH09c. Values significantly different from 0 were grayed out. * $P < 0.05$ *** $P < 0.01$ **** $P < 0.001$.

(a)

	MAY1	MAY2	MAY3	GLO1	GLO2	GLO3	GLO4	JDN1	JDN2	JDN3	JDN4	JDN5	BAS1	EUR1	EUR2	EUR3	EUR4	MAD1	MAD2	MAD3	REU1	REU2	REU3	REU3b	REU4	REU5	ROD1	ROD2	ROD3	ROD4	TRO1	
MAY2	0.020**	-																														
MAY3	0.009	0.011	-																													
GLO1	0.001	0.014	0.009	-																												
GLO2	0.000	0.004	0.002	0.003	-																											
GLO3	0.011	0.013	0.016*	0.012	0.005	-																										
GLO4	0.004	0.005	0.004	0.003	0.000	0.007	-																									
JDN1	0.000	0.016*	0.000	0.006	0.000	0.013*	0.004	-																								
JDN2	0.004	0.014*	0.000	0.005	0.000	0.011	0.005	0.000	-																							
JDN3	0.002	0.012	0.013*	0.002	0.000	0.005	0.005	0.007	0.006	-																						
JDN4	0.019*	0.001	0.012	0.020*	0.002	0.006	0.008	0.015*	0.014*	0.011	-																					
JDN5	0.007	0.005	0.009	0.005	0.000	0.010	0.003	0.006	0.004	0.004	0.008	-																				
BAS1	0.026**	0.010	0.026**	0.012	0.004	0.022**	0.013	0.024**	0.024**	0.015*	0.015	0.005	-																			
EUR1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-																		
EUR2	0.003	0.010	0.004	0.000	0.000	0.001	0.002	0.001	0.005	0.006	0.009	0.000	0.011	0.000	-																	
EUR3	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-																
EUR4	0.032*	0.026	0.049**	0.018	0.028	0.004	0.028	0.041**	0.045**	0.016	0.022	0.017	0.034*	0.000	0.017	0.014	-															
MAD1	0.022***	0.022***	0.031***	0.026***	0.016*	0.007	0.018*	0.024***	0.021***	0.015**	0.019**	0.020**	0.038***	0.000	0.015**	0.000	0.013	-														
MAD2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	-													
MAD3	0.026***	0.021**	0.034***	0.023**	0.013	0.002	0.023*	0.028***	0.027***	0.012*	0.019**	0.017*	0.030***	0.000	0.013*	0.000	0.007	0.000	0.000	-												
REU1	0.015*	0.000	0.007	0.012	0.002	0.013*	0.003	0.011*	0.011*	0.011*	0.006	0.004	0.020**	0.000	0.006	0.000	0.024	0.018***	0.001	0.019**	-											
REU2	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-											
REU3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-										
REU3b	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-									
REU4	0.000	0.025**	0.008	0.014	0.005	0.018*	0.010	0.000	0.010	0.014*	0.022**	0.011	0.030**	0.000	0.013	0.001	0.043**	0.030***	0.000	0.036***	0.015*	0.000	0.000	0.000	0.000	-						
REU5	0.000	0.005	0.000	0.000	0.000	0.006	0.000	0.000	0.000	0.000	0.013*	0.000	0.009	0.004	0.000	0.013	0.026	0.010*	0.000	0.019**	0.004	0.000	0.000	0.000	0.000	0.000	-					
ROD1	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-					
ROD2	0.016*	0.020**	0.029**	0.002	0.000	0.000	0.003	0.009*	0.013*	0.007	0.012	0.006	0.021**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.009	0.000	0.000	0.000	0.013*	0.000	0.000	-				
ROD3	0.012	0.001	0.027**	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.009	0.000	0.000	0.005	-			
ROD4	0.027***	0.032***	0.035***	0.015*	0.015	0.005	0.012	0.015**	0.019**	0.015*	0.024**	0.011	0.036***	0.000	0.008	0.009	0.003	0.002	0.000	0.008	0.020**	0.000	0.000	0.000	0.017*	0.000	0.000	0.009	-			
TRO1	0.029*	0.011	0.023*	0.020	0.009	0.036**	0.005	0.022*	0.022*	0.020*	0.019	0.020	0.019	0.015	0.016	0.003	0.054***	0.032**	0.002	0.040**	0.011	0.000	0.000	0.000	0.030**	0.010	0.011	0.031**	0.008	0.040**	-	
TRO2	0.058***	0.079***	0.048***	0.068***	0.078***	0.082***	0.054**	0.064***	0.059***	0.080***	0.069***	0.093***	0.107***	0.040**	0.062***	0.047*	0.116***	0.096***	0.081***	0.109***	0.076***	0.036*	0.021	0.025	0.058***	0.071***	0.072***	0.100***	0.129***	0.116***	0.075***	

(b)

	NCA01	NCA02	NCA06	NCA07	NCA08	NCA09	NCA10	LOY3	LOY4	LOY5	LOY6
NCA01	-										
NCA02	0.047*	-									
NCA06	0.072*	0.069**	-								
NCA07	0.120***	0.048*	0.009	-							
NCA08	0.021	0.001	0.004	0.011	-						
NCA09	0.066*	0.009	0.000	0.000	0.004	-					
NCA10	0.079	0.000	0.000	0.000	0.000	0.000	-				
LOY3	0.114***	0.086***	0.020	0.077***	0.073**	0.041*	0.000	-			
LOY4	0.101***	0.084***	0.015	0.000	0.021	0.000	0.023	0.064***	-		
LOY5	0.084***	0.027*	0.017	0.014	0.016	0.001	0.000	0.049**	0.029**	-	
LOY6	0.056**	0.050***	0.000	0.000	0.000	0.000	0.000	0.018	0.010	0.013	-

(c)

	CHE01	CHE02	CHE03	CHE04	CHE05	CHE06	CHE07	CHE08	CHE09	CHE10	NCA01	NCA02	NCA03	NCA04	NCA05	NCA06	NCA07	NCA08	NCA09	NCA10	LOY1	LOY2	LOY4	
CHE02	0.058***	-																						
CHE03	0.013	0.032**	-																					
CHE04	0.020*	0.074***	0.016	-																				
CHE05	0.004	0.053***	0.007	0.008	-																			
CHE06	0.000	0.061***	0.012	0.024**	0.010	-																		
CHE07	0.008	0.065***	0.006	0.018*	0.006	0.005	-																	
CHE08	0.004	0.069***	0.010	0.012	0.006	0.006	0.000	-																
CHE09	0.009	0.052***	0.002	0.007	0.005	0.009	0.005	0.003	-															
CHE10	0.005	0.052***	0.000	0.011	0.002	0.006	0.006	0.006	0.010	-														
NCA01	0.248***	0.294***	0.261***	0.308***	0.279***	0.249***	0.270***	0.288***	0.276***	0.247***	-													
NCA02	0.144***	0.189***	0.179***	0.221***	0.189***	0.156***	0.187***	0.192***	0.184***	0.154***	0.019	-												
NCA03	0.259***	0.285***	0.255***	0.308***	0.287***	0.262***	0.277***	0.287***	0.275***	0.258***	0.030	0.000	-											
NCA04	0.269***	0.283***	0.262***	0.324***	0.295***	0.275***	0.297***	0.304***	0.288***	0.271***	0.051*	0.000	0.026	-										
NCA05	0.295***	0.320***	0.293***	0.349***	0.316***	0.307***	0.326***	0.329***	0.320***	0.299***	0.090***	0.018	0.066***	0.041**	-									
NCA06	0.274***	0.303***	0.273***	0.332***	0.299***	0.287***	0.301***	0.309***	0.300***	0.278***	0.000	0.000	0.000	0.000	0.000	-								
NCA07	0.246***	0.274***	0.243***	0.311***	0.269***	0.264***	0.273***	0.287***	0.269***	0.246***	0.008	0.000	0.000	0.000	0.000	0.000	-							
NCA08	0.259***	0.290***	0.257***	0.317***	0.285***	0.274***	0.288***	0.297***	0.284***	0.264***	0.055**	0.000	0.011	0.017	0.021**	0.000	0.000	-						
NCA09	0.249***	0.274***	0.243***	0.299***	0.276***	0.255***	0.272***	0.275***	0.266***	0.251***	0.038*	0.000	0.000	0.013	0.053***	0.000	0.000	0.000	-					
NCA10	0.260***	0.282***	0.247***	0.299***	0.280***	0.264***	0.277***	0.287***	0.271***	0.261***	0.065**	0.041	0.014	0.046**	0.074***	0.000	0.000	0.005	0.015	-				
LOY1	0.215**	0.244**	0.211**	0.271**	0.253**	0.230**	0.247**	0.248**	0.233**	0.222**	0.058*	0.013	0.021	0.031	0.060**	0.000	0.000	0.000	0.005	0.050	-			
LOY2	0.227**	0.266**	0.233**	0.294**	0.257**	0.241**	0.259**	0.272**	0.256**	0.236**	0.046	0.000	0.047**	0.026	0.042**	0.000	0.000	0.013	0.032*	0.052**	0.023	-		
LOY4	0.099**	0.164**	0.121**	0.168**	0.131**	0.115**	0.140**	0.141**	0.144**	0.116**	0.191**	0.073**	0.166**	0.174**	0.187**	0.169**	0.137**	0.149**	0.170**	0.165**	0.195**	0.115**	-	
LOY5	0.047**	0.111**	0.086**	0.125**	0.083**	0.072**	0.087**	0.088**	0.098**	0.070**	0.216**	0.067*	0.190**	0.216**	0.213**	0.196**	0.180**	0.182**	0.182**	0.191**	0.191**	0.158**	0.043*	-

Table S2.

(a) SSH09c-1

	NCA01-2	NCA03-4	NCA05	NCA06	NCA07	NCA08	NCA09	NCA10	LOY1-2
NCA01-2 (18)	-								
NCA03-4 (35)	0.000	-							
NCA05 (50)	0.021*	0.027**	-						
NCA06 (43)	0.024*	0.016*	0.002	-					
NCA07 (29)	0.000	0.000	0.000	0.000	-				
NCA08 (46)	0.013	0.009	0.000	0.000	0.000	-			
NCA09 (25)	0.015	0.000	0.025*	0.012	0.000	0.000	-		
NCA10 (33)	0.067***	0.031**	0.045***	0.045	0.000	0.016	0.018	-	
LOY1-2 (23)	0.005	0.032**	0.010	0.021*	0.000	0.005	0.027*	0.054**	-

(b) SSH09c-2

	CHE01	CHE02	CHE03	CHE04	CHE05	CHE06	CHE07	CHE08	CHE09	CHE10
CHE01 (19)	-									
CHE02 (22)	0.081***	-								
CHE03 (22)	0.000	0.046**	-							
CHE04 (26)	0.012	0.093***	0.004	-						
CHE05 (38)	0.010	0.067***	0.000	0.006	-					
CHE06 (26)	0.018	0.088***	0.002	0.009	0.004	-				
CHE07 (35)	0.008	0.090***	0.006	0.018*	0.005	0.005	-			
CHE08 (21)	0.009	0.094***	0.000	0.011	0.005	0.002	0.006	-		
CHE09 (35)	0.002	0.074***	0.000	0.003	0.003	0.005	0.005	0.002	-	
CHE10 (33)	0.014	0.074***	0.000	0.009	0.003	0.010	0.011	0.010	0.009	-

(c) SSH09c-3

	CHE01	CHE02	CHE05	CHE06	CHE07	CHE09	CHE10	LOY4	LOY5
CHE01 (17)	-								
CHE02 (17)	0.058**	-							
CHE05 (18)	0.027	0.062**	-						
CHE06 (22)	0.002	0.043**	0.016	-					
CHE07 (13)	0.016	0.034	0.047*	0.006	-				
CHE09 (14)	0.021	0.027	0.018	0.000	0.009	-			
CHE10 (14)	0.017	0.048*	0.007	0.007	0.032	0.020	-		
LOY4 (19)	0.110***	0.130***	0.104***	0.066***	0.081***	0.093***	0.037***	-	
LOY5 (16)	0.017	0.059***	0.035*	0.009	0.002	0.021	0.002	0.059**	-

Table S3. Pairwise F_{ST} values between populations pairs for SSH09b-1. The number of colonies per population is indicated in parentheses *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$

	CHE	NCA01-2	NCA03-4	NCA05-6-7	NCA08-9-10	LOY3	LOY4	LOY5	LOY6
CHE (30)	-								
NCA01-2 (18)	0.021	-							
NCA03-4 (12)	0.003	0.000	-						
NCA05-6-7 (29)	0.066***	0.043**	0.041**	-					
NCA08-9-10 (27)	0.015	0.000	0.000	0.020*	-				
LOY3 (14)	0.035**	0.029*	0.027	0.058***	0.032**	-			
LOY4 (28)	0.071***	0.022	0.029*	0.026**	0.013	0.066***	-		
LOY5 (16)	0.060***	0.022	0.025	0.047***	0.010	0.039**	0.028**	-	
LOY6 (34)	0.029**	0.000	0.000	0.022*	0.000	0.017	0.019*	0.017	-

Table S4.

(a) SSH09a-1

	MAY	GLO	JDN	EUR	MAD	REU	ROD
MAY (45)	-						
GLO (57)	0.000	-					
JDN (92)	0.000	0.003	-				
EUR (57)	0.000	0.003	0.000	-			
MAD (80)	0.008*	0.008*	0.006*	0.004	-		
REU (99)	0.000	0.000	0.000	0.000	0.000	-	
ROD (82)	0.008	0.000	0.000	0.000	0.000	0.000	

(b) SSH09a-2

	MAY	GLO	JDN	EUR	MAD	REU	ROD	TRO
MAY(45)								
GLO (42)	0.000	-						
JDN (74)	0.000	0.002	-					
EUR (46)	0.000	0.000	0.000	-				
MAD (59)	0.026***	0.004	0.009**	0.000	-			
REU (71)	0.000	0.000	0.000	0.006	0.000	-		
ROD (69)	0.030***	0.000	0.000	0.000	0.000	0.000	-	
TRO (14)	0.048***	0.058***	0.059***	0.052***	0.091***	0.049***	0.088***	

(c) SSH09a-3

	MAY	GLO	JDN	EUR	MAD	REU	ROD
MAY (27)	-						
GLO (25)	0.000	-					
JDN (41)	0.000	0.004	-				
EUR (15)	0.000	0.000	0.000	-			
MAD (25)	0.021**	0.037***	0.020**	0.005	-		
REU (46)	0.000	0.006	0.000	0.006	0.015*	-	
ROD (25)	0.000	0.000	0.000	0.000	0.000	0.000	-

SYNTHESE GENERALE ET PERSPECTIVES

La protection et la conservation des écosystèmes nécessitent une certaine compréhension de leur fonctionnement, notamment en ce qui concerne le maintien dans le temps et dans l'espace des populations et des communautés, ainsi que les processus évolutifs sous-jacents. Ainsi, pour établir des plans de conservation efficaces, l'étude des mécanismes de diversification et de maintien de la diversité génétique des populations animales et végétales est essentielle. Estimer les diversités spécifiques en utilisant différentes approches pour tenter de s'approcher le plus possible de la réalité et étudier le maintien dans le temps des espèces notamment en estimant les degrés de connectivité entre localités afin d'estimer leurs capacités de résilience sont autant de données qui permettront de dresser un premier état des lieux dans un processus de conservation.

Dans cette optique, les coraux du genre *Pocillopora* représentent un modèle très intéressant. En effet, ces coraux sont répartis sur tous les récifs des océans Indien et Pacifique. Dix-sept espèces ont été décrites sur la base de critères morphologiques (Veron, 2000), ce qui représente une diversité potentielle formidable pour étudier les capacités de dispersion et d'adaptation de chacune d'entre elles. Cependant, les premières études comparant les données morphologiques et les données génétiques ont révélé un manque de congruence évident entre ces deux types de données, soulignant l'importance de procéder à des analyses de délimitation d'espèces chez ces coraux avant de mener d'autres types d'analyses. Par ailleurs, certaines morpho-espèces sont encore très peu étudiées à ce jour et l'apport de connaissances demeure fondamental pour comprendre leur fonctionnement, mais aussi pour estimer leur vulnérabilité ou celle de l'écosystème dans son ensemble.

Ce travail de thèse avait pour objectif d'estimer la diversité spécifique des *Pocillopora* en utilisant des méthodes de délimitation d'espèces à partir de marqueurs mitochondriaux et nucléaires, ainsi que d'étudier les flux de gènes pour deux hypothèses d'espèces présentant des écologies contrastées. Les méthodes de délimitation d'espèce n'avaient encore jamais été utilisées pour étudier ce genre. Ainsi, à partir d'un échantillonnage réalisé dans le sud de l'aire de répartition des *Pocillopora* en combinaison avec les données de la littérature, cette étude se veut la plus exhaustive possible afin d'apprécier le mieux possible la diversité génétique du genre. Sur un total de plus de 8 000 colonies échantillonnées au cours de cette thèse, 943 ont été choisies pour représenter la variabilité des morphes de ce genre. Les analyses ont permis d'identifier au moins 18 hypothèses d'espèces secondaires. En outre, les résultats confirment

que la morphologie n'est clairement pas un caractère diagnostique pour l'identification des espèces et l'analyse globale a permis d'identifier 3 lignées complètement nouvelles, représentant possiblement des espèces cryptiques. L'étude de la structuration des populations a permis de révéler des patrons de structuration surprenants et différents pour les deux hypothèses d'espèces étudiées.

1. Synthèse des résultats et discussion

1.1. La délimitation des espèces chez *Pocillopora*

L'analyse de 943 colonies sélectionnées pour représenter la variabilité morphologique du genre, originaires de trois provinces biogéographiques (le Sud-Ouest de l'Océan Indien, le Sud-Ouest du Pacifique Tropical et le Sud-Est de la Polynésie), a permis d'identifier 33 haplotypes ORF dont 18 n'avaient jamais été trouvés. Par ailleurs, 22 haplotypes de la littérature qui n'avaient pas été recensés au sein de l'échantillonnage ont été ajoutés aux analyses, représentant un total de 55 haplotypes différents. Les séquences incluses dans l'analyse proviennent de localités variées et couvrent une grande partie de l'aire de répartition des *Pocillopora*.

La combinaison de plusieurs méthodes de délimitation d'espèces basées sur le marqueur mitochondrial ORF a permis d'identifier 16 hypothèses d'espèces primaires (PSH). La confrontation des PSHs aux résultats des tests d'assignement, conduits à partir des données microsatellites, permet d'identifier plusieurs hypothèses d'espèces secondaires (SSH) parmi les PSHs. En effet, dans le cas où le nombre d'individus correspondant à chaque SSH est suffisant, les microsatellites permettent de confirmer les PSHs identifiées à partir des marqueurs mitochondriaux et parfois de redéfinir plus finement les hypothèses d'espèces, par exemple PSH05, PSH09, PSH13. Ainsi, cette étude révèle quatre cas :

(1) la détection de nouvelles espèces rares et restreintes géographiquement. Cela concerne SSH07, SSH08 et SSH12 qui présentent des haplotypes divergents n'ayant pas été reportés dans les études antérieures et étant restreints à Moorea, Juan de Nova et Nouvelle-Calédonie, respectivement. On peut également y ajouter la SSH01, bien que son identification basée sur la morphologie diverge selon les auteurs (*effusus*, endémique dans le Pacifique Est et *fungiformis*, endémique du Sud de Madagascar), cette étude révèle qu'il s'agirait en réalité d'une seule et même espèce. En effet, ce travail a permis de montrer que cette hypothèse d'espèce est en réalité composée de colonies présentant différents morphes dont un très particulier, encroûtant et

formant des chandelles. Sur un peu plus de 8 000 échantillons collectés, et presque 1 000 colonies séquencées, seules 10 colonies présentent cet haplotype particulier (ORF01), en faisant une espèce extrêmement rare. Cependant, les données acquises permettent d'envisager la réévaluation du statut d'endémisme pour ces deux morpho-espèces (*P. effusus* et *P. fungiformis*) et par conséquent de réévaluer son aire de répartition. En effet, la comparaison des données de la littérature a révélé que cet haplotype avait été précédemment trouvé dans le Pacifique Tropical Est (Pinzón et al., 2013), à Hawaï (Flot et al., 2008; Marti-Puig et al., 2014) et dans l'archipel des Kiribati (Îles Phoenix ; Marti-Puig et al., 2014). Ainsi, ce travail permet également d'agrandir son aire de répartition grâce à l'identification de cet haplotype aux îles Chesterfield dans le Pacifique Ouest ainsi qu'à La Réunion et à Madagascar dans le Sud-Ouest de l'océan Indien.

(2) la confirmation de certaines espèces mais pour lesquelles l'aire de répartition s'en trouve agrandie. C'est notamment le cas pour SSH04, appelée *P. damicornis* type α par (Schmidt-Roach et al., 2013; 2014b) qui est bien considérée comme une seule et même entité génétique homogène et qui, jusqu'à présent, avait été identifiée uniquement dans le Pacifique. Cette étude permet d'agrandir son aire de répartition au Sud-Ouest de l'océan Indien où de rares individus ont été trouvés à Rodrigues.

(3) l'invalidation de certaines morpho-espèces. En effet, l'espèce décrite par Schmidt-Roach et al. (2014b) et nommée *P. bairdi* ne paraît pas pouvoir recevoir le statut d'espèce, aucune méthode de délimitation ne parvenant à l'isoler à partir des marqueurs mitochondriaux. En outre, étant donné que cette morpho-espèce n'a pas été détectée dans notre échantillonnage, les colonies n'ont pas pu être génotypées et les tests d'assignement ne permettent pas d'apporter d'informations supplémentaires.

(4) la présence de complexe d'espèces cryptiques préalablement identifiées comme une seule espèce à distribution large. C'est notamment le cas pour les PSHs pour lesquelles le nombre d'individus est grand : PSH05, PSH09 et PSH13. À titre d'exemple, la PSH09 correspondant au complexe *P. eydouxi/meandrina* se subdivise en trois groupes génétiques homogènes : SSH09a, SSH09b et SSH09c avec les deux dernières qui se retrouvent en sympatrie dans le Pacifique au niveau du site et SSH09a qui est trouvé exclusivement dans l'océan Indien. Ainsi, ce complexe se découperait en deux ou trois espèces génétiques (SSH). Le même type de résultat est observé pour la PSH05, qui se découperait également en plusieurs espèces génétiques. Par ailleurs, la PSH13 se subdivise également en plusieurs groupes génétiques homogènes : SSH13a, SSH13b et SSH13c. Le cas de la SSH13b est particulièrement intéressant. En effet, elle est trouvée en sympatrie avec SSH13a (restreinte à l'océan Indien) et

SSH13c (restreinte à l'océan Pacifique) et présente des caractères morphologiques très particuliers et facilement reconnaissables sur le terrain (aspect velouté, polypes très étirés avec des petites verrues serrées). Ainsi, la SSH13b pourrait représenter une espèce génétique à part entière, à laquelle un nom devra être donné (ou réattribué).

Afin de confirmer les hypothèses d'espèces identifiées dans le Chapitre 1, des analyses de structuration des populations ont été menées sur deux d'entre-elles (les PSH05 et PSH09). Ces analyses permettront également de comparer les résultats obtenus pour deux hypothèses d'espèces supposées avoir des traits de vies différents.

1.2. Spécificité des patterns modes de reproduction en fonction des hypothèses d'espèces

Les deux hypothèses d'espèces sur lesquelles ce travail a porté plus précisément sont *P. damicornis* β (SSH05) et *P. eydouxi* (SSH09). *P. damicornis* β est une espèce plutôt lagunaire vivant dans des eaux peu profondes, protégée des fortes houles mais pouvant être soumise à des régimes hydrodynamiques importants, notamment au niveau des crêtes récifales, et à des variations de luminosité et de températures journalières et saisonnières. Les lagons peuvent être considérés comme des milieux en partie clos, à cause de la barrière corallienne qui peut limiter en partie les flux de gènes. En revanche, *P. eydouxi* a été échantillonné principalement sur la pente externe des récifs entre 0 et 25 m sous la surface. Ces deux espèces présentent également des morphologies différentes, *P. damicornis* β possède un squelette plutôt fin et fragile avec des verrues très allongées alors que *P. eydouxi* correspond à des colonies beaucoup plus massives avec un squelette épais et des grosses verrues. L'étude de la structuration génétique de ces deux espèces a révélé des modes de reproduction et des patterns de structuration génétique contrastés.

Ainsi, les résultats révèlent une différence dans le mode de reproduction de ces deux espèces. En ce qui concerne *P. damicornis* β , une forte propension à la propagation clonale a été identifiée puisqu'en effet, à tous les endroits où elle a été échantillonnée (Mayotte, Juan de Nova, La Réunion et Nouvelle-Calédonie) des clones ont été identifiés. Cette particularité a déjà été mise en évidence par le passé chez *P. damicornis sensu lato* (Adjeroud and Tsuchiya, 1999; Ayre and Miller, 2004; Gorospe and Karl, 2013; Stoddart, 1983; Yeoh and Dai, 2010). Cependant, à la lumière des nouvelles données génétiques, l'identification des espèces étudiées dans la littérature précédemment citée se pose et il est important de considérer avec prudence ces résultats antérieurs. Dans certains cas, l'identification des types de *P. damicornis* peut se

faire *a posteriori* à partir des données de l'ADN mitochondrial (Souter et al., 2009). Plusieurs études ont clairement noté l'existence de propagation clonale chez *P. damicornis* β (Souter et al., 2009; Torda et al., 2013a; Torda et al., 2013b). Cependant, jusqu'à ce jour, une seule étude s'était intéressée à la propagation clonale sur de grandes échelles géographiques (Adjeroud et al., 2014), identifiant la présence de colonies appartenant au même clone sur deux sites distants de 200 km. Les résultats obtenus dans le Chapitre 2 suggèrent que la propagation clonale est possible à des échelles modérées, ce qui est concordant avec les résultats de Souter et al. (2009) et Adjeroud et al. (2014). Cependant, dans le chapitre 3, il n'a pas été possible d'identifier, sur des îles différentes, des colonies appartenant au même clone. Les cinq îles échantillonnées étant distantes au minimum de 530 km (approximativement), cela laisse supposer que la dispersion des larves parthénogénétique est limitée voire inexistante à cette échelle géographique. Cependant, l'échantillonnage des colonies provenant des îles de Juan de Nova, de Mayotte de Nouvelle-Calédonie n'a pas été réalisé dans le but d'identifier des clones et il n'est pas possible d'exclure qu'avec un échantillonnage plus conséquent sur chacune des îles, la présence de clones partagés aurait été révélée entre les différentes îles. La propagation clonale peut être le résultat de la fracture du squelette (par fragmentation ; Highsmith, 1982), sous l'action de vagues puissantes ou de piétinement des baigneurs supposé plus important dans les zones très touristiques. En effet, un morceau détaché de la colonie mère est capable, comme les plantes sont capables de bouturage, de se fixer à nouveau sur le substrat pour former une nouvelle colonie. La propagation de clones peut également être le résultat de bourgeonnement, de l'expulsion des polypes de la colonie (Kvitt et al., 2015; Sammarco, 1982) ou de la production de larves parthénogénétiques (issues de reproduction asexuée et potentiellement identiques au parent). Ce dernier phénomène a été mis en évidence à plusieurs occasions et dans différents endroits (Adjeroud et al., 2014; Souter et al., 2009; Torda et al., 2013b).

En ce qui concerne la seconde espèce, *P. eydouxi* (PSH09), il existait extrêmement peu de données concernant sa reproduction. La seule donnée qui était disponible était qu'il était *broadcast-spawner*, car il émet à la fois ses gamètes mâles et femelles dans la colonne d'eau (Hirose et al., 2001). Les résultats de ce travail montrent que la propagation clonale est inexistante à l'échelle de cet échantillonnage, bien qu'elle ait été mise en évidence dans le Pacifique Tropical Est (Baums et al., 2014; Pinzón et al., 2012) pour des colonies présentant un haplotype ORF type 1 (*sensu* Pinzón et al., 2013). De façon surprenante, dans ces deux dernières études les colonies étudiées ont été décrites comme ressemblant à *P. damicornis sensu lato* alors qu'au sein de notre échantillonnage ce morphe n'a été attribué qu'anecdotiquement aux colonies appartenant à la PSH09. Cette contradiction, laisse supposer que les colonies de

ces études seraient en fait des hybrides type 1/type 5 ou type 1/type 3 comme cela a été mis en évidence dans Combosch and Vollmer (2015). Dans le cas où ces colonies seraient effectivement des hybrides, cela laisse supposer que les traits biologiques tels que la reproduction pourraient être hérités du génome nucléaire, expliquant que les colonies du TEP, bien que présentant un haplotype mitochondrial associé à *P. eydouxi*, soient capables de reproduction clonale.

1.3. Spécificité des patterns de structuration génétique en fonction des hypothèses d'espèces

Les schémas de structuration génétique et de connectivité sont également différents pour les deux espèces de l'étude. En ce qui concerne *P. damicornis* β , les résultats révèlent des valeurs de différenciation génétique plus faibles pour des populations distantes de 400 km, suggérant une dispersion limitée des gènes au-delà de cette distance. Des résultats similaires ont également été mis en évidence en Australie (Torda et al., 2013a), révélant une différenciation génétique plus faible au sein d'un récif (populations séparées de 100 km au maximum) qu'entre récifs (séparés de 500 à 1000 km). En outre, sur de très petites distances, inférieures à 2 km, les populations ne présentent pas de différenciation génétique comme le suggèrent nos résultats au sein du complexe récifal de La Saline/L'ermitage sur la côte ouest de La Réunion et ayant également été mis en évidence à Hawaï chez *P. damicornis sensu lato* (Gorospe and Karl, 2013). Ces différents résultats semblent suggérer un isolement par la distance des populations. Par ailleurs, dans le cas des populations clonales, le problème de la gestion des clones pour les analyses se pose. En effet, conserver les répétitions de chaque génotype risque d'entraîner des biais dans les résultats, notamment par la surreprésentation de certains génotypes. Cependant, ne garder qu'un seul représentant de chaque génotype pour les analyses risque de révéler une vision tronquée de la réalité, notamment si les différentes colonies d'un même clone participent aux événements de reproduction sexuée. Ainsi, afin d'approximer au mieux la structure génétique des populations, réaliser les analyses sur deux jeux de données différents permet ainsi d'essayer de s'approcher le plus possible de la réalité. Ainsi, les résultats révèlent que la différenciation génétique est globalement plus faible quand les estimations sont obtenues à partir du jeu de données tronqué (en ne conservant qu'un seul représentant de chaque clone, supposé représenter la reproduction sexuée). La même tendance a été observée par Adjeroud et al. (2014). Cela pourrait être le reflet de capacités dispersives meilleures pour les larves produites de façon sexuée. Cependant, la dispersion des larves est en

débat. En effet, chez les scléreactiniaires, Ayre & Miller (2004) ont trouvé que les larves ont tendance à s'installer près de leurs parents alors que Sherman et al. (2005) trouvent que l'installation des larves près des parents était rarement une réussite, supposant que l'allocation serait favorisée. En outre, le résultat des tests d'assignement (réalisés pour les deux jeux de données) révèlent un pattern de structuration surprenant, notamment parce que les populations de Nouvelle-Calédonie et de Mayotte se trouvent assignées dans le même cluster génétique. Ceci suggère qu'il pourrait exister des flux de gènes entre ces deux zones, mais qui seraient unidirectionnels, allant de l'Est vers l'Ouest en lien avec la circulation océanique globale de l'océan Indien. Par ailleurs, la structuration observée pourrait être le résultat d'un isolement marqué des populations où agirait la dérive génétique et stabiliserait les mêmes allèles dans les populations par le fait du hasard. Ceci pourrait être conforté par l'important déséquilibre de liaison observé au sein de chaque population qui est supposé devenir grand quand la migration est limitée (Ohta, 1982a, b).

En ce qui concerne *P. eydouxi*, les tests d'assignement révèlent un joli puzzle à résoudre. L'analyse de la structure génétique révèle un partitionnement fractal des différentes SSHs. En effet, chaque SSH se divise en clusters qui eux-mêmes se divisent en sous-clusters et ce, avec les deux méthodes d'assignement utilisées (STRUCTURE and DAPC). Cependant, la comparaison des deux différentes méthodes révèle que les assignements diffèrent pour les niveaux de structuration les plus bas. Ainsi, il a été choisi de ne considérer que les niveaux de structuration congruents entre les différentes méthodes en suivant un principe de parcimonie. Cette approche permet d'identifier huit différents clusters pour l'ensemble des SSHs (trois clusters pour SSH09a et SSH09c et deux pour SSH09b). Les différents groupes génétiques identifiés présentent la particularité d'être retrouvés dans différentes populations, de façon plus ou moins homogène (en fonction des SSHs), sans lien avec la géographie. En outre, Plusieurs hypothèses ont été proposées pour expliquer la structuration observée. En effet, cela pourrait résulter d'une spéciation allopatrique et de la mise en place de barrières à la reproduction empêchant les colonies de se reproduire et ainsi d'homogénéiser les fréquences alléliques au sein de chaque population. Par ailleurs, un isolement historique aurait pu entraîner une divergence des lignées ce qui peut résulter en une hétérogénéité de la différenciation génétique le long du génome, comme cela a été montré chez deux populations de bars entre l'Atlantique et la Méditerranée (Tine et al., 2014). Une telle variabilité le long des génomes peut entraîner des problèmes d'interprétations, notamment lorsque les génomes sont complètement inconnus, comme c'est le cas pour le génome nucléaire des *Pocillopora*. Ainsi, si les microsatellites utilisés pour estimer la différenciation des populations sont situés dans une zone peu contrainte

du génome, les groupes génétiques révélés apparaissent comme très différenciés pouvant correspondre à des espèces différentes, alors que si les microsatellites sont localisés dans des régions très contraintes, les groupes génétiques pourraient ne plus apparaître, menant à la conclusion qu'il n'existerait qu'un seul groupe génétique correspondant à une seule population panmictique. La sélection par l'habitat peut également expliquer la structuration observée. En effet, les individus pourraient s'être adaptés à des micro-habitats différents en fonction de l'exposition à la lumière, aux courants locaux, à la salinité ou encore à la disponibilité en nourriture. A titre d'exemple, la profondeur pourrait jouer ce rôle structurant, comme cela a été mis en évidence chez *Seriatopora hystrix* en Australie (Van Oppen et al., 2011), limitant ainsi les flux de gènes et par conséquent la connectivité entre les différents habitats. Cependant, en ce qui concerne les colonies de PSH09, elles ont été échantillonnées à des gammes de profondeur similaires (entre 8 et 14 m), ce qui rend peu probable le fait que ce soit la profondeur qui soit à l'origine des différents groupes génétiques observés. La connectivité génétique fait référence non seulement au transport larvaire mais aussi à la capacité à se sédentariser (recrutement), survivre et de participer à la génération suivante dans le nouvel habitat. En d'autres mots, des habitats défavorables peuvent diminuer les niveaux de connectivité même en cas d'échanges importants d'individus entre populations. Ceci est appelé le « *phenotype-environment mismatch* » (inadéquation phénotype/environnement ; Marshall et al., 2010) et peut présenter une barrière biologique aux flux de gènes pour des organismes dont l'échelle d'hétérogénéité spatiale est plus petite que l'échelle du transport larvaire et où la mortalité non aléatoire arrive après la dispersion.

Une autre hypothèse proposée au cours de ce travail pour expliquer les patrons de structuration surprenants observés chez *P. eydouxi* serait l'hybridation. En effet, chez les coraux, les exemples d'hybridation entre espèces ne manquent pas comme par exemple chez les *Acropora* (Márquez et al., 2002; van Oppen et al., 2000), les *Stylophora* (Arrigoni et al., 2016; Flot et al., 2011) ou encore les *Madracis* (Frade et al., 2010). Chez les coraux du genre *Pocillopora*, l'hybridation a été soupçonnée dans plusieurs études menées dans le Pacifique Tropical Est (Combosch et al., 2008; Pinzón and LaJeunesse, 2011) avant d'être formellement mise en évidence grâce à des données RADSeq (Combosch and Vollmer, 2015). La chose intéressante dans ces dernières études est le fait que chaque fois que l'hybridation a été mise en évidence chez *Pocillopora*, les coraux possédant un haplotype ORF type 1 étaient impliqués, laissant supposer qu'il n'y aurait pas ou peu de barrières à la reproduction chez cette lignée mitochondriale. En outre, différentes morpho-espèces de *Pocillopora* sont trouvées en sympatrie sur tous les sites échantillonnés suggérant la possibilité que différentes morpho-

espèces puissent féconder les ovules des lignées maternelles de type 1 et résultant en différents « lots » génétiques différents et retrouvés grâce aux marqueurs microsatellites.

A ce jour, il n'est pas possible de privilégier une hypothèse plutôt qu'une autre et d'autres investigations doivent être conduites pour pouvoir conclure avec certitude sur l'origine de ces différents clusters.

2. Conclusions et perspectives

Ce travail a permis d'apporter des connaissances sur la diversité génétique des *Pocillopora* ainsi que sur la structuration génétique des populations de deux hypothèses d'espèces. Cependant, la plupart des résultats obtenus soulèvent de nombreuses questions et les perspectives d'études pour les résoudre sont nombreuses.

Pour répondre aux différentes hypothèses proposées dans les chapitres de structuration des populations, notamment sur l'origine de la structuration observée pour chacune des deux espèces étudiées au cours de cette thèse. L'utilisation de données de RADseq permettrait de déterminer l'origine évolutive de ces groupes génétiques et quantifier les taux d'introgession entre eux, s'il y a lieu. Les données RADseq pourraient également permettre de confirmer ou infirmer les hypothèses d'espèces proposées dans le Chapitre 1. En outre, la quantité de données récoltées au cours de ce travail est très importante et toutes n'ont pas encore été exploitées. Ainsi, il serait possible d'évaluer la connectivité et la structuration des populations pour deux autres PSHs (PSH04 et PSH13) qui présentent un nombre suffisant d'individus et permettant encore d'enrichir les connaissances sur les différentes hypothèses espèces de ce genre.

Par ailleurs, des études récentes révèlent des cas de chimérisme et de mosaïcisme chez les coraux et notamment chez *Pocillopora* (Rinkevich et al., 2016; Schweinsberg et al., 2015). Ces études sont encore rares chez les coraux, et il serait intéressant d'étudier les variations génotypiques au sein d'une même colonie hôte. Le chimérisme correspond à la fusion de deux génomes d'origines différentes tandis que le mosaïcisme résulte de mutations somatiques entraînant la présence de génomes distincts différenciés de quelques allèles. Ces deux processus donnent alors naissance à une colonie physique présentant plusieurs génotypes (Rinkevich et al., 2016; Schweinsberg et al., 2015). Cette diversité génétique intra-colonie doit être explorée avec des outils moléculaires car ce pourrait être une piste pour expliquer les difficultés qui résident dans la définition des espèces chez les coraux du genre *Pocillopora*. D'autant que chez *P. damicornis sensu lato* l'existence de larves chimériques d'origine asexuée a été mise en

évidence (Rinkevich et al., 2016) et que ces larves chimériques pourraient présenter une fitness plus élevée que les autres. En effet, celles-ci représentent une alliance entre deux génotypes qui pourrait générer des avantages adaptatifs (Rinkevich et al., 2016). Étudier ces phénomènes pourrait donc apporter des pistes dans la compréhension des capacités d'adaptation des individus et leur maintien dans le temps.

Enfin, étant donné la complexité qui réside dans la compréhension des organismes coralliens, il paraît intéressant de ne pas regarder que l'hôte mais de regarder l'holobionte dans son ensemble (corail, algues, champignons, bactéries, Archéobactéries, virus). En effet, les coraux vivent en association avec de nombreux organismes tels que les algues symbiotiques photosynthétiques unicellulaires du genre *Symbiodinium* (appelées les zooxanthelles). Ces micro-algues fournissent les produits de la photosynthèse à l'hôte et permettent à ce dernier de croître plus rapidement. Les zooxanthelles sont identifiées grâce à différents marqueurs moléculaires (ITS2, par exemple) que l'on distingue alors selon leurs séquences et que l'on nomme type. Chaque colonie hôte peut abriter différents types, avec certains types dominants et d'autres plus cryptiques. Ce serait d'ailleurs ces types cryptiques qui pourraient permettre à l'hôte corail de récupérer après un événement de blanchissement. Par ailleurs, d'autres micro-organismes pourraient avoir un rôle dans la vie des organismes coralliens. En effet, certains champignons pourraient jouer un rôle dans le métabolisme du carbone et de l'azote, comme cela a été supposé chez *Porites asteroides* (Wegley et al., 2007), mais ce rôle demeure encore méconnu. Par ailleurs, la présence des bactéries pourrait avoir une action « anti-fouling » (anti-encrassement) sur les coraux en déposant un biofilm protecteur (par exemple, Golberg et al., 2013), mais aussi également avoir un rôle dans la calcification car l'association de certaines bactéries avec des dinoflagellés en culture entraîne la calcification des algues unicellulaires (Frommlet et al., 2015). Ainsi, étudier l'ensemble de l'holobionte, pourrait permettre d'avoir une vision plus large de la survie des organismes coralliens face aux changements globaux.

Références

- Adjeroud, M., Guérécheau, A., Vidal-Dupiol, J., Flot, J.-F., Arnaud-Haond, S., Bonhomme, F., 2014. Genetic diversity, clonality and connectivity in the scleractinian coral *Pocillopora damicornis*: a multi-scale analysis in an insular, fragmented reef system. *Marine Biology* 161, 531-541.
- Adjeroud, M., Tsuchiya, M., 1999. Genetic variation and clonal structure in the scleractinian coral *Pocillopora damicornis* in the Ryukyu Archipelago, southern Japan. *Marine Biology* 134, 753-760.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *J Mol Biol* 215.
- Amar, K.O., Chadwick, N.E., Rinkevich, B., 2007. Coral planulae as dispersion vehicles: biological properties of larvae released early and late in the season. *Marine Ecology Progress Series* 350, 71-78.
- Apprill, A., Marlow, H.Q., Martindale, M.Q., Rappe, M.S., 2009. The onset of microbial associations in the coral *Pocillopora meandrina*. *ISME J* 3, 685-699.
- Arnaud-Haond, S., Moalic, Y., Hernández-García, E., Eguiluz, V.M., Alberto, F., Serrão, E.A., Duarte, C.M., 2014. Disentangling the influence of mutation and migration in clonal seagrasses using the genetic diversity spectrum for microsatellites. *Journal of Heredity*.
- Arrigoni, R., Benzoni, F., Terraneo, T.I., Caragnano, A., Berumen, M.L., 2016. Recent origin and semi-permeable species boundaries in the scleractinian coral genus *Stylophora* from the Red Sea. *Scientific Reports* 6, 34612.
- Ayata, S.D., 2010. Importance relative des facteurs hydroclimatiques et des traits d'histoire de vie sur la dispersion larvaire et la connectivité à différentes échelles spatiales (Manche, Golfe Gascogne). Université Pierre et Marie Curie, p. 412.
- Ayre, D.J., Hughes, T.P., 2000. Genotypic diversity and gene flow in the brooding and spawning corals along the Great Barrier Reef, Australia. *Evolution* 54, 1590-1605.
- Ayre, D.J., Miller, K.J., 2004. Where do clonal coral larvae go? Adult genotypic diversity with reproductive effort in the brooding coral *Pocillopora damicornis*. *Marine Ecology Progress Series* 277, 95-105.
- Ayre, D.J., Resing, J.M., 1986. Sexual and asexual production of planulae in reef corals. *Marine Biology* 90, 187-190.
- Baird, A.H., Guest, J.R., Willis, B.L., 2009. Systematic and Biogeographical Patterns in the Reproductive Biology of Scleractinian Corals. *Annual Review of Ecology, Evolution, and Systematics* 40, 551-571.
- Balloux, F., Lehmann, L., De Meeûs, T., 2003. The population genetics of clonal and partially clonal diploids. *Genetics* 164, 1635-1644.
- Barluenga, M., Stölting, K.N., Salzburger, W., Muschick, M., Meyer, A., 2006. Sympatric speciation in Nicaraguan crater lake cichlid fish. *Nature* 439, 719-723.
- Baum, D.A., Shaw, K.L., 1995. Genealogical perspectives on the species problem. *Experimental and molecular approaches to plant biosystematics* 53, 123-124.
- Baums, I.B., Devlin-Durante, M., Laing, B.A.A., Feingold, J., Smith, T., Bruckner, A., Monteiro, J., 2014. Marginal coral populations: the densest known aggregation of *Pocillopora* in the Galápagos Archipelago is of asexual origin. *Frontiers in Marine Science* 1.
- Birky, C.W., 1996. Heterozygosity, heteromorphy, and phylogenetic trees in asexual eukaryotes. *Genetics* 144, 427-437.
- Bohonak, A.J., 1999. Dispersal, gene flow and population structure. *Q Rev Biol* 74, 21-44.

- Bowen, B.W., Rocha, L.A., Toonen, R.J., Karl, S.A., 2013. The origins of tropical marine biodiversity. *Trends in Ecology & Evolution* 28, 359-366.
- Braby, M.F., Eastwood, R., Murray, N., 2012. The subspecies concept in butterflies: has its application in taxonomy and conservation biology outlived its usefulness? *Biological Journal of the Linnean Society* 106, 699-716.
- Bruvo, R., Michiels, N.K., D'Souza, T.G., Schulenburg, H., 2004. A simple method for the calculation of microsatellite genotype distances irrespective of ploidy level. *Mol Ecol* 13, 2101-2106.
- Budd, A.F., Romano, S.L., Smith, N.D., Barbeitos, M.S., 2010. Rethinking the phylogeny of scleractinian corals: a review of morphological and molecular data. *Integr Comp Biol* 50, 411-427.
- Cardinale, B.J., Duffy, J.E., Gonzalez, A., Hooper, D.U., Perrings, C., Venail, P., Narwani, A., Mace, G.M., Tilman, D., Wardle, D.A., Kinzig, A.P., Daily, G.C., Loreau, M., Grace, J.B., Larigauderie, A., Srivastava, D.S., Naeem, S., 2012. Biodiversity loss and its impact on humanity. *Nature* 486, 59-67.
- Carlson, D.B., 1999. The evolution of mating systems in tropical reef corals. *Tree* 14, 491-495.
- Cbol Plant Working Group, Hollingsworth, P.M., Forrest, L.L., Spouge, J.L., Hajibabaei, M., Ratnasingham, S., van der Bank, M., Chase, M.W., Cowan, R.S., Erickson, D.L., Fazekas, A.J., Graham, S.W., James, K.E., Kim, K.-J., Kress, W.J., Schneider, H., van AlphenStahl, J., Barrett, S.C.H., van den Berg, C., Bogarin, D., Burgess, K.S., Cameron, K.M., Carine, M., Chacón, J., Clark, A., Clarkson, J.J., Conrad, F., Devey, D.S., Ford, C.S., Hedderson, T.A.J., Hollingsworth, M.L., Husband, B.C., Kelly, L.J., Kesanakurti, P.R., Kim, J.S., Kim, Y.-D., Lahaye, R., Lee, H.-L., Long, D.G., Madriñán, S., Maurin, O., Meusnier, I., Newmaster, S.G., Park, C.-W., Percy, D.M., Petersen, G., Richardson, J.E., Salazar, G.A., Savolainen, V., Seberg, O., Wilkinson, M.J., Yi, D.-K., Little, D.P., 2009. A DNA barcode for land plants. *Proceedings of the National Academy of Sciences of the United States of America* 106, 12794-12797.
- 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., 2011. Population genetics of an ecosystem-defining reef coral *Pocillopora damicornis* in the Tropical Eastern Pacific. *PLoS One* 6, e21200.
- 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.
- Cowen, R.K., Foundation, N.S., 2003. Population Connectivity in Marine Systems: Report of a Workshop to Develop Science Recommendations for the National Science Foundation, November 4-6, 2002, Durango, Colorado. American Geophysical Union.
- Crow, J.F., 1994. Advantages of sexual reproduction. *Developmental genetics* 15, 205-213.
- Crow, J.F., 1999. The omnipresent process of sex. *Journal of Evolutionary Biology* 12, 1023-1025.
- Crutzen, P.J., 2006. The "Anthropocene". In: Ehlers, E., Krafft, T. (Eds.), *Earth System Science in the Anthropocene*. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 13-18.
- Dagg, A.I., Foster, J.B., 1976. *The Giraffe: Its Biology, Behaviour and Ecology*. Van Nostrand Reinhold.
- Dayrat, B., 2005. Towards integrative taxonomy. *Biological Journal of the Linnean Society* 85, 407-415.
- de Meeûs, T., Prugnolle, F., Agnew, P., 2007. Asexual reproduction: Genetics and evolutionary aspects. *Cellular and Molecular Life Sciences* 64, 1355-1372.

- De Queiroz, K., 1998. The general lineage concept of species, species criteria, and the process of speciation: A conceptual unification and terminological recommendations. In: Howard, D.J., Berchler, S.H. (Eds.), *Endless Forms: Species and Speciation*. Oxford University Press, Oxford, England.
- De Queiroz, K., 2007. Species concepts and species delimitation. *Systematic Biology* 56, 879-886.
- Di Rienzo, A., Peterson, A.C., Garza, J.C., Valdes, A.M., Slatkin, M., Freimer, N.B., 1994. Mutational processes of simple-sequence repeat loci in human populations. *Proceedings of the National Academy of Sciences of the United States of America* 91, 3166-3170.
- Donoghue, M.J., 1985. A critique of the biological species concept and recommendations for a phylogenetic alternative. *Bryologist* 88, 172-181.
- Doyle, J.J., 1995. The irrelevance of allele tree topologies for species delimitation, and a non-topological alternative. *Systematic Botany* 20, 574-588.
- Evanno, G., Regnaut, S., Goudet, J., 2005. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14, 2611-2620.
- Fabricius, K., Metzner, J., 2004. Scleractinian walls of mouths: Predation on coral larvae by corals. *Coral Reefs* 23.
- Fadlallah, Y.H., 1983. Sexual Reproduction, Development and Larval Biology in Scleractinian Corals. *Coral Reef* 2, 129-150.
- Fennessy, J., Bidon, T., Reuss, F., Kumar, V., Elkan, P., Nilsson, Maria A., Vamberger, M., Fritz, U., Janke, A., 2016. Multi-locus analyses reveal four giraffe species instead of one. *Current Biology* 26, 1-7.
- Fisher, R.A., 1930. *The genetical theory of natural selection: a complete variorum edition*. Oxford University Press.
- Fitzpatrick, B.M., 2002. Molecular correlates of reproductive isolation. *Evolution* 56, 191-198.
- Flot, J.-F., Couloux, A., Tillier, S., 2010. Haplowebs as a graphical tool for delimiting species: a revival of Doyle's "field for recombination" approach and its application to the coral genus *Pocillopora* in Clipperton. *BMC Evolutionary Biology* 10, 1-14.
- Flot, J.-F., Hespels, B., Li, X., Noel, B., Arkhipova, I., Danchin, E.G.J., Hejnol, A., Henrissat, B., Koszul, R., Aury, J.-M., Barbe, V., Barthelemy, R.-M., Bast, J., Bazykin, G.A., Chabrol, O., Couloux, A., Da Rocha, M., Da Silva, C., Gladyshev, E., Gouret, P., Hallatschek, O., Hecox-Lea, B., Labadie, K., Lejeune, B., Piskurek, O., Poulain, J., Rodriguez, F., Ryan, J.F., Vakhrusheva, O.A., Wajnberg, E., Wirth, B., Yushenova, I., Kellis, M., Kondrashov, A.S., Welch, D.B.M., Pontarotti, P., Weissenbach, J., Wincker, P., Jaillon, O., Van Doninck, K., 2013. Genomic evidence for ameiotic evolution in the bdelloid rotifer *Adineta vaga*. *Nature* 500, 453-457.
- Flot, J.F., Blanchot, J., Charpy, L., Cruaud, C., Licuanan, W.Y., Nakano, Y., Payri, C., Tillier, S., 2011. Incongruence between morphotypes and genetically delimited species in the coral genus *Stylophora*: phenotypic plasticity, morphological convergence, morphological stasis or interspecific hybridization? *BMC Ecol* 11, 22.
- 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.
- Flot, J.F., Tillier, S., 2006. Molecular phylogeny and systematics of the scleractinian coral genus *Pocillopora* in Hawaii. *Proceedings of 10th international coral reef symposium*.
- Fontaneto, D., Herniou, E.A., Boschetti, C., Caprioli, M., Melone, G., Ricci, C., Barraclough, T.G., 2007. Independently evolving species in asexual bdelloid rotifers. *PLoS Biol* 5, e87.
- Frade, P.R., Reyes-Nivia, M.C., Faria, J., Kaandorp, J.A., Luttikhuisen, P.C., Bak, R.P., 2010. Semi-permeable species boundaries in the coral genus *Madracis*: introgression in a brooding coral system. *Mol Phylogenet Evol* 57, 1072-1090.

- Frommlet, J.C., Sousa, M.L., Alves, A., Vieira, S.I., Suggett, D.J., Serôdio, J., 2015. Coral symbiotic algae calcify ex hospite in partnership with bacteria. *Proceedings of the National Academy of Sciences* 112, 6158-6163.
- Fujisawa, T., Barraclough, T.G., 2013. Delimiting species using single-locus data and the Generalized Mixed Yule Coalescent approach: a revised method and evaluation on simulated data sets. *Systematic Biology* 62, 707-724.
- Fujita, M.K., Leaché, A.D., Burbrink, F.T., McGuire, J.A., Moritz, C., 2012. Coalescent-based species delimitation in an integrative taxonomy. *Trends in Ecology & Evolution* 27, 480-488.
- Gao, H., Bryc, K., Bustamante, C.D., 2011. On identifying the optimal number of population clusters via the Deviance Information Criterion. *Plos One* 6, e21014.
- Gleason, D.F., Hofmann, D.K., 2011. Coral larvae: from gametes to recruits. *Journal of Experimental Marine Biology and Ecology* 408, 42-57.
- Golberg, K., Pavlov, V., Marks, R.S., Kushmaro, A., 2013. Coral-associated bacteria, quorum sensing disrupters, and the regulation of biofouling. *Biofouling* 29, 669-682.
- Goldstein, D.B., Linares, A.R., Cavalli-Sforza, L.L., Feldman, M.W., 1995. An evaluation of genetic distances for use with microsatellite loci. *Genetics* 139, 463-471.
- Gorospe, K.D., Karl, S.A., 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.
- Grigg, R.W., 1983. Community structure, succession and development of coral reefs in Hawaii. *Marine Ecology Progress Series* 11, 1-14.
- Grimm, V., Reise, K., Strasser, M., 2003. Marine metapopulations: a useful concept? *Helgoland Marine Research* 56, 222-228.
- Guichoux, E., Lagache, L., Wagner, S., Chaumeil, P., LÉGer, P., Lepais, O., Lepoittevin, C., Malausa, T., Revardel, E., Salin, F., Petit, R.J., 2011. Current trends in microsatellite genotyping. *Molecular Ecology Resources* 11, 591-611.
- Hanski, I., Gilpin, M., 1991. Metapopulation dynamics: brief history and conceptual domain. *Biological Journal of the Linnean Society* 42, 3-16.
- Harrison, P.L., 2011. Sexual reproduction of scleractinian corals. *Coral reefs: an ecosystem in transition*. Springer Netherlands, pp. 59-85.
- Hatta, M., Fukami, H., Wang, W., Omori, M., Shimoike, K., Hayashibara, T., Ina, Y., Sugiyama, T., 1999. Reproductive and genetic evidence for reticulate evolutionary history of mass-spawning corals. *Mol. Biol. Evol* 16, 1607-1613.
- Hebert, P.D.N., Cywinska, A., Ball, S.L., deWaard, J.R., 2003a. Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences* 270, 313-321.
- Hebert, P.D.N., Ratnasingham, S., de Waard, J.R., 2003b. Barcoding animal life: cytochrome *c* oxidase subunit 1 divergences among closely related species. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270, S96.
- Hellberg, M.E., 2006. No variation and low synonymous substitution rates in coral mtDNA despite high nuclear variation. *BMC Evol Biol* 6, 24.
- Heyward, A.J., Babcock, R.C., 1986. Self- and cross-fertilization in scleractinian corals. *Marine Biology* 90, 191-195.
- Heyward, A.J., Negri, A.P., 1999. Natural inducers for coral larval metamorphosis. *Coral Reef* 18, 273-279.
- Highsmith, R.C., 1982. Reproduction by fragmentation in corals. *Marine Ecology Progress Series* 7, 207-226.
- Hirose, M., Kinzie III, R.A., Hidaka, M., 2001. Timing and process of entry of zooxanthellae into oocytes of hermatic corals. *Coral Reef* 20, 273-280.

- Hirose, M., Yamamoto, H., Nonaka, M., 2007. Metamorphosis and acquisition of symbiotic algae in planula larvae and primary polyps of *Acropora* spp. *Coral Reefs* 27, 247-254.
- Isomura, N., Nishihira, M., 2001. Size variation of planulae and its effect on the lifetime of planulae in three pocilloporid corals. *Coral Reefs* 20, 309-315.
- Jackson, G.A., 1986. Interaction of physical and biological processes in the settlement of planktonic larvae. *Bulletin of Marine Science* 39, 202-212.
- 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.
- Kingman, J.F.C., 1982. On the genealogy of large populations. *J Appl Probab* 19, 27-43.
- Knowlton, N., Maté, J.L., Guzmán, H.M., Rowan, R., Jara, J., 1997. Direct evidence for reproductive isolation among the three species of the *Montastraea annularis* complex in Central America (Panamá and Honduras). *Marine Biology* 127, 705-711.
- Kondrashov, A.S., 1988. Deleterious mutations and the evolution of sexual reproduction. *Nature* 336, 435-440.
- Kruger, A., Schleyer, M.H., 1998. Sexual reproduction in the coral *Pocillopora verrucosa* (Cnidaria: Scleractinia) in KwaZulu-Natal, South Africa. *Marine Biology* 132, 703-710.
- Kvitt, H., Kramarsky-Winter, E., Maor-Landaw, K., Zandbank, K., Kushmaro, A., Rosenfeld, H., Fine, M., Tchernov, D., 2015. Breakdown of coral colonial form under reduced pH conditions is initiated in polyps and mediated through apoptosis. *Proc Natl Acad Sci U S A* 112, 2082-2086.
- Legendre, P., 2000. Comparison of permutation methods for the partial correlation and partial mantel tests. *J. Statist. Comput. Simul.* 67, 37-73.
- Levins, R., 1969. Some demographic and genetic consequences of environmental heterogeneity for biological control. *Bulletin of the Entomological society of America* 15, 237-240.
- Levins, R., 1970. Extinction. In: M, G. (Ed.), *Some mathematical questions in biology*. American Mathematical Society, Providence, R.I.
- Lynch, M., 1990. The rate of morphological evolution in mammals from the standpoint of the neutral expectation. *American Naturalist*, 727-741.
- Magalon, H., Adjeroud, M., Veuille, M., 2005. Patterns of genetic variation do not correlate with geographical distance in the reef building coral *Pocillopora meandrina* in the South Pacific. *Molecular Ecology* 14, 1861-1868.
- Malay, M.C.D., Paulay, G., 2010. Peripatric speciation drives diversification and distributional pattern of reef hermit crabs (Decapoda: Diogenidae: Calcinus). *Evolution* 64, 634-662.
- Mallet, J., 1995. A species definition for the modern synthesis. *Trends in Ecology & Evolution* 10, 294-299.
- Mallet, J., 2008. Hybridization, ecological races and the nature of species: empirical evidence for the ease of speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences* 363, 2971.
- Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. *Cancer Research* 27, 209.
- Márquez, L.M., Van Oppen, M.J., Willis, B., Reyes, A., Miller, D.J., 2002. The highly cross-fertile coral species, *Acropora hyacinthus* and *Acropora cytherea*, constitute statistically distinguishable lineages. *Molecular Ecology* 11, 1339-1349.
- Marshall, D.J., Monroe, K., Bode, M., Keough, M.J., Swearer, S., 2010. Phenotype–environment mismatches reduce connectivity in the sea. *Ecology Letters* 13, 128-140.
- Marti-Puig, P., Forsman, Z.H., Haverkort-Yeh, R.D., Knapp, I.S.S., Maragos, J.E., Toonen, R.J., 2014. Extreme phenotypic polymorphism in the coral genus *Pocillopora*; micro-

- morphology corresponds to mitochondrial groups, while colony morphology does not. *Bulletin of Marine Science* 90, 211-231.
- Mather, P., Bennett, I., 1993. *A Coral reef handbook: a guide to the geology, flora and fauna of the Great Barrier Reef*, University of California.
- Mayden, R.L., 2002. On biological species, species concepts and individuation in the natural world. *Fish and Fisheries* 3, 171-196.
- Mayr, E., 1942. *Systematics and the origin of species, from the viewpoint of a zoologist*. Harvard University Press.
- McCormick, M.I., 2003. Consumption of coral propagules after mass spawning enhances larval quality of damselfish through maternal effects. *Oecologia* 136, 37-45.
- Meier, R., Willmann, R., 2000. The Hennigian Species Concept. In: *Species concepts and phylogenetic theory: A debate*. Columbia University Press, United States of America.
- Mesak, F., Tatarenkov, A., Earley, R.L., Avise, J.C., 2014. Hundreds of SNPs vs. dozens of SSRs: which dataset better characterizes natural clonal lineages in a self-fertilizing fish? *Frontiers in Ecology and Evolution* 2, 74.
- Michener, D., 1976. Diverse approaches to systematics. *Evol Biol* 4, 1-38.
- Miller, K.J., Ayre, D.J., 2004. The role of sexual and asexual reproduction in structuring high latitude populations of the reef coral *Pocillopora damicornis*. *Heredity* 92, 557-568.
- Miller, K.J., Mundy, C.N., 2005. In situ fertilisation success in the scleractinian coral *Goniastrea favulus*. *Coral Reefs* 24, 313-317.
- Mittermeier, R.A., Myers, N., Thomsen, J.B., 2008. Biodiversity hotspots and major tropical wilderness areas: approaches to setting conservation priorities. *Conservation Biology* 12, 516-520.
- Moalic, Y., Arnaud-Haond, S., Perrin, C., Pearson, G.A., Serrao, E.A., 2011. Travelling in time with networks: Revealing present day hybridization versus ancestral polymorphism between two species of brown algae, *Fucus vesiculosus* and *F. spiralis*. *BMC Evolutionary Biology* 11, 33-33.
- Muir, P.R., 1984. Periodicity and asexual planulae production in *Pocillopora damicornis* (Linnaeus) at Magnetic Island. James Cook University, p. 60.
- Muller, H.J., 1932. Some genetic aspects of sex. *Am Nat* 66, 119-138.
- Muller, H.J., 1964. The relation of recombination to mutational advance. *Mutat Res* 1, 2-9.
- Muscatine, L., Porter, J.W., 1977. Reef corals: mutualistic symbioses adapted to nutrient-poor environments. *BioScience* 27, 454-460.
- Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J., 2000. Biodiversity hotspots for conservation priorities. *Nature* 403, 853-858.
- Nishikawa, A., Katoh, M., Sakai, K., 2003. Larval settlement rates and gene flow of broadcast-spawning (*Acropora tenuis*) and planula-brooding (*Stylophora pistillata*) corals. *Marine Ecology Progress Series* 256, 87-97.
- Nosil, P., Harmon, L.J., Seehausen, O., 2009. Ecological explanations for (incomplete) speciation. *Trends in Ecology & Evolution* 24, 145-156.
- Ohta, T., 1982a. Linkage disequilibrium due to random genetic drift in finite subdivided populations. *Proc Natl Acad Sci U S A* 79, 1940-1944.
- Ohta, T., 1982b. Linkage disequilibrium with the island model. *Genetics* 101, 139-155.
- Padial, J.M., Miralles, A., De la Riva, I., Vences, M., 2010. The integrative future of taxonomy. *Frontiers in Zoology* 7, 1-14.
- Palumbi, S.R., 1992. Marine speciation on a small planet. *Trends in Ecology & Evolution* 7, 114-118.
- Park, E., Hwang, D.S., Lee, J.S., Song, J.I., Seo, T.K., Won, Y.J., 2012. Estimation of divergence times in cnidarian evolution based on mitochondrial protein-coding genes and the fossil record. *Mol Phylogenet Evol* 62, 329-345.

- Paulay, G., Meyer, C., 2002. Diversification in the Tropical Pacific: comparisons between marine and terrestrial systems and the importance of founder speciation. *Integrative and Comparative Biology* 42, 922-934.
- Paz-García, D.A., Chávez-Romo, H.E., Correa-Sandoval, F., Reyes-Bonilla, H., López-Pérez, A., Medina-Rosas, P., Hernández-Cortés, M.P., 2012. Genetic connectivity patterns of corals *Pocillopora damicornis* and *Porites panamensis* (Anthozoa: Scleractinia) along the west coast of Mexico. *Pacific Science* 66, 43-61.
- Pineda, J., Hare, J.A., Sponaugle, S., 2007. Larval transport and dispersal in the coastal ocean and consequences for population connectivity. *Oceanography* 20, 22-39.
- 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., Reyes-Bonilla, H., Baums, I.B., LaJeunesse, T.C., 2012. Contrasting clonal structure among *Pocillopora* (Scleractinia) communities at two environmentally distinct sites in the Gulf of California. *Coral Reefs* 31, 765-777.
- Pinzón, J.H., Sampayo, E., Cox, E., Chauka, L.J., Chen, C.A., Voolstra, C.R., LaJeunesse, T.C., Maggs, C., 2013. Blind to morphology: genetics identifies several widespread ecologically common species and few endemics among Indo-Pacific cauliflower corals (*Pocillopora*, Scleractinia). *Journal of Biogeography* 40, 1595-1608.
- Pons, J., Barraclough, T.G., Gomez-Zurita, J., Cardoso, A., Duran, D.P., Hazell, S., Kamoun, S., Sumlin, W.D., Vogler, A.P., 2006. Sequence-based species delimitation for the DNA taxonomy of undescribed insects. *Systematic Biology* 55, 595-609.
- Postaire, B., Magalon, H., Bourmaud, C.A.F., Bruggemann, J.H., 2016. Molecular species delimitation methods and population genetics data reveal extensive lineage diversity and cryptic species in Aglaopheniidae (Hydrozoa). *Molecular Phylogenetics and Evolution* 105, 36-49.
- Pritchard, J.K., Stephens, M., Donnelly, P., 2000. Inference of population structure using multilocus genotype data. *Genetics* 155, 954-959.
- Puillandre, N., Lambert, A., Brouillet, S., Achaz, G., 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Mol Ecol* 21, 1864-1877.
- Ratnasingham, S., Hebert, P.D.N., 2007. bold: The Barcode of Life Data System (<http://www.barcodinglife.org>). *Molecular Ecology Notes* 7, 355-364.
- Raymond, M., Rousset, F., 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumnicism. *J. Heredity* 86, 248-249.
- Reaka-Kudla, M.L., Wilson, D., Wilson, E., 1997. The global biodiversity of the coral reefs: A comparison with rain forests. *Biodiversity II: Understanding and protecting our biological resources*. Joseph Henry Press, Washington, DC, pp. 83-108.
- Richmond, R.H., Hunter, C.L., 1990. Reproduction and recruitment of corals: comparisons among the Caribbean, the Tropical Pacific, and the Red Sea. *Marine Ecology Progress Series* 60, 185-203.
- Richmond, R.H., Jokiell, P.L., 1984. Lunar periodicity in larva release in the reef coral *Pocillopora damicornis* at Enewetak and Hawai'i. *Bulletin of Marine Science* 34, 280-287.
- Ridgway, T., Hoegh-Guldberg, O., Ayre, D.J., 2001. Panmixia in *Pocillopora verrucosa* from South Africa. *Marine Biology* 139, 175-181.
- Ridgway, T., Riginos, C., Davis, J., Hoegh-Guldberg, O., 2008. Genetic connectivity patterns of *Pocillopora verrucosa* in southern African Marine Protected Areas. *Marine Ecology Progress Series* 354, 161-168.
- Rinkevich, B., Shaish, L., Douek, J., Ben-Shlomo, R., 2016. Venturing in coral larval chimerism: a compact functional domain with fostered genotypic diversity. *Sci Rep* 6, 19493.

- Roberts, C.M., McClean, C.J., Veron, J.E.N., Hawkins, J.P., Allen, G.R., McAllister, D.E., Mittermeier, C.G., Schueler, F.W., Spalding, M., Wells, F., Vynne, C., Werner, T.B., 2002. Marine biodiversity hotspots and conservation priorities for tropical reefs. *Science* 295, 1280.
- Rowan, R., Powers, D.A., 1991. Molecular genetic identification of symbiotic dinoflagellates (zooxanthellae). *Mar. Ecol. Prog. Ser.* 71, 65-73.
- Rozenfeld, A.F., Arnaud-Haond, S., Hernández-García, E., Eguíluz, V.M., Matías, M.A., Serrao, E., Duarte, C.M., 2007. Spectrum of genetic diversity and networks of clonal organisms. *Journal of the Royal Society Interface* 4, 1093-1102.
- Rueffler, C., Van Dooren, T.J.M., Leimar, O., Abrams, P.A., 2006. Disruptive selection and then what? *Trends in Ecology & Evolution* 21, 238-245.
- Sale, P.F., Cowen, R.K., Danilowicz, B.S., Jones, G.P., Kritzer, J.P., Lindeman, K.C., Planes, S., Polunin, N.V.C., Russ, G.R., Sadovy, Y.J., Steneck, R.S., 2005. Critical science gaps impede use of no-take fishery reserves. *Trends in Ecology & Evolution* 20, 74-80.
- Samadi, S., Barberousse, A., 2006. The tree, the network, and the species. *Biological Journal of the Linnean Society* 89, 509-521.
- Sammarco, P.W., 1982. Polyp bail-out: an escape response to environmental stress and new means of reproduction in corals. *Marine Ecology Progress Series* 10, 57-65.
- Schlick-Steiner, B.C., Steiner, F.M., Seifert, B., Stauffer, C., Christian, E., Crozier, R.H., 2009. Integrative Taxonomy: A Multisource Approach to Exploring Biodiversity. *Annual Review of Entomology* 55, 421-438.
- Schluter, D., 2009. Evidence for Ecological Speciation and Its Alternative. *Science* 323, 737.
- Schmidt-Roach, S., Johnston, E., Fontana, S., Jury, C.P., Forsman, Z., 2014a. Daytime spawning of *Pocillopora* species in Kaneohe Bay, Hawai'i. *Galaxea* 16, 11-12.
- Schmidt-Roach, S., Lundgren, P., Miller, K.J., Gerlach, G., Noreen, A.M.E., Andreakis, N., 2013. Assessing hidden species diversity in the coral *Pocillopora damicornis* from Eastern Australia. *Coral Reefs* 32, 161-172.
- Schmidt-Roach, S., Miller, K.J., Lundgren, P., Andreakis, N., 2014b. With eyes wide open: a revision of species within and closely related to the *Pocillopora damicornis* species complex (Scleractinia; Pocilloporidae) using morphology and genetics. *Zoological Journal of the Linnean Society* 170, 1-33.
- Schoch, C.L., Seifert, K.A., Huhndorf, S., Robert, V., Spouge, J.L., Levesque, C.A., Chen, W., Fungal Barcoding, C., Fungal Barcoding Consortium Author, L., Bolchacova, E., Voigt, K., Crous, P.W., Miller, A.N., Wingfield, M.J., Aime, M.C., An, K.-D., Bai, F.-Y., Barreto, R.W., Begerow, D., Bergeron, M.-J., Blackwell, M., Boekhout, T., Bogale, M., Boonyuen, N., Burgaz, A.R., Buyck, B., Cai, L., Cai, Q., Cardinali, G., Chaverri, P., Coppins, B.J., Crespo, A., Cubas, P., Cummings, C., Damm, U., de Beer, Z.W., de Hoog, G.S., Del-Prado, R., Dentinger, B., Diéguez-Uribeondo, J., Divakar, P.K., Douglas, B., Dueñas, M., Duong, T.A., Eberhardt, U., Edwards, J.E., Elshahed, M.S., Fliiegerova, K., Furtado, M., García, M.A., Ge, Z.-W., Griffith, G.W., Griffiths, K., Groenewald, J.Z., Groenewald, M., Grube, M., Gryzenhout, M., Guo, L.-D., Hagen, F., Hambleton, S., Hamelin, R.C., Hansen, K., Harrold, P., Heller, G., Herrera, C., Hirayama, K., Hirooka, Y., Ho, H.-M., Hoffmann, K., Hofstetter, V., Högnabba, F., Hollingsworth, P.M., Hong, S.-B., Hosaka, K., Houbraken, J., Hughes, K., Huhtinen, S., Hyde, K.D., James, T., Johnson, E.M., Johnson, J.E., Johnston, P.R., Jones, E.B.G., Kelly, L.J., Kirk, P.M., Knapp, D.G., Kõljalg, U., Kovács, G.M., Kurtzman, C.P., Landvik, S., Leavitt, S.D., Liggenstoffer, A.S., Liimatainen, K., Lombard, L., Luangsa-ard, J.J., Lumbsch, H.T., Maganti, H., Maharachchikumbura, S.S.N., Martin, M.P., May, T.W., McTaggart, A.R., Methven, A.S., Meyer, W., Moncalvo, J.-M., Mongkolsamrit, S., Nagy, L.G., Nilsson, R.H., Niskanen, T., Nyilasi, I., Okada, G., Okane, I., Olariaga, I., Otte, J., Papp, T., Park, D., Petkovits, T., Pino-Bodas, R., Quaedvlieg, W.,

- Raja, H.A., Redecker, D., Rintoul, T.L., Ruibal, C., Sarmiento-Ramírez, J.M., Schmitt, I., Schüßler, A., Shearer, C., Sotome, K., Stefani, F.O.P., Stenroos, S., Stielow, B., Stockinger, H., Suetrong, S., Suh, S.-O., Sung, G.-H., Suzuki, M., Tanaka, K., Tedersoo, L., Telleria, M.T., Tretter, E., Untereiner, W.A., Urbina, H., Vágvölgyi, C., Vialle, A., Vu, T.D., Walther, G., Wang, Q.-M., Wang, Y., Weir, B.S., Weiß, M., White, M.M., Xu, J., Yahr, R., Yang, Z.L., Yurkov, A., Zamora, J.-C., Zhang, N., Zhuang, W.-Y., Schindel, D., 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences of the United States of America* 109, 6241-6246.
- Schweinsberg, M., Weiss, L.C., Striewski, S., Tollrian, R., Lampert, K.P., 2015. More than one genotype: how common is intracolony genetic variability in scleractinian corals? *Molecular Ecology* 24, 2673-2685.
- Shearer, T.L., Van Oppen, M.J.H., Romano, S.L., Wörheide, G., 2002. Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). *Molecular Ecology* 11, 2475-2487.
- Sherman, C.D.H., Ayre, D.J., Miller, K.J., 2005. Asexual reproduction does not produce clonal populations of the brooding coral *Pocillopora damicornis* on the Great Barrier Reef, Australia. *Coral Reefs* 25, 7-18.
- Sier, C.J.S., Olive, P.J.W., 1994. Reproduction and reproductive variability in the coral *Pocillopora verrucosa* from the Republic of Maldives. *Marine Biology* 118, 713-722.
- Simpson, G.G., 1951. The species concept. *Evolution* 5, 285-298.
- Slatkin, M., 1993. Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* 47, 264-279.
- Slatkin, M., 1995. A measure of population subdivision based on microsatellite allele frequencies. *Genetics* 139, 457-462.
- Smith, B.D., Zeder, M.A., 2013. The onset of the Anthropocene. *Anthropocene* 4, 8-13.
- Souter, P., Henriksson, O., Olsson, N., Grahn, M., 2009. Patterns of genetic structuring in the coral *Pocillopora damicornis* on reefs in East Africa. *BMC Ecol* 9, 19.
- Stimson, J., 1978. Mode and timing of reproduction in some common hermatypic corals of Hawaii and Enewetak. *Marine Biology* 48, 173-184.
- Stoddart, J.A., 1983. Asexual production of planulae in the coral *Pocillopora damicornis*. *Marine Biology* 76, 279-284.
- Stoeckel, S., Masson, J.-P., 2014. The exact distributions of F_{IS} under partial asexuality in small finite populations with mutation. *PLoS ONE* 9, e85228.
- Streelman, J.T., Danley, P.D., 2003. The stages of vertebrate evolutionary radiation. *Trends in Ecology & Evolution* 18, 126-131.
- Suzuki, G., 2012. Simultaneous spawning of *Pocillopora* and *Goniopora* corals in the morning time. *Galaxea, Journal of Coral Reef Studies* 14, 115-116.
- Tanner, J.E., 1996. Seasonality and lunar periodicity in the reproduction of Pocilloporid corals. *Coral Reefs* 15, 59-66.
- Templeton, A.R., 1998. Species and speciation: geography, population structure, ecology, and gene trees. In: Berlocher, D.J.H.a.S.H. (Ed.), *Endless Forms: Species and Speciation*. Oxford University Press, pp. 32-43.
- Tine, M., Kuhl, H., Gagnaire, P.-A., Louro, B., Desmarais, E., Martins, R.S.T., Hecht, J., Knaust, F., Belkhir, K., Klages, S., Dieterich, R., Stueber, K., Piferrer, F., Guinand, B., Bierne, N., Volckaert, F.A.M., Bargelloni, L., Power, D.M., Bonhomme, F., Canario, A.V.M., Reinhardt, R., 2014. European sea bass genome and its variation provide insights into adaptation to euryhalinity and speciation. *Nature Communications* 5, 5770.
- Torda, G., Lundgren, P., Willis, B.L., van Oppen, M.J., 2013a. Revisiting the connectivity puzzle of the common coral *Pocillopora damicornis*. *Mol Ecol* 22, 5805-5820.

- Torda, G., Lundgren, P., Willis, B.L., van Oppen, M.J.H., 2013b. Genetic assignment of recruits reveals short and long distance larval dispersal in *Pocillopora damicornis* on the Great Barrier Reef. *Molecular Ecology* 22, 5821–5834.
- Valdes, A.M., Slatkin, M., Freimer, N., 1993. Allele frequencies at microsatellite loci: the stepwise mutation model revisited. *Genetics* 133, 737-749.
- van Oppen, M.J., Willis, B.L., Vugt, H.W., Miller, D.J., 2000. Examination of species boundaries in the *Acropora cervicornis* group (Scleractinia, cnidaria) using nuclear DNA sequence analyses. *Mol Ecol* 9, 1363-1373.
- Van Oppen, M.J.H., Bongaerts, P.I.M., Underwood, J.N., Peplow, L.M., Cooper, T.F., 2011. The role of deep reefs in shallow reef recovery: an assessment of vertical connectivity in a brooding coral from west and east Australia. *Molecular Ecology* 20, 1647-1660.
- Van Valen, L., 1976. Ecological species, multispecies, and oaks. *Taxon* 25, 233-239.
- Verity, R., Nichols, R.A., 2016. Estimating the number of subpopulations (K) in structured populations. *Genetics* 203, 1827-1839.
- Veron, J.E.N., 1995. Corals in space and time: the biogeography and evolution of the Scleractinia. Comstock/Cornell.
- Veron, J.E.N., 2000. Corals of the world. Australian Institute of Marine Science.
- Veron, J.E.N., Pichon, M.M., 1976a. Scleractinia of Eastern Australia. Part I: Families Thamnasteriidae, Astrocoeniidae, Pocilloporidae. In: Science, A.I.o.M. (Ed.).
- Veron, J.J., Pichon, M.M., 1976b. Scleractinia of Eastern Australia. Part I: Families Thamnasteriidae, Astrocoeniidae, Pocilloporidae.
- Villanueva, R.D., Yap, H.T., Montaña, M.N.E., 2008. Timing of planulation by pocilloporids corals in the northwestern Philippines. *Marine Ecology Progress Series* 370, 111-119.
- Vitousek, P.M., Mooney, H.A., Lubchenco, J., Melillo, J.M., 1997. Human Domination of Earth's Ecosystems. *Science* 277, 494.
- Wegley, L., Edwards, R., Rodriguez-Brito, B., Liu, H., Rohwer, F., 2007. Metagenomic analysis of the microbial community associated with the coral *Porites astreoides*. *Environmental Microbiology* 9, 2707-2719.
- Weir, B., Cockerham, C., 1984. Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358-1370.
- Welch, D.B.M., Meselson, M., 2000. Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. *Science* 288, 1211.
- Willis, B.L., Babcock, R.C., Harrison, P.L., Wallace, C.C., 1997. Experimental hybridization and breeding incompatibilities within the mating systems of mass spawning reef corals. *Coral Reefs* 16, S53-S65.
- Wright, S., 1931. Evolution in mendelian populations. *Genetics* 16, 97-159.
- Wright, S., 1943. Isolation by distance. *Genetics* 28, 114-138.
- Yeoh, S.-R., Dai, C.-F., 2010. The production of sexual and asexual larvae within single broods of the scleractinian coral, *Pocillopora damicornis*. *Marine Biology* 157, 351-359.
- Yu, J.K., Wang, H.Y., Lee, S.C., Dai, C.F., 1999. Genetic structure of a scleractinian coral, *Mycidium elephantotus*, Taiwan. *Marine Biology* 133, 21-28.
- Yule, G.U., 1924. A mathematical theory of evolution, based on the conclusions of Dr. J.C. Willis, F.R.S. *Philosophical Transactions of the Royal Society London B* 213, 21-87.
- Yund, P.O., 2000. How severe is sperm limitation in natural populations of marine free-spawners? *Trends Ecol Evol* 15, 10-13.
- Zhang, J., Kapli, P., Pavlidis, P., Stamatakis, A., 2013. A general species delimitation method with applications to phylogenetic placements. *Bioinformatics* 29, 2869-2876.

ANNEXES

1. Isolation and characterization of 22 microsatellite loci from two coral species: *Acropora muricata* (Linnaeus, 1758) (Scleractinia, Acroporidae) and *Porites lutea* Milne-Edwards & Haime, 1851 (Scleractinia, Poritidae)

Pauline Gélín^{1,2}, Clément Rougeux^{1,3}, Vincent Mehn¹, Mireille M. M. Guillaume^{4,2,1}, J. Henrich Bruggemann^{1,2}, Hélène Magalon^{1,2,*}

L'Annexe 1 décrit le développement de marqueurs microsatellites pour deux espèces de coraux durs : *Acropora muricata* et *Porites lutea*. Cet article est publié dans *Conservation Genetic Resources*.

2. The fine-scale genetic structure of the malaria vectors *Anopheles funestus* and *Anopheles gambiae* (Diptera: Culicidae) in the north-eastern part of Tanzania

P. Gélín¹, H. Magalon¹, C. Drakeley^{2,3}, C. Maxwell^{2,4}, S. Magesa⁴, W. Takken⁵ and C. Boëte^{3,5,6*}

L'annexe 2 décrit la structuration génétique des populations de moustiques (*Anopheles funestus* et *Anopheles gambiae*) dans le nord-est de la Tanzanie. Cet article est publié dans *International Journal of Tropical Insect Science*.

3. One species for one island? Unexpected diversity and weak connectivity in a widely distributed tropical hydrozoan

Bautisse Postaire^{1,2,3}, Pauline Gelin^{1,2}, J. Henrich Bruggemann^{1,2} and Hélène Magalon^{1,2}

L'annexe 3 décrit la structuration génétique des populations d'une espèce hydrozoaire, *Lytocarpia brevirostris*, dans les océans Indien et Pacifique. Cet article est accepté pour publication dans *Heredity*.

Microsatellite records for volume 7, issue 4

© Springer Science+Business Media Dordrecht 2015

Electronic supplementary material The online version of this article (doi:[10.1007/s12686-015-0493-8](https://doi.org/10.1007/s12686-015-0493-8)) contains supplementary material, which is available to authorized users.

Species	Locus	Primer sequence (5'–3')	Repeat motif	Size range (bp)	Sample size (N)	Number of alleles (N _A)	H _o (H _e)
<i>Porites lutea</i> Milne-Edwards & Haime, 1851 (Scleractinia, Poritidae)	Am09 ^h	F: GAAGGCTCTTGTTGCGAT R: CGTTGATTTTGTGACTTCAA	(CA) ₈	113–129	46	4	0.43 (0.60) ^{NS}
	Am10 ^h	F: CTCCCAACAGATGCTATTTAAGAGA R: CGACCAACCAATAACCACTT	(AG) ₉	145–151	46	3	0.14 (0.37) ⁵
	Am11 ^h	F: GCAATGCCATGGTTTCCA R: TTGCAGCATCAAAGACCC	(TC) ₉	136–140	46	3	0.21 (0.20) ^{NS}
	P101 ^h	F: TCATTCAAATACCTTCTCAAGATTCA R: TGGTATTTTCATACATAATTTCCCTTG	(AG) ₁₁	232–276	48	7	0.53 (0.74) ^{NS}
	P102 ^h	F: GTCATCGTCATCACCATCCA R: GAGCCGAACAGATTTCAACC	(ACC) ₆	91–112	48	3	0.47 (0.54) ^{NS}
	P103 ^h	F: TTGCCCATTTCCAATAACTG R: GGAAAAGACGAAAATTAATAGCCC	(AAC) ₇	185–200	48	6	0.37 (0.59) ^{NS}
	P104 ^h	F: TTTTATGGAGTGGCAITTTG R: TTCAGTGTCAAAAAGTGCAAGAAA	(ACA) ₇	144–153	48	3	0.05 (0.40) ⁵
	P105 ^h	F: TGGTATATTATATTTCTGTTCCTTTT R: CTCGACATGAGCATTTGCCT	(TC) ₈	131–143	48	2	0.16 (0.49) ^{NS}
	P106 ^h	F: ACGTGCAGGTTGAATGTATGC R: GGTTTATAGATTCATCACTAACCAA	(GT) ₁₀	190–204	48	5	0.58 (0.71) ^{NS}
	P107 ^h	F: CAACCAACCACTTCGCTACA R: TACTGCTTCTCAACATGCGG	(TC) ₆	200–204	48	2	0.05 (0.15) ^{NS}
	P108 ^h	F: GACCGTGATCAAAGGCTTA R: CGAAGAACCAGAAACAGGAAG	(CT) ₇	182–226	48	3	0.05 (0.15) ^{NS}
P109 ^h	F: TTGACTTTTGGGCTTGAAA R: CACGACTACCGGGTGAAGTT	(AG) ₇	179–181	48	2	0.05 (0.05) ^{NS}	
P110 ^h	F: CACCATAATCATGAGATTTACTAATTGA R: GAATCAACCAATGGCAGTCC	(AC) ₉	123–143	48	4	0.53 (0.73) ^{NS}	
P111 ^h	F: GCTGTTCAATGCATTTCTGG R: GCTGTGAGCTTTAAACCGACC	(AC) ₉	141–149	48	2	0.32 (0.51) ^{NS}	
<i>Pelteobagrus vachelli</i>	PV5 ⁱ	F: GGACTGGAACAACAACAGGCT R: CCCAAACCGCACAITATTTT	(CA) ₁₀	214–222	32	5	0.688 (0.581)
	KP735111 ⁱ	F: GCTTTTATCATGCAGTGTGG R: GAAAATTCAGCATTTCCAAGC	(GA) ₁₀	258–278	32	8	0.938 (0.802)
	PV11 ⁱ	F: CAGGCTAACCTTGGCTTGAC R: CAGTCCCAGTCAGCTCTGC	(GT) ₉	275–281	32	4	0.563 (0.537)
	KP735112 ⁱ	F: CCGTACACAGTTGCTCTCCA R: CAAAAACACTCCAGCAGTCCA	(AC) ₉	227–245	32	10	0.844 (0.782)
	PV19 ⁱ	F: CTGTACACTCGCCTTCAAA R: TTACTCGCAGTGCATTTTGG	(GT) ₉	251–255	32	3	0.188 (0.177)
	KP735113 ⁱ						
	PV20 ⁱ						
	KP735114 ⁱ						
	PV22 ⁱ						
	KP735115 ⁱ						

Species	Locus	Primer sequence (5'–3')	Repeat motif	Size range (bp)	Sample size (N)	Number of alleles (N_A)	H_o (H_e)
	PV23 ⁱ	F: GCTGAGAACTGAAACCCTGC R: ACATGGACCTCCCATTTGTGT	(GA) ₉	241–251	32	6	0.688 (0.672)
	KP735116 ⁱ	F: ITTCAATGGCCCTCCGTGATTT R: TTCTTCCCTTTTCCCTCGT	(CT) ₉	273–283	32	6	0.563 (0.711)
	PV28 ⁱ	F: ITTCAATGGCCCTCCGTGATTT R: TTCTTCCCTTTTCCCTCGT	(TA) ₉	279–291	32	7	0.844 (0.712)
	KP735117 ⁱ	F: GCCAAGACGACATCACAGAA R: TGTGCCAGGTGTGTACGTT	(TA) ₈	237–245	32	5	0.438 (0.643)
	PV30 ⁱ	F: GCCAAGACGACATCACAGAA R: TGTGCCAGGTGTGTACGTT	(TC) ₈	220–222	32	2	0.031 (0.031)
	KP735118 ⁱ	F: TCGGTCAGTTCAGATCACCA R: CACACAACACAAATCTGCCAA	(GT) ₈	285–315	32	12	0.625 (0.79)
	PV34 ⁱ	F: CACACAACACAAATCTGCCAA R: GATGAAAAGAAACCCGGAAACA	(AC) ₈	251–257	32	4	0.5 (0.563)
	KP735119 ⁱ	F: GATGAAAAGAAACCCGGAAACA R: TGGAGAGAAAGAAAGGCCAGA	(AC) ₈	258–268	32	5	0.75 (0.655)
	PV35 ⁱ	F: TGGAGAGAAAGAAAGGCCAGA R: CAACGACTGGAGGTCAAACA	(GT) ₈	258–260	32	2	0.125 (0.222)
	KP735120 ⁱ	F: CAACGACTGGAGGTCAAACA R: AGCTGACGGAACTGTTACTGC	(TG) ₈	275–281	32	3	0.344 (0.304)
	PV36 ⁱ	F: AGCTGACGGAACTGTTACTGC R: TGGATTGTTGAGGTTGGA	(TG) ₈	257–261	32	3	0.188 (0.294)
	KP735121 ⁱ	F: TGAATAGCTTGCCGTTGTG R: GGGATTAGGTTGAGGAGC	(AC) ₈	249–279	32	13	0.813 (0.864)
	PV37 ⁱ	F: TGGATTGTTGAGGTTGGA R: CCCTAAGTTGAGCCCTTTA	(AC) ₈	272–278	32	4	0.563 (0.586)
	KP735122 ⁱ	F: GGGATTAGGTTGAGGAGC R: TCACTACCCGGGACTGACTT	(AA) ₈	268–289	32	7	0.938 (0.748)
	PV39 ⁱ	F: TCACTACCCGGGACTGACTT R: TAAAAATTACCCGGCCATTC	(TCA) ₇	231–237	32	3	0.125 (0.121)
	KP735123 ⁱ	F: TAAAAATTACCCGGCCATTC R: AGCATTGGAGCCGATCATAC	(CAT) ₇	254–260	32	3	0.125 (0.354)
	PV40 ⁱ	F: AGCATTGGAGCCGATCATAC R: AGGCACCCCTCAGTAAGACA	(TCC) ₆	230–233	32	2	0.406 (0.484)
	KP735124 ⁱ	F: AGGCACCCCTCAGTAAGACA R: AGCAGCTCTTCAACCGTTA	(TTG) ₆	228–243	32	6	0.938 (0.79)
	PV43 ⁱ	F: AGCAGCTCTTCAACCGTTA R: CAGGTAGAGTGGGAGTGATGG	(ATG) ₆	218–220	32	2	0.781 (0.484)
	KP735125 ⁱ	F: CAGGTAGAGTGGGAGTGATGG R: TCAAACTGGGCAAAAACCTCC	(GTT) ₆	259–268	32	3	0.594 (0.549)
	PV44 ⁱ	F: TCAAACTGGGCAAAAACCTCC R: AACGAAAGGGGTTTTTCAGAT					
	KP735126 ⁱ	F: AACGAAAGGGGTTTTTCAGAT R: ITTCTGAAAGGAGACGTCGG					
	PV45 ⁱ	F: ITTCTGAAAGGAGACGTCGG R: CCAGGCAGGATGGTTACCT					
	KP735127 ⁱ	F: CCAGGCAGGATGGTTACCT R: CTTCCCAGAAGTCTGAACGG					
	PV49 ⁱ	F: CTTCCCAGAAGTCTGAACGG R: TCCAGGATCAGGACATCACACA					
	KP735128 ⁱ	F: TCCAGGATCAGGACATCACACA R: CCACTGGCAAACAATTGAAA					
	PV50 ⁱ	F: CCACTGGCAAACAATTGAAA R: TCAGACCCGCTTAATATGC					
	KP735129 ⁱ	F: TCAGACCCGCTTAATATGC R: ATGATCAAACAAGTGGGCA					
	PV65 ⁱ	F: ATGATCAAACAAGTGGGCA R: AGAAGCGGGAGGAGAGAGAG					
	KP735130 ⁱ	F: AGAAGCGGGAGGAGAGAGAG R: TGCTTCTCTCCGCTAATGT					
	PV71 ⁱ	F: TGCTTCTCTCCGCTAATGT R: GCAGCCACTAATATGAGCCC					
	KP735131 ⁱ	F: GCAGCCACTAATATGAGCCC R: CTGGGAAACATCCTGGAAA					
	PV84 ⁱ	F: CTGGGAAACATCCTGGAAA R: AACCATCTGCAGGTGAGAC					
	KP735132 ⁱ	F: AACCATCTGCAGGTGAGAC R: CACATGATCGTCAACCTCGTC					
	PV86 ⁱ	F: CACATGATCGTCAACCTCGTC R: TGAGGATGATTTCTGCACCTG					
	KP735133 ⁱ	F: TGAGGATGATTTCTGCACCTG R: ITCCCTAAAATGACCTCGTGC					
	PV88 ⁱ	F: ITCCCTAAAATGACCTCGTGC R: ITCCCTAAAATGACCTCGTGC					
	KP735134 ⁱ	F: ITCCCTAAAATGACCTCGTGC R: ITCCCTAAAATGACCTCGTGC					
	PV90 ⁱ	F: ITCCCTAAAATGACCTCGTGC R: ITCCCTAAAATGACCTCGTGC					

Isolation and characterization of 22 microsatellite loci from two coral species: *Acropora muricata* (Linnaeus, 1758) (Scleractinia, Acroporidae) and *Porites lutea* Milne-Edwards & Haime, 1851 (Scleractinia, Poritidae)

Pauline Gélín^{1,2}, Clément Rougeux^{1,3}, Vincent Mehn¹, Mireille M. M. Guillaume^{4,2,1}, J. Henrich Bruggemann^{1,2}, Hélène Magalon^{1,2,*}

¹UMR ENTROPIE Université de La Réunion/CNRS/IRD 9220, 15 avenue René Cassin, CS 92003, 97744 St Denis Cedex 09, La Réunion, France

²Laboratory of Excellence CORAIL, www.labex-corail.fr

³present address: GIROQ, Département de Biologie, Pavillon Vachon, Université Laval, Sainte-Foy, Québec, Canada G1K 7P4

⁴Département MPA, UMR BOrEA CNRS-MNHN-UPMC-IRD-UCBN-UAG, Museum National d'Histoire Naturelle, 43 rue Cuvier, 75005 Paris, France

Keywords: microsatellite, scleractinia, coral, reef, *Acropora muricata*, *Porites lutea*

*Corresponding author: Hélène Magalon, helene.magalon@univ-reunion.fr, Phone: +262 (0)262 93 86 76; Fax: +262 (0)262 93 86 85; UMR ENTROPIE, Université de La Réunion, Faculté des sciences et technologies, 15 avenue René Cassin, CS 92003, 97744 St Denis Cedex 09, La Réunion, France

Total genomic DNA of eight colonies for each of the two species, *Acropora muricata* and *Porites lutea*, sampled on the west coast of Reunion Island, was isolated using DNeasy Blood & Tissue kit (Qiagen™) following the manufacturer's instructions and sent to GenoScreen, Lille, France (www.genoscreen.fr). We did not perform any separation of coral and zooxanthellae tissues, thus we extracted both DNA simultaneously; markers specificity was checked *a posteriori*. One µg was used for the development of microsatellite libraries through 454 GS-FLX Titanium pyrosequencing of enriched DNA libraries as described in Malausa et al. (2011). A total of 6840 and 5824 sequences containing microsatellite motifs were identified using the software QDD (Meglecz et al. 2010) and 532 and 529 primer pairs recognized for *A. muricata* and *P. lutea*, respectively. For each of the two species, 96 primer pairs were selected for amplification test depending on the motif, the number of repeats (≥ 5) and the product size (>100 pb). Amplification tests were performed on 4 individuals and revealed on agarose gel (2%). The 52 and 40 primer pairs amplified for *A. muricata* and *P. lutea* respectively, were tested for polymorphism on 8 individuals for each species on an ABI 3730 XL (Gentyane Plateforme, INRA,

Clermont-Ferrand, France). To assess their specificity (coral or zooxanthellae), the 35 remaining markers for *A. muricata* and the 25 ones for *P. lutea* were then tested on DNA extracted from zooxanthellae cultures isolated from the 8 colonies used to build the library. Finally, 11 coral-specific microsatellite loci were isolated from each of the two species (from Am01 to Am11 for *A. muricata* and from Pl01 to Pl11 for *P. lutea*).

Characterization of the new developed loci was performed by genotyping two populations from the west coast of Reunion Island: one in Saint-Leu (21°10'51.14"S, 55°17'10.34"E) for *Acropora muricata* (n = 46) and one in Etang-Salé (21°16'13.31"S, 55°19'56.48"E) for *Porites lutea* (n = 48). PCR amplifications were performed in 10 µL containing 5 µL of MasterMix 2x (Applied Biosystems), 2.5 µL of ultra-pure water, 0.25 µL of each primer (10 µM) and 2 µL of DNA (10 ng/µL). The thermocycling program was: 94°C for 5 min + 7 x (94°C for 30 s, 62°C [-1°C at each cycle] for 30 s, 72°C for 30 s) + 30 x (94°C for 30 s, 55°C for 30 s, 72°C for 30 s) + 72°C for 5 min.

Diversity indices, Hardy-Weinberg equilibrium and linkage disequilibrium were assessed using Arlequin v 3.5.1.2. No linkage disequilibrium was found after FDR corrections.

The development of these loci will be very useful in studying the genetic diversity and gene flow between the populations of the staghorn coral *A. muricata* and the massive coral *P. lutea*, both remaining poorly documented.

Acknowledgments

The authors want to thank Vianney Denis and Lionel Bigot for assistance in collecting samples and stimulating discussions and master students Gaëlle Perazio, Laura Benestan, Juliette Fremery and Christophe Demichelis for their help in this study. Acknowledgments also go to the Gentyane Plateforme team (INRA, Clermont-Ferrand, France) and Mayalen Zubia for assisting us in zooxanthellae culturing.

Accession Numbers

Primer sequences are deposited on GenBank with accession numbers from KP298675 to KP298685 for *Porites lutea* and from KP298686 to KP298696 for *Acropora muricata*

Electronic Supplementary Material

The microsatellite sequences are available as electronic supplementary material (ESM_1.pdf)

Compliance with Ethical Standards

Research was conducted with permission of the regional authorities of marine affairs (decisions 53/DMSOI/2011 and 15-2013/DEAL/SEB/UBIO) and the marine park authorities of Reunion Island (RNMR).

The authors declare that they have no conflict of interest.

References

- Malausa T, Gilles A, Meglecz E, Blanquart H, Duthoy S, Costedoat C, Dubut V, Pech N, Castagnone-Sereno P, Delye C, Feau N, Frey P, Gauthier P, Guillemaud T, Hazard L, Le Corre V, Lung-Escarmant B, Male PJ, Ferreira S, Martin JF (2011) High-throughput microsatellite isolation through 454 GS-FLX Titanium pyrosequencing of enriched DNA libraries. *Mol Ecol Resour* 11:638-644
- Meglecz E, Costedoat C, Dubut V, Gilles A, Malausa T, Pech N, Martin JF (2010) QDD: a user-friendly program to select microsatellite markers and design primers from large sequencing projects. *Bioinformatics* 26:403-404
- Tang P-W, Wei NV, Chen C-W, Wallace CC, Chen CA (2010) Comparative study of genetic variability of AAT and CT/GT microsatellites in staghorn coral, *Acropora* (Scleractinia: Acroporidae). *Zool Stud* 49:657-668

The fine-scale genetic structure of the malaria vectors *Anopheles funestus* and *Anopheles gambiae* (Diptera: Culicidae) in the north-eastern part of Tanzania

P. Gélín¹, H. Magalon¹, C. Drakeley^{2,3}, C. Maxwell^{2,4}, S. Magesa⁴,
W. Takken⁵ and C. Boëte^{3,5,6*}

¹UMR ENTROPIE Université de La Réunion/IRD/CNRS, Faculté des Sciences et Technologies, Université de La Réunion, 15 Bd René Cassin, CS 92003, 97744 St Denis Cedex 09, La Réunion, France; ²Faculty of Infectious and Tropical Diseases, London School of Hygiene and Tropical Medicine, WC1E 7HT, London, UK; ³Joint Malaria Programme, PO Box 2228 Moshi, Tanzania; ⁴Amani Medical Research Centre, National Institute for Medical Research, PO Box 81, Muheza, Tanzania; ⁵Laboratory of Entomology, Wageningen University and Research Centre, PO Box 16, 6700, AA Wageningen, the Netherlands; ⁶UMR “Emergence des Pathologies Virales”, EPV Aix-Marseille Université, IRD 190 - Inserm 1207 - EHESP, 27 Bd Jean Moulin, 13005, Marseille, France

(Accepted 17 May 2016)

Abstract. Understanding the impact of altitude and ecological heterogeneity at a fine scale on the populations of malaria vectors is essential to better understand and anticipate eventual epidemiological changes. It could help to evaluate the spread of alleles conferring resistance to insecticides and also determine any increased entomological risk of transmission in highlands due to global warming. We used microsatellite markers to measure the effect of altitude and distance on the population genetic structure of *Anopheles funestus* and *Anopheles gambiae* s.s. in the Muheza area in the north-eastern part of Tanzania (seven loci for each species). Our analysis reveals strong gene flow between the different populations of *An. funestus* from lowland and highland areas, as well as between populations of *An. gambiae* sampled in the lowland area. These results highlight for *An. funestus* the absence of a significant spatial subpopulation structuring at small-scale, despite a steep ecological and altitudinal cline. Our findings are important in the understanding of the possible spread of alleles conferring insecticide resistance through mosquito populations. Such information is essential for vector control programmes to avoid the rapid spread and fixation of resistance in mosquito populations.

Key words: *Anopheles funestus*, *Anopheles gambiae*, altitude, malaria, microsatellite, population genetics, Tanzania

Introduction

Environmental heterogeneity is known to be a driver of population structure (Temunović *et al.*,

2012; Stein *et al.*, 2014), and malaria mosquitoes are no exception (Touré *et al.*, 1994). Without doubt, altitude is a major parameter affecting it because of the abiotic (temperature and humidity) as well as the biotic conditions, including the lower presence of pathogens such as *Plasmodium* spp., as shown in

*E-mail: cboete@gmail.com

bird populations (van Rooyen *et al.*, 2013). As global warming is significantly affecting species distribution as well as their mutual interactions, altitude can be considered as a proxy to understand its impact on the structure of populations of terrestrial insects (Hodkinson, 2005), including mosquitoes. In the case of malaria, the large increase in household ownership of long-lasting insecticidal nets (LLINs) has contributed to an important decrease in the number of malaria cases in the last decade (WHO, 2014), but at the same time, this is associated with a greater risk of the spread of resistance to pyrethroids, the sole insecticidal class used for LLIN treatment (Ranson *et al.*, 2011). Understanding, at a fine scale, how resistance can spread in mosquito populations is of critical importance to optimize current and future vector control. Genetic data can, not only help obtain this information, but can also provide a better understanding of the role of connectivity between populations on the dynamics of disease transmission (Donnelly *et al.*, 2004). Finally, and on a more hypothetical aspect of the control of vector-borne diseases, it is a prerequisite for any release of genetically modified mosquitoes or sterile insects aimed to control the disease at a large scale.

Along the coast of Tanzania, the major malaria vectors are *Anopheles funestus* and *Anopheles gambiae* s.s. (Magesa *et al.*, 1991; Temu *et al.*, 1998; Mboera and Magesa, 2001; Maliti *et al.*, 2014). Here, we present a comparative study of the population genetics of these two major vectors from the Muheza district in the north-eastern part of Tanzania, to investigate the spatial population structure of these two anopheline species. This region is a rural area known for its holoendemic transmission of malaria in the lowland area (Mboera and Magesa, 2001; Alilio *et al.* 2004), and it also presents a steep ecological cline with the presence of highland areas close to the town of Muheza and in the vicinity of the Amani Nature Reserve in the Tanga Region of Tanzania, East Africa.

Materials and methods

Collection site

Adult mosquitoes were collected in Muheza district, Tanga Region in the north-eastern part of Tanzania. Six sites were chosen that could be organized in two clusters according to their altitudes: the lowland cluster (mean altitude = 236 m \pm 11.8 m) included three villages (Mamboleo, Songa Kibaoni, Zeneth), and the highland area (mean altitude = 952 m \pm 63.2 m) with three villages (Mikwinini, Mlesa, Ndola) (Fig. 1 and Table 1). Pairwise distances between villages inside each cluster ranged from 5 to 10 km (Table 2).

Mosquito sampling, species identification and oocysts collection

The collections were performed between May 2005 and August 2005 where indoor-resting blood-fed anopheline females were collected with aspirators in the morning regularly during this period, i.e. between 15 and 19 times in the lowland and the highland areas, respectively. At each visit, the same four houses were visited for 15 to 20 min in each village. Mosquitoes were brought to the insectary in Muheza, fed with a sugar solution *ad libitum*, and maintained for 7 days in small pots. On the seventh day after capture, midguts of the surviving mosquitoes were dissected with tweezers under a light microscope and the presence/absence of *Plasmodium falciparum* oocysts recorded. All the collected mosquitoes did not survive the period of 7 days in the insectary. Performing an analysis of the mortality rate neither revealed any significant differences between areas and species, nor their interactions (see Additional File 1). This reveals the absence of any bias in the mortality. Similarly, a number of mosquitoes survived the period of 7 days in the insectary, but they were untested for genotyping (conservation issues). Again the statistical analysis (see Additional File 2) revealed no significant difference between the groups, and this did not affect the survey.

Morphological identification based on the taxonomic keys for the identification of Afrotropical mosquitoes (Gillies and de Meillon, 1968; Gillies and Coetzee, 1987) was done to identify *An. funestus* group and *An. gambiae* s.l. DNA extraction from whole insects was performed using the DNeasy Blood and Tissue kit (Qiagen) according to the manufacturer's instructions and DNA concentration was measured with a Nanodrop[®] spectrophotometer. Identification of subspecies inside each complex was performed by polymerase chain reaction (PCR) according to a protocol based on Koekemoer *et al.* (2002) within the *An. funestus* group and Wilkins *et al.* (2006) within the *An. gambiae* complex. Sampling details are summarized in Table 1.

Microsatellite genotyping

A total of 142 *An. funestus* were genotyped using 10 microsatellite loci: AF2, AF3, AF5, AF19, AF20 (Sinkins *et al.*, 2000), FunF, FunG, FunL, FunO, FunR (Cohuet *et al.*, 2002). A total of 181 *An. gambiae* individuals were genotyped using 13 microsatellite loci: 11 loci from Zheng *et al.* (1996) [H29, H46, H88, H131, H143, H197, H249, H555, H675, H678, H1D1 (abbreviated names from Midega *et al.* (2010))] and two microsatellite loci from the NOS and cecropin 3' UTR genes (Luckhart *et al.*, 2003), which function in the innate immune system of insects.



Fig. 1. Map of the area where collections of malaria mosquitoes were performed.

Table 1. Number of collected adult (*Anopheles funestus* and *An. gambiae*) surviving 7 days after collection and their infection status

Species	Village	Latitude	Longitude	Altitude (m)	Area of collection (altitude)	Number of collected mosquitoes	Number of surviving mosquitoes after 7 days	Number of mosquitoes used in the genetic analysis
<i>An. funestus</i>	Ndola	S 05°06.400'	E 038°34.325'	1060 m	Highland	129	40	28
	Mlesa	S 05°07.921'	E 038°37.365'	845 m	Highland	16	9	4
	Mikwinini	S 05°08.432'	E 038°35.164'	989 m	Highland	1	0	–
	Mamboleo	S 05°13.123'	E 038°43.244'	220 m	Lowland	8	3	3
	Songa Kibaoni	S 05°15.204'	E 038°38.639'	259 m	Lowland	20	6	4
	Zeneth	S 05°13.492'	E 038°39.620'	229 m	Lowland	616	176	103
<i>An. gambiae</i>	Ndola	S 05°06.400'	E 038°34.325'	1060 m	Highland	42	14	–
	Mlesa	S 05°07.921'	E 038°37.365'	845 m	Highland	11	8	–
	Mamboleo	S 05°13.123'	E 038°43.244'	220 m	Lowland	126	66	43
	Songa Kibaoni	S 05°15.204'	E 038°38.639'	259 m	Lowland	66	34	27
	Zeneth	S 05°13.492'	E 038°39.620'	229 m	Lowland	366	215	102

Table 2. Matrix of pairwise geographic distance amongst sampled sites. Distances are expressed in km

	Highland		Lowland		
	Ndola	Mlesa	Zeneth	Songa Kibaoni	Mamboleo
Mlesa	6.0	–			
Zeneth	16.4	11.3	–		
Songa Kibaoni	17.9	13.5	3.5	–	
Mamboleo	20.5	15.5	6.5	9.2	–

For each species, the forward primer of each locus was labelled with a fluorescent dye using a M13 labelled tail added to the 5'-end of the oligonucleotide (four colours: FAM, VIC, NED, PET). PCR were performed in a volume of 10 µl with 5 µl of MasterMix 2x (Applied BiosystemsTM; Thermo Fisher Scientific Inc., Waltham, Massachusetts, USA), 0.25 µl of each primer (10 µM) and 0.25 µl of labelled M13 tail (10 µM), 2 µl of DNA template (10 ng/µl) using the following cycling: 2 min at 95 °C; followed by 30 cycles of 30 sec at 95°C, 30 sec at 55 °C and 30 sec at 72 °C); then 5 min at 95 °C. Mixes of PCR products were then made according to amplicon size and dye, and genotyping was performed using an ABI PRISM 3730 XL DNA sequencer (Applied BiosystemsTM). Results were analysed with GeneMapperTM v. 4.0 software (Applied BiosystemsTM).

Data analyses

Estimates of linkage disequilibrium were performed using FSTAT v. 2.9.3 (Goudet, 1995). The presence of null alleles was assessed using MICRO-CHECKER v. 2.2.3 (Van Oosterhout *et al.*, 2004).

Genetic diversity of mosquito populations was assessed by the number of alleles (N_a), observed

heterozygosity (H_o), and expected heterozygosity (H_e) using Arlequin v. 3.5.1.2 (Excoffier *et al.*, 2005). Tests of deviation from Hardy–Weinberg equilibrium (HWE) were performed using FSTAT v. 2.9.3 (Goudet, 1995).

The significance of genetic differentiation between populations based on allelic distribution across populations was examined using a Fisher exact test with FSTAT v. 2.9.3 (Goudet, 1995). The pairwise F_{ST} statistic between populations was calculated after Weir and Cockerham (1984) using FSTAT v. 2.9.3 (Goudet, 1995) and Arlequin v. 3.5.1.2 (Excoffier *et al.*, 2005).

For each species, without taking into account their geographic origin, *An. funestus* and *An. gambiae* individuals were clustered on the basis of their genetic relatedness using the Bayesian clustering approach implemented in STRUCTURE v. 2.3.3 (Pritchard *et al.*, 2000; Falush *et al.*, 2003). Simulations were performed using the admixture and correlated frequencies model. We estimated the number K of genetic clusters (here between $K = 1$ and $K = 4$) to which the individuals should be assigned. For all simulations, we did not force the model with predefined allele frequencies for source clusters. Five independent runs were conducted to assess the consistency of the results across runs,

Table 3. Summary statistics of *Anopheles funestus* populations

Population	N	N_a	H_O	H_E	F_{IS}	r
Highland						
AF2	30	9	0.600	0.785	0.239	0.11 (yes)
AF19	31	6	0.645	0.555	-0.165	-0.14 (no)
AF20	29	3	0.552	0.639	0.138	0.07 (no)
FunF	28	4	0.429	0.442	0.031	-0.01 (no)
FunG	30	4	0.733	0.629	-0.170**	-0.13 (no)
FunO	31	5	0.484	0.526	0.081	0.03 (no)
FunR	32	4	0.406	0.496	0.183	0.08 (no)
All	32	5 ± 2	0.550 + 0.120	0.583 + 0.110	0.056	
Lowland						
AF2	110	11	0.718	0.824	0.129**	0.06 (yes)
AF19	107	6	0.411	0.503	0.183**	0.10 (yes)
AF20	106	4	0.396	0.542	0.268***	0.12 (yes)
FunF	107	5	0.579	0.603	0.040	0.01 (no)
FunG	106	6	0.651	0.725	0.102***	0.04 (no)
FunO	108	6	0.565	0.569	0.008	0.00 (no)
FunR	105	4	0.305	0.505	0.400***	0.18 (yes)
All	110	6 ± 2.38	0.518 + 0.150	0.610 + 0.121	0.151**	

N : number of amplified individuals; N_a : number of alleles; H_O : observed heterozygosity; H_E : expected heterozygosity under Hardy–Weinberg equilibrium; F_{IS} : inbreeding coefficient. Next to the F_{IS} is indicated the significance of the P -values for deviation to Hardy–Weinberg equilibrium: *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$; r : Null allele frequencies (presence or absence is indicated next by yes or no).

and all runs were based on 10^6 iterations after a burn-in period of 10^5 iterations. We then identified the number of genetically homogeneous clusters as described by Evanno *et al.* (2005).

As analyses using STRUCTURE are based on strong assumptions (Panmixia, no linkage disequilibrium), we performed a dissimilarity analysis calculating the Shared Allele Distance (DAS) between pairs of genotypes and drawing a neighbour-joining tree with DARwin v 5.0 (Perrier and Jacquemoud-Collet, 2006).

Results

Species identification

Of the 142 *An. funestus* individuals, all were identified as *An. funestus sensu stricto* following PCR identification tests. Amongst 181 *An. gambiae* individuals, 172 were identified as *An. gambiae sensu stricto* and nine as *An. arabiensis*. *Anopheles arabiensis* were excluded from further analyses.

Linkage disequilibrium and test for Hardy–Weinberg equilibrium

Due to a large number of missing data for some of the loci, which may be due to null alleles, data from seven microsatellite loci from 10 were retained for further analyses for *An. funestus* (AF2, AF19, AF20, FunF, FunG, FunO and FunL). Amongst 21

linkage equilibrium tests, none was significant at the 5% level after Bonferroni correction for multiple testing. As such, these seven loci were considered as independent and kept for further analyses. Because of small sampling size for some locations and low *Plasmodium* prevalence over the dates of collection, all individuals were pooled into two populations: ‘Highland’ ($n = 32$) and ‘Lowland’ ($n = 110$).

For *An. gambiae*, due to a large number of missing data for some of the loci, which may also be due to null alleles, data from seven microsatellite loci from 13 were retained for further analyses (H29, H88, H131, H249, H678, H1D, NOS). Amongst 21 linkage equilibrium tests, none was significant at the 5% level after Bonferroni correction for multiple testing. As such, these seven loci were considered as independent and kept for further analyses. Due to low *Plasmodium* prevalence, infected individuals were pooled with non-infected for each village. For *An. gambiae*, the term ‘population’ referred to individuals sampled in the same village during the time of the experiment.

Regarding the two *An. funestus* populations, the highland population did not show deviation from Hardy–Weinberg equilibrium over all loci, whilst the lowland population showed a significant excess of homozygotes ($F_{IS} = 0.151^{**}$; Table 3). Additionally, the three populations of *An. gambiae* showed significant excess of homozygotes over all loci (F_{IS} ranged from 0.162** to 0.179**, Table 4).

Table 4. Summary statistics of *Anopheles gambiae* populations

Population	<i>N</i>	<i>N_a</i>	<i>H_O</i>	<i>H_E</i>	<i>F_{IS}</i>	<i>r</i>
Zeneth						
H29	98	6	0.255	0.320	0.205**	0.08 (yes)
H88	99	7	0.364	0.656	0.436***	0.20 (yes)
H249	100	7	0.630	0.686	0.077**	0.03 (no)
H1D1	101	2	0.465	0.487	0.044	0.02 (no)
H131	98	6	0.633	0.728	0.130*	0.06 (yes)
H678	100	11	0.790	0.806	0.021	0.01 (no)
NOS	100	9	0.300	0.511	0.414***	0.19 (yes)
All	102	6.8 ± 2.8	0.491 ± 0.200	0.598 ± 0.167	0.178**	
Songa Kibaoni						
H29	26	3	0.308	0.370	0.172	0.07 (no)
H88	22	6	0.455	0.644	0.285*	0.13 (no)
H249	24	7	0.583	0.682	0.136	0.06 (no)
H1D1	27	2	0.444	0.453	0.019	0.00 (no)
H131	21	5	0.571	0.723	0.223	0.11 (no)
H678	23	9	0.739	0.778	0.051	0.00 (no)
NOS	24	6	0.250	0.430	0.424*	0.16 (yes)
All	27	5.4 ± 2.4	0.479 ± 0.169	0.583 ± 0.162	0.179**	
Mamboleo						
H29	43	3	0.302	0.285	−0.062	−0.02 (no)
H88	43	7	0.372	0.554	0.331**	0.14 (yes)
H249	41	7	0.585	0.778	0.248	0.12 (yes)
H1D1	43	2	0.488	0.502	0.026	0.01 (no)
H131	40	7	0.525	0.682	0.242	0.11 (yes)
H678	41	8	0.756	0.813	0.071	0.03 (no)
NOS	41	5	0.341	0.396	0.140	0.06 (no)
All	43	5.6 ± 2.3	0.482 ± 0.159	0.573 ± 0.196	0.162**	

N: number of amplified individuals; *N_A*: number of alleles; *H_O*: observed heterozygosity; *H_E*: expected heterozygosity under Hardy–Weinberg equilibrium; *F_{IS}*: inbreeding coefficient. Next to the *F_{IS}* is indicated the significance of the *P*-values for deviation to Hardy–Weinberg equilibrium: *, *P* < 0.05; **, *P* < 0.01; ***, *P* < 0.001; *r*: Null allele frequencies (presence or absence is indicated next by yes or no).

Genetic diversity and population differentiation

The number of alleles per locus ranged from four (AF20) to 12 (AF2) for *An. funestus* and from two (H1D1) to 11 (H678) for *An. gambiae* (Tables 3 and 4, respectively).

Tests of genetic differentiation between populations were performed for each species. For *An. funestus*, the highland population was undifferentiated from the lowland population (*F_{ST}* = 0.012, *P* = 0.05; Fisher exact test not significant). Moreover, allelic richness of the highland population was not significantly different from the lowland population (permutation test, *P* = 0.67).

For *An. gambiae*, the allelic richness of the three populations was not significantly different from each other (permutation test, all *P* > 0.05). *F_{ST}* estimates were not significantly different from 0 (Table 5). Thus, no genetic differentiation was found between the three populations, nor when considering only the NOS locus implied in the

Table 5. Pairwise multilocus estimates of *F_{ST}* for *Anopheles gambiae* populations. All comparisons for Fisher exact tests were not significant at the 0.05 level

	Mamboleo	Songa Kibaoni
Songa Kibaoni	0.003 ^{NS}	–
Zeneth	0.001 ^{NS}	0.009 ^{NS}

immune response (*F_{ST}* ranged from 0.001 to 0.018, all *P* > 0.05).

Assignment tests

A Bayesian clustering analysis was performed using STRUCTURE. For both *An. funestus* and *An. gambiae*, the clustering analysis indicated that the posterior distribution of the allele frequencies amongst clusters was best explained with a grouping into

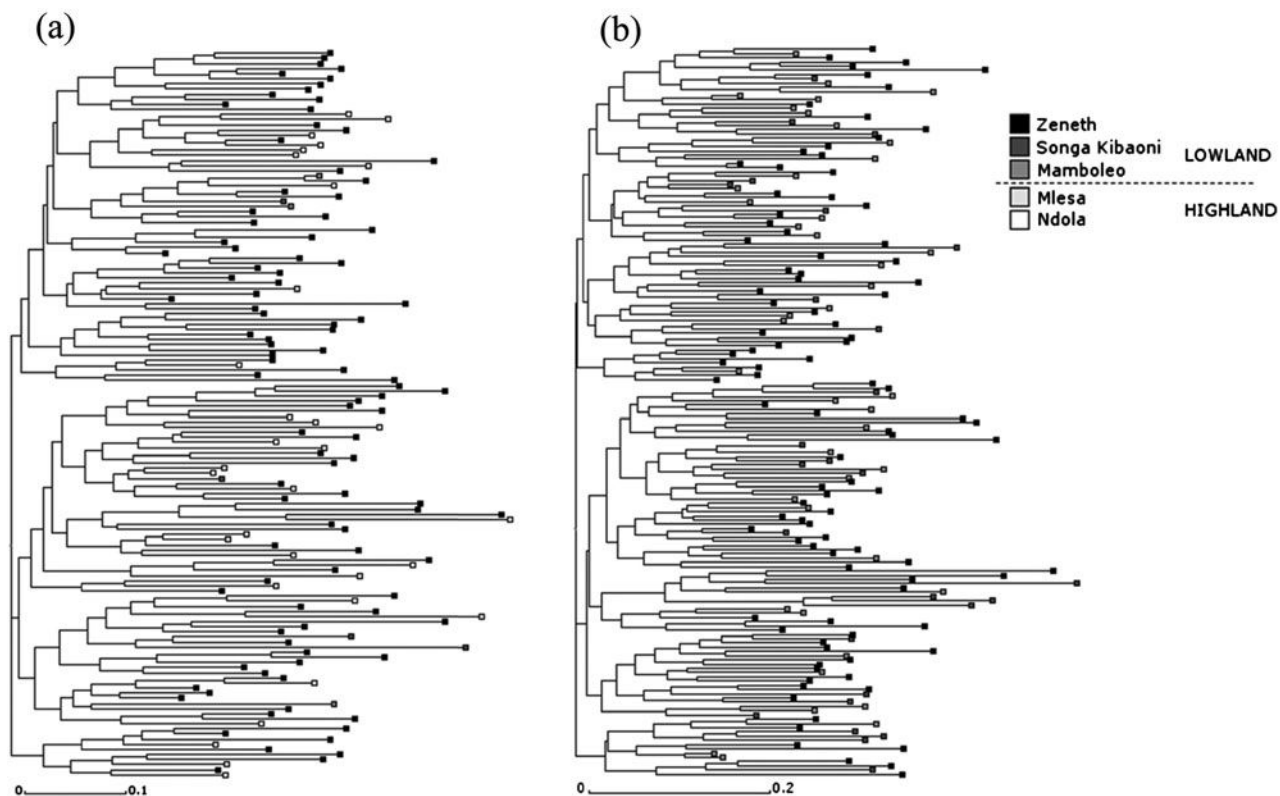


Fig. 2. Neighbour joining trees based on shared allele distance (DAS) for each species. (a) *Anopheles funestus* and (b) *Anopheles gambiae*.

$K = 1$ genetic cluster (data not shown). This is confirmed by the construction of a neighbour-joining tree based on DAS between pairs of multilocus genotypes for each species. Indeed, neither different genetic clusters nor spatial clustering for both species could be observed (Fig. 2a and b).

Discussion

Low prevalence of Plasmodium

We could not perform tests to assess genetic differentiation between 'infected' and 'uninfected' individuals due to the low prevalence found, but it is interesting to note that all infected mosquitoes (10 *An. gambiae* and six *An. funestus*) were caught in the lowland, suggesting that altitude could be a structuring factor for immunity gene. We should, however, remain cautious as temperature in altitude could also preclude *Plasmodium* infection. This would then release the pressure favouring a stronger immune response against malaria parasites. Obviously, we have no information about a potential higher mortality in the infected mosquitoes that were collected but died before dissection 7 days after capture.

High gene flow between populations

Our results showed that the highland population of *An. funestus* was not genetically differentiated from the lowland population. We had hypothesized that differences, both in altitude and in ecological factors, could lead to significant genetic differentiation, but this does not appear to be true despite a steep ecological cline. This is important because it shows that altitude does not act as a strong barrier to gene flow.

Likewise, for *An. gambiae*, the populations were not significantly differentiated from each other in the lowland area (pairwise distances ranged from 3 to 9 km; Table 2). This is consistent with previous studies that have revealed a deme size greater than 50 km for *An. gambiae* (Lehmann *et al.*, 1996; Kamau *et al.*, 1998). The similarity between topography and climate of the different locations probably explains this lack of differentiation.

Furthermore, the high positive F_{IS} found for both species could be explained by the presence of null alleles or by the Wahlund effect, either spatial (pool of individuals originated from different houses or different foci within houses) and/or temporal (pool of successive cohorts over a 4-month sampling period). One could suggest that we pooled partially

diverging clades from *An. funestus* (Michel *et al.*, 2005), but this hypothesis can be rejected as this should have been revealed by the Bayesian assignment tests. The hypothesis of possible local scale and/or temporal genetic heterogeneity cannot be ruled out. This is, however, not in contradiction with our result of an absence of genetic differentiation at a larger scale, as already found in other organisms (Torda *et al.*, 2013; Gorospe *et al.*, 2013, 2015). To test for local structure, the sampling and genotyping efforts (number of individuals and number of loci) should be enhanced for the different houses and foci to allow fine-scale structuring analyses.

Nevertheless, for both *An. funestus* and *An. gambiae*, this is a major point of concern when monitoring the spread of resistance against insecticides, as well as the resistance against drugs in *Plasmodium*. In a more hypothetical way, such high gene flow between locations presenting different ecological conditions would facilitate the spread of a genetic mutation. The lack of ecological barriers could be an advantage in the use of genetically modified mosquitoes for malaria control (Boëte, 2006; Levy, 2007; Wang and Jacobs-Lorena, 2013).

Conclusion

Our study found evidence of high gene flow between populations of malaria vectors in northeastern Tanzania. In the case of *An. funestus*, the genetic structure is unaffected by altitude, at least in this part of East Africa. However, it might be a factor influencing *An. gambiae* distribution with a decrease of mosquito number as altitude increases. Moreover, and given the importance of the spatio-temporal dynamics of vectors in disease transmission (Chaves *et al.*, 2015), there is a crucial need to understand such vector population dynamics in areas with contrasted patterns of transmission (Kamugisha, 2005), to better design tailor-made solutions to control disease transmission.

Ethics

The study received ethical approval by the Tanzanian Commission for Science and Technology (COSTECH) with research permit number 2005-51-NA-2004-53 granted to C.B. All anopheline collections were undertaken following verbal consent of householders.

Supplementary material

To view supplementary material for this article, please visit <http://dx.doi.org/10.1017/S1742758416000175>.

Acknowledgements

We are grateful to the team of entomologists for their dedication during the fieldwork and to the villagers who kindly accepted our regular visits in their homes. We thank the GENTYANE platform (INRA, Clermont, France) for their technical support for genotyping. During this work, CB was supported by a Marie Curie Intra-European Fellowship MEIF-CT-2003-501659 at Wageningen University, the Netherlands.

References

- Alilio M. S., Kitua A., Njunwa K., Medina M., Rønn A.M., Mhina J., Msuya F., Mahundi J., Depinay J. M., Whyte S., Krasnik A. and Bygbjerg I. C. (2004) Malaria control at the district level in Africa: The case of the Muheza District in Northeastern Tanzania. *American Journal of Tropical Medicine and Hygiene* 71(2 Suppl), 205–213.
- Boëte C. and Joint Malaria Programme, KCMC, Tanzania (2006) *Genetically Modified Mosquitoes for Malaria Control*, <http://www.landesbioscience.com>. ISBN: 1-58706-096-5.
- Chaves L. F., Imanishi N. and Hoshi T. (2015) Population dynamics of *Armigeres subalbatus* (Diptera: Culicidae) across a temperate altitudinal gradient. *Bulletin of Entomological Research* 105, 589–597. doi: [10.1017/S0007485315000474](https://doi.org/10.1017/S0007485315000474).
- Cohuet A., Simard F., Berthomieu A., Raymond M., Fontenille D. and Weill M. (2002) Isolation and characterization of microsatellite DNA markers in the malaria vector *Anopheles funestus*. *Molecular Ecology Notes* 2, 498–500. doi: [10.1046/j.1471-8286.2002.00290.x](https://doi.org/10.1046/j.1471-8286.2002.00290.x).
- Donnelly M. J., Pinto J., Girod R., Besansky N. J. and Lehmann T. (2004) Revisiting the role of introgression *vs* shared ancestral polymorphisms as key processes shaping genetic diversity in the recently separated sibling species of the *Anopheles gambiae* complex. *Heredity* 92, 61–68. doi: [10.1038/sj.hdy.6800377](https://doi.org/10.1038/sj.hdy.6800377).
- Evanno G., Regnaut S. and Goudet J. (2005) Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology* 14, 2611–2620.
- Excoffier L., Laval G. and Schneider S. (2005) Arlequin (Version 3.0): An integrated software package for population genetics data analysis. *Evolutionary Bioinformatics* 1, 47–50.
- Falush D., Stephens M. and Pritchard J. K. (2003) Inference of population structure using multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* 164, 1567–1587.
- Gillies M. T. and Coetzee M. (1987) *A Supplement to the Anophelinae of Africa South of Sahara (Afrotropical Region)*. South African Institute for Medical Research, Johannesburg, South Africa. 143 pp.

- Gillies M. T. and De Meillon B. (1968) *The Anophelinae of Africa South of the Sahara (Ethiopian Zoogeographical Region)*. South African Institute for Medical Research, Johannesburg, South Africa. 343 pp.
- Gorospe K. D. and Karl S. A. (2013) Genetic relatedness does not retain spatial pattern across multiple spatial scales: Dispersal and colonization in the coral, *Pocillopora damicornis*. *Molecular Ecology* 22, 3721–3736. doi:10.1111/mec.12335.
- Gorospe K. D. and Karl S. A. (2015) Depth as an organizing force in *Pocillopora damicornis*: Intra-reef genetic architecture. *PLoS ONE* 10, e0122127. doi:10.1371/journal.pone.0122127.
- Goudet J. (1995) FSTAT (Version 1.2): A computer program to calculate F-Statistics. *Journal of Heredity* 86, 485–486.
- Hodkinson I. D. (2005) Terrestrial insects along elevation gradients: Species and community responses to altitude. *Biological Reviews* 80, 489–513. doi:10.1017/S1464793105006767.
- Kamau L., Lehmann T., Hawley W. A., Orago A. S. and Collins F. H. (1998) Microgeographic genetic differentiation of *Anopheles gambiae* mosquitoes from Asembo bay, western Kenya: A comparison with Kilifi in coastal Kenya. *The American Journal of Tropical Medicine and Hygiene* 58, 64–69.
- Kamugisha M. L., Maxwell C. and Curtis C. F. (2005) Characteristics of malaria among children living in lowlands and highlands of Muheza District, north-east Tanzania. *Tanzania Health Research Bulletin* 7, 67–72.
- Koekemoer L. L., Kamau L., Hunt R. H. and Coetzee M. (2002) A cocktail polymerase chain reaction assay to identify members of the *Anopheles funestus* (Diptera: Culicidae) group. *The American Journal of Tropical Medicine and Hygiene* 66, 804–811.
- Lehmann T., Hawley W. A., Kamau L., Fontenille D., Simard F. and Collins F. H. (1996) Genetic differentiation of *Anopheles gambiae* populations from East and West Africa: Comparison of microsatellite and allozyme loci. *Heredity* 77, 192–208.
- Levy S. (2007) Mosquito modifications: New approaches to controlling malaria. *BioScience* 57, 816–821. doi:10.1641/B571003.
- Luckhart S., Li K., Dunton R., Lewis E. E., Crampton A. L., Ryan J. R. and Rosenberg R. (2003) *Anopheles gambiae* immune gene variants associated with natural *Plasmodium* infection. *Molecular and Biochemical Parasitology* 128, 83–86. doi: 10.1016/S0166-6851(03)00016-1.
- Magesa S. M., Wilkes T. J., Mnzava A. E., Njunwa K. J., Myamba J., Kivuyo M. D., Hill N., Lines J. D. and Curtis C. F. (1991) Trial of pyrethroid impregnated bednets in an area of Tanzania holoendemic for malaria. Part 2. Effects on the malaria vector population. *Acta Tropica* 49, 97–108.
- Maliti D., Ranson H., Magesa S., Kisinza W., Mcha J., Haji K., Killeen G. and Weetman D. (2014) Islands and stepping-stones: Comparative population structure of *Anopheles gambiae sensu stricto* and *Anopheles arabiensis* in Tanzania and implications for the spread of insecticide resistance. *PLoS ONE* 9(10), e110910. doi:10.1371/journal.pone.0110910.
- Mboera L. and Magesa S. (2001) The rise and fall of malarial sporozoite rates in *Anopheles gambiae* s.l. and *An. funestus* in north-eastern Tanzania, between 1934 and 1999. *Annals of Tropical Medicine & Parasitology* 95, 325–330.
- Michel A. P., Guelbeogo W. M., Grushko O., Schemerhorn B. J., Kern M., Willard M. B., Sagnon N'F., Costantini C. and Besansky N. J. (2005) Molecular differentiation between chromosomally defined incipient species of *Anopheles funestus*. *Insect Molecular Biology* 14, 375–387. doi: 10.1111/j.1365-2583.2005.00568.x.
- Midega J. T., Muturi E. J., Baliraine F. N., Mbogo C. M., Githure J., Beier J. C. and Yan G. (2010) Population structure of *Anopheles gambiae* along the Kenyan coast. *Acta Tropica* 114, 103–108. doi:10.1016/j.actatropica.2010.01.011.
- Perrier X. and Jacquemoud-Collet J. P. (2006) *DARwin Software v 5.0*. <http://darwin.cirad.fr/>.
- Pritchard J. K., Stephens M. and Donnelly P. (2000) Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.
- Ranson H., N'guessan R., Lines J., Moiroux N., Nkuni Z. and Corbel V. (2011) Pyrethroid resistance in African anopheline mosquitoes: What are the implications for malaria control? *Trends in Parasitology* 27, 91–98.
- Sinkins S. P., Hackett B. J., Costantini C., Vulule J., Ling Y.-Y., Collins F. H. and Besansky N. J. (2000) Isolation of polymorphic microsatellite loci from the malaria vector *Anopheles funestus*. *Molecular Ecology* 9, 490–492. doi:10.1046/j.1365-294x.2000.00871-2.
- Stein A., Gerstner K. and Kreft H. (2014) Environmental heterogeneity as a universal driver of species richness across taxa, biomes and spatial scales. *Ecology Letters* 17, 866–880.
- Temu E. A., Minjas J. N., Coetzee M., Hunt R. H. and Shiff C. J. (1998) The role of four anopheline species (Diptera: Culicidae) in malaria transmission in coastal Tanzania. *Transactions of the Royal Society of Tropical Medicine and Hygiene* 92, 152–158.
- Temunović M., Franjić J., Satovic Z., Grgurev M., Frascaria-Lacoste N. and Fernández-Manjarrés J. F. (2012) Environmental heterogeneity explains the genetic structure of continental and Mediterranean populations of *Fraxinus angustifolia* Vahl. *PLoS ONE* 7(8), e42764. doi:10.1371/journal.pone.0042764.
- Torda G., Lundgren P., Willis B. L. and van Oppen M. J. H. (2013) Revisiting the connectivity puzzle of the common coral *Pocillopora damicornis*. *Molecular Ecology* 22, 5805–5820. doi:10.1111/mec.12540.
- Touré Y. T., Petrarca V., Traoré S. F., Coulibaly A., Maïga H. M., Sankaré O., Sow M., Di Deco M. A. and Coluzzi M. (1994) Ecological genetic studies in the chromosomal form Mopti of *Anopheles gambiae s.str.* in Mali, West Africa. *Genetica* 94, 213–223.
- Van Oosterhout C., Hutchinson W. F., Wills D. P. M. and Shipley P. (2004) Micro-checker: Software for

- identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4, 535–538.
- van Rooyen J., Lalubin F., Glazot O. and Christe P. (2013) Altitudinal variation in haemosporidian parasite distribution in great tit populations. *Parasites & Vectors* 6, 139. doi: [10.1186/1756-3305-6-13](https://doi.org/10.1186/1756-3305-6-13).
- Wang S. and Jacobs-Lorena M. (2013) Genetic approaches to interfere with malaria transmission by vector mosquitoes. *Trends in Biotechnology* 31, 185–193. doi: [10.1016/j.tibtech.2013.01.001](https://doi.org/10.1016/j.tibtech.2013.01.001).
- Weir B. S. and Cockerham C. (1984) Estimating F-statistics for the analysis of population structure. *Evolution* 38, 1358–1370.
- WHO [World Health Organization] (2014) *World Malaria Report 2014*. World Health Organization, Geneva, Switzerland. 242 pp. ISBN: 978 92 4 156483 0.
- Wilkins E. E., Howell P. I. and Benedict M. Q. (2006) IMP PCR primers detect single nucleotide polymorphisms for *Anopheles gambiae* species identification, Mopti and Savanna rDNA types, and resistance to dielidrin in *Anopheles arabiensis*. *Malaria Journal* 5, 125. doi: [10.1186/1475-2875-5-125](https://doi.org/10.1186/1475-2875-5-125).
- Zheng L., Benedict M. Q., Cornel A. J., Collins F. H. and Kafatos F. C. (1996) An integrated genetic map of the African human malaria vector mosquito, *Anopheles gambiae*. *Genetics* 143, 941–952.

One species for one island? Unexpected diversity and weak connectivity in a widely distributed tropical hydrozoan

Bautisse Postaire^{1,2,3}, Pauline Gelin^{1,2}, J. Henrich Bruggemann^{1,2} and H el ene Magalon^{1,2}

1 Universit e de La R eunion, UMR ENTROPIE Universit e de La R eunion-CNRS-IRD (Saint Denis, France); **2** Laboratoire d'Excellence CORAIL, (Perpignan, France); **3** present adress : Aix Marseille Universit e, CNRS, IRD, Avignon Universit e, IMBE UMR 7263, (Marseille, France)

Corresponding author: Bautisse Postaire. Address: Aix Marseille Universit e, CNRS, IRD, Avignon Universit e, IMBE UMR 7263, Station marine d'Endoume, Chemin de la batterie des lions, 13007, Marseille, France. Email: bautisse.postaire@imbe.fr

Running title: Population genetics of a marine brooding hydrozoan

Word count: 5355

Abstract

Isolation by distance (IBD) is one of the main modes of differentiation in marine species, above all in species presenting low dispersal capacities. This paper reports the genetic structuring in the tropical hydrozoan *Lytocarpia brevirostris* α (*sensu* Postaire et al. 2016b), a brooding species, from 13 populations in the Western Indian Ocean and one from New Caledonia (Tropical Southwestern Pacific). At the local scale, populations rely on asexual propagation at short distance, which was not found at larger scales; identical genotypes were restricted to single populations. After the removal of repeated genotypes, all populations presented significant positive F_{IS} values (between 0.094*** and 0.335***). Gene flow was extremely low at all spatial scales, between sites within islands (< 10 km distance) and among islands (100 to > 11 000 km distance), with significant pairwise F_{ST} values (between 0.012*** and 0.560***). A general pattern of IBD was found at the Indo-Pacific scale, but also within sampled ecoregions of the Western Indian Ocean province. Clustering analyses identified each sampled island as an independent population, while analysis of molecular variance indicated that population genetic differentiation was significant at small (within island) and intermediate (among islands within province) spatial scales. The high population differentiation might reflect the life cycle of this brooding hydrozoan, possibly preventing regular dispersal at distances more than a few kilometres and probably leading to high cryptic diversity, each island housing an independent evolutionary lineage.

Keywords: Hydrozoa, *Lytocarpia*, microsatellite, population genetics, brooding species, cryptic diversity

Introduction

Seascape connectivity, the process linking habitat patches and populations through the exchange of organisms across the marine environment, is a key driver of population dynamics, genetic structuring and diversification processes of marine organisms (Palumbi 1992; Paulay and Meyer 2002; Cowen et al. 2007; Boissin et al. 2011; Bowen et al. 2013). Knowledge of seascape and population connectivity ideally forms the basis for the definition of management and conservation units (Cowen et al. 2007; Christie et al. 2010; White et al. 2010; Thomas et al. 2012). Indeed, a high proportion of marine organisms rely on the dispersal of their larval stage by ocean currents for population maintenance over time and colonization of new habitats. Stretches of open ocean are thus often regarded as environmental barriers reducing the dispersal (*i.e.* including the transport and recruitment phases) of propagules (gametes and larvae) over large geographic scales for organisms presenting short pelagic larval duration, direct development, brooding and/or holobenthic life cycles (*e.g.* Mokhtar-Jamaï et al. 2011). Hence, marine species with low dispersal capacities are expected to have narrow distribution ranges and strong population differentiation with geographic distance, *i.e.* an isolation by distance (IBD) pattern (Slatkin 1993). This pattern is thought to increase the number of speciation opportunities, mainly allopatric, by favouring vicariant events and the formation of independent evolutionary lineages over time (Paulay and Meyer 2002; Malay and Paulay 2010) leading *in fine* to speciation (De Queiroz 1998).

The relative importance of environmental barriers opposing the dispersal of organisms is often estimated via the assessment of population connectivity across geographic distances. This connectivity represents the genetic linking of local populations through the effective dispersal (*i.e.* transport, recruitment and reproduction) of individuals (larvae, juveniles or adults) among them (Sale et al. 2005). It represents a continuum, from an absence of connectivity (closed populations, mainly self-recruiting) to high connectivity (open populations, most of the recruitment happens via migration of individuals), in which the life cycle and reproductive strategy are important traits that shape population differentiation and connectivity. Nevertheless, Shanks (2009) observed that the relationship between direct or indirect development, pelagic larval duration, swimming capacity and dispersal capacity is not always straightforward, as a large body of literature exists, detailing examples of lecithotrophs, brooders and direct developers, with biogeographic ranges spanning oceans (Ayre and Hughes 2000; Kyle and Boulding 2000; Boissin et al. 2008). Thus other aspects, *e.g.* larval behaviour and species ecology, must be considered to explain the patterns of marine

population connectivity (Shanks 2009), as well as oceanic circulation and historical sea level variations (Tremblay et al. 2007; Ayre et al. 2009; Schiavina et al. 2014). In order to achieve a more comprehensive understanding of the maintenance of natural populations over time, it is necessary to multiply the number of biological models studied.

Hydrozoans are ubiquitous in all marine ecosystems with species often presenting broad biogeographic distributions and a variety of life cycles and reproductive features (Bouillon et al. 2006). The Aglaopheniidae family (Marktanner-Turneretscher 1890) represents one of the largest, with over 250 valid species (Bouillon et al. 2006) and is particularly diversified in tropical marine ecosystems. The diversity of this family is still under assessment, as recent publications highlighted that morphological characters provide little clues to the evolutionary history and species richness of this taxon, presenting extensive cases of morphological convergence and low anagenesis (Leclère et al. 2007; Moura et al. 2012; Postaire et al. 2016a; Postaire et al. 2016b). Aglaopheniidae species generally do not have a medusa stage but brood their larvae in dedicated structures of the colony (named gonothecae; Millard 1975), and planulae are released only when mature (Boero et al. 1992). Hydrozoan planulae usually crawl rather than swim and settle in less than 24 h, suggesting low dispersal capacities (see Gili and Hughes (1995) for a review of hydrozoan ecology). Intuitively, this feature contradicts the extensive geographic ranges spanning several major biogeographic provinces of these species. *Lytocarpia brevirostris* (Busk 1852) is a typical tropical Aglaopheniidae brooding species, found on coral reefs throughout the Indo-Pacific region (Millard 1975; Gravier-Bonnet 2006; Gravier-Bonnet and Bourmaud 2006a, 2006b, 2012; Di Camillo et al. 2011). This orange feather-like colonial hydrozoan is composed of two sympatric cryptic species, which can be identified using molecular markers (either typical barcoding markers or microsatellites; Postaire et al. 2016b). In this study, we focus on one of these cryptic species, named *L. brevirostris* α in Postaire et al. (2016b). Hydrozoans are among the least studied cnidarians for population genetics, but the few studies available to date (Darling and Folino-Rorem 2009, Schuchert 2014) highlighted the high genetic differentiation among individuals sampled from different regions. Considering these previous results, one could argue that each population restricted to a small geographic region potentially represents an independent species (*sensu* Samadi and Barberousse 2006). Indeed, the classification of organisms into species may be hampered by variation of genetic and morphological characters within and among populations. Furthermore, gene flows across ocean basins are known to vary in intensity and direction according to oceanic circulation (Dawson 2001; Hohenlohe 2004; Cowen et al. 2007; Weersing and Toonen 2009), influencing both population differentiation and speciation processes. Thus, the

wide distributions of marine brooding hydrozoans could be (1) a taxonomic artefact and species represent complexes of sibling species with limited dispersal abilities, or (2) a single species with an extensive population subdivision (Schuchert 2014). One of the solutions to start answering the problem is to conduct an analysis of population structuring based on an extensive hierarchical sampling of populations at large geographic scales.

In this study, we used microsatellites to assess the population structuring of *L. brevirostris* α . We sampled populations in two biogeographic provinces as defined by Spalding et al. (2007): the Western Indian Ocean (WIO: Reunion Island, Rodrigues, Madagascar, Scattered Islands) and the Tropical Southwestern Pacific (TSP: New Caledonia). We evaluated patterns of connectivity across three spatial scales: within sampling site (short distance connectivity), among islands within a biogeographic province (intra-regional comparisons) and between biogeographic provinces (inter-regional connectivity). To our knowledge, this is the first study on genetic connectivity in marine hydrozoans.

Method

Sample collection and DNA extraction

Each sampling site was explored randomly using SCUBA diving. *Lytocarpia brevirostris* individuals (defined as a plume), which present a scarce and patchy distribution on outer reef slopes, often found in shaded caverns or on vertical cliffs, were sampled between three and 25 m depth. When encountered, individuals were collected and placed individually in sequentially numbered plastic bags to approximate distances between individuals (the closeness in number reflects the proximity of individuals). *Lytocarpia brevirostris* grows by asexual propagation and individual plumes are sometimes linked by a stolon, thus forming colonies (defined hereafter as the aggregation of individuals forming a genet). Consequently, to minimize the probability of sampling members from the same genet, we collected, where possible, individuals at least few centimetres to several decimetres apart. Larger individuals (3-10 cm high) with visible reproductive structures were preferentially collected to avoid misidentification with other species from the genus *Lytocarpia*. A total of 593 collected *L. brevirostris* α individuals (Figure 1, Table 1) were primarily identified using morphological characters (Millard 1975) and verified using microsatellite data (Postaire et al. 2016b). *Lytocarpia brevirostris* was relatively rare in New Caledonia compared to the WIO and only three sites presented more than five individuals (Figure 1, Table 1) among 18 explored sites all around the island.

Table 1: *Lytocarpia brevirostris* α samples (N = 593) used in this study. For each population are given: total sample size (N), number of unique multi-locus genotypes (MLG) (N_{MLG}), clonal richness $R = [(N_{MLG} - 1) / (N-1)]$, observed (H_O) and expected (H_E) heterozygosities, inbreeding coefficient (F_{IS}), allelic richness $Ar(26) \pm se$ (standard error) and private allelic richness $Ap(26) \pm se$. GPS coordinates of sampling sites are in decimal degrees. H_O , H_E and F_{IS} were calculated keeping one representative per MLG per population. With the F_{IS} , is indicated the test significance for deviation to Hardy-Weinberg equilibrium: NS: non-significant; *: $P < 0.05$; **: $P < 0.01$; ***: $P < 0.001$.

Marine Province	Ecoregion	Island	Site name	Population	Latitude	Longitude	N	N_{MLG}	R	H_O	H_E	F_{IS}	$Ar(26)$	$Ap(26)$	
WIO	Mascarene Islands	Reunion Island	Cap La Houssaye	RUN1	-21.01766	55.23836	56	36	0.636	0.374	0.416	0.104*	2.919 \pm 0.142	0.009 \pm 0.009	
			Passe de l'Ermitage	RUN2	-21.08659	55.22065	36	13	0.342	0.344	0.447	0.238**	2.753 \pm 0.231	0.148 \pm 0.085	
			Saint Leu	RUN3	-21.18419	55.28418	57	42	0.732	0.438	0.488	0.104**	3.322 \pm 0.277	0.011 \pm 0.007	
			Manapany	RUN4	-21.37500	55.58358	35	23	0.647	0.354	0.477	0.263*	3.073 \pm 0.463	0.121 \pm 0.112	
	Rodrigues	Petits Patés	ROD1	-19.65386	63.41501	46	13	0.266	0.452	0.475	0.051***	3.086 \pm 0.602	0.237 \pm 0.130		
		Ilot Cocos	ROD2	-19.74016	63.28660	46	11	0.222	0.331	0.433	0.238***	2.933 \pm 0.611	0.161 \pm 0.102		
	Western/Northern Madagascar	Juan de Nova Island	Juan 2	JUA1	-16.95337	42.76011	41	34	0.825	0.489	0.636	0.235***	4.472 \pm 0.565	0.200 \pm 0.102	
			Juan 6	JUA2	-17.08177	42.72536	52	39	0.745	0.545	0.661	0.177***	5.051 \pm 0.638	0.144 \pm 0.083	
			P7	JUA3	-17.03280	42.73580	43	36	0.833	0.462	0.647	0.290***	5.292 \pm 0.567	0.313 \pm 0.139	
			Biodiv 7	JUA4	-17.07470	42.76650	33	25	0.750	0.572	0.707	0.194***	5.600 \pm 0.510	0.220 \pm 0.105	
Madagascar		Anakao	MAD1	-23.66586	43.59985	45	27	0.590	0.395	0.472	0.166***	3.978 \pm 0.631	0.129 \pm 0.055		
		Ifaty	MAD2	-23.19096	43.58291	26	18	0.680	0.350	0.526	0.341***	4.087 \pm 0.681	0.068 \pm 0.054		
		Tuléar	MAD3	-23.40370	43.63818	46	24	0.511	0.363	0.512	0.294***	3.723 \pm 0.555	0.189 \pm 0.109		
TSP	New Caledonia	Maré	Sud Cap Coster	NCA1	-21.48673	168.11940	31	31	1	0.454	0.592	0.239***	4.291 \pm 0.632	1.691 \pm 0.393	
			Grande Terre	Népoui	NCA2	-21.41855	164.96713	10	10	1	0.433	0.592	0.177***		
			Grande Terre	Kouma	NCA3	-20.81090	164.30625	6	6	1	0.509	0.608	0.280***		

Specimens were fixed and preserved in 90% ethanol for later DNA extraction. Prior to DNA extraction, all reproductive structures were removed from the individuals to avoid genotyping progeny issued from sexual reproduction and thus distinct individuals. DNA was extracted from one or two primary branches of the hydrocaulus per individual using DNeasy Blood & Tissue kit (Qiagen, Hilden, Germany) following the manufacturer's protocol.

Microsatellite genotyping

We used 16 microsatellite loci specific for the *L. brevis* species complex (Postaire et al. 2015a). Amplifications, thermocycling and genotyping conditions were the same as in Postaire et al. (2015b). Identical multi-locus genotypes (MLG) were identified with GENCLONE v. 2.0 (Arnaud-Haond and Belkhir 2007) using the maximum set of loci for each sampling site and keeping only the individuals without missing data. For subsequent analyses, we kept only the loci presenting less than 10% of missing data for the entire dataset (*i.e.* 10 loci: Lb01, Lb02, Lb03, Lb05, Lb06, Lb07, Lb08, Lb10, Lb11, Lb16; see Results) and only one representative per MLG (*i.e.* one individual per colony; Table 1).

All tests in this study were corrected for false discovery rate (FDR) in multiple tests (Benjamini and Hochberg 1995). We used MICROCHECKER v. 2.3 (van Oosterhout et al. 2004) to check for scoring errors and to estimate null allele frequencies. Linkage disequilibrium (LD) was tested using Arlequin v. 3.5 (Excoffier et al. 2005) among all pairs of loci in each population with a permutation test ($n = 10^3$). Observed (H_o), expected (H_e) heterozygosities, F_{IS} indices and tests for Hardy-Weinberg equilibrium (HWE) were computed using the software Arlequin v. 3.5 (Excoffier et al. 2005) for all populations and over all loci. Mean allelic richness [$Ar(g)$] and mean private allelic richness [$Ap(g)$] were calculated. We applied a rarefaction method to obtain estimates of both values independently of sample size using ADZE v. 1.0 (Szpiech et al. 2008). We chose to compare both values to a common sample size of $g = 26$ as it was the minimum number of individuals found per site, excluding NCA2 and NCA3 as these populations were too small (≤ 10 individuals; Table 1).

Clustering individuals and populations, analyses of population differentiation and gene flow

We investigated the population structure using different approaches: measures of differentiation and a Bayesian clustering method. First, the geographic origin of samples (*i.e.* site) was treated as an *a priori* defined population. Due to low sampling size, NCA2 and NCA3 were

pruned from this set of analyses. Population-level pairwise F_{ST} comparisons and Fisher's exact tests of population differentiation were performed in Arlequin v. 3.5. (Excoffier et al. 2005). The significance of the observed F_{ST} statistics was tested using the null distribution generated from 5×10^3 non-parametric random permutations. To infer mechanisms that may be responsible for the observed patterns of population structure, we compared estimates of genetic differentiation to geographic distances among sites of the WIO and the TSP. Euclidean distances between sampling locations were measured with Google Earth v. 7.1 (<http://earth.google.fr/>) using site coordinates (Table 3) and taking into account the regional pattern of oceanic currents (Schott et al. 2009). We used a Mantel test (Mantel 1967) to evaluate the correlation between linearized genetic differentiation (Slatkin's distance = $F_{ST} / (1 - F_{ST})$) and the \log_{10} of the geographic distance between sites (Table 3). This relationship is expected to be positive and linear in the context of a two-dimensional IBD model (Rousset 1997). All Mantel tests were performed using the program GENODIVE (Meirmans and van Tienderen 2004) with 10^4 random permutations to assess significance. Population differentiation was also assessed without *a priori* stratification of samples.

Then, the geographic origin of samples was no more considered, allowing us to use samples from NCA2 and NCA3. We performed a discriminant analysis of principal components (DAPC) in the R (R Development Core Team 2004) package *adegenet* (Jombart 2008; Jombart et al. 2010). DAPC is a non-model-based method that maximizes the differences between groups while minimizing variation within groups without prior information on individuals' origin. This method does not assume HWE or absence of LD. We used the function `find.clusters()` to assess the optimal number of groups with the Bayesian information criterion (BIC) method (*i.e.* K with the lowest BIC value is ideally the optimal number of clusters). We tested values of K ranging from 1 to 30, but BIC values may keep decreasing after the true K value in case of genetic clines and hierarchical structure (Jombart et al. 2010). Therefore, the rate of decrease in BIC values was visually examined to identify values of K beyond which BIC values decreased only slightly (Jombart et al. 2010). The `dapc()` function was then executed using the best grouping, retaining axes of PCA sufficient to explain $\geq 90\%$ of total variance of data. Afterwards, Bayesian clustering analyses were performed to estimate the most probable number of populations (K) given the data, as implemented in the program STRUCTURE v. 2.3.2 (Pritchard et al. 2000), using the admixture model with correlated allele frequencies (Falush and Pritchard 2003). This analysis assumes that, the dataset is composed of K populations and individuals are assigned to each putative population under HWE and minimized LD. We studied the assignment of samples again using a hierarchical approach. Four independent runs were conducted for each value of K from 1 to 10 with a burn-in period of 5×10^4

steps followed by 5×10^5 Markov chain Monte Carlo iterations. We used the statistic proposed by Evanno et al. (2005) to estimate the number of clusters K implemented in STRUCTURE HARVESTER (Earl and vonHoldt 2012). The STRUCTURE outputs of the best number of K were summarized with CLUMPP v. 1.0 (Jakobsson and Rosenberg 2007) and formatted with DISTRUCT v. 1.1 (Rosenberg 2004). Finally, Arlequin v. 3.5 was used to perform hierarchical analysis of molecular variance using clusters identified by STRUCTURE as groups. A hierarchical analysis of molecular variance (AMOVA) (Excoffier et al. 1992), using provinces as groups and islands as populations was finally performed.

Results

Multi-locus genotyping and potential importance of asexual reproduction

Over the 16 available loci for *L. brevirostris* α , 15 loci amplified correctly (*i.e.* presented less than 10% of missing data) on samples from Reunion Island (all except Lb13), 12 loci (all except Lb04, Lb09 and Lb12, Lb13) on samples from Juan de Nova Island, Madagascar and Rodrigues (WIO except Reunion Island) and 10 loci (all except Lb04, Lb09, Lb12, Lb13, Lb14 and Lb15) on samples from New Caledonia (TSP). Our analysis of 609 individuals yielded 470 MLGs, indicating that asexual propagation occurred in the sampled populations. The individuals sharing the same MLG were found close together (*i.e.* small difference in field numbers) (Figure 2). Moreover, none of the MLGs were shared by different populations. In other words, clones were confined to their sites. Consequently, only one representative of each MLG was used for further analyses.

Genetic variability

Global significant LD among loci was detected ($P < 0.05$, 68 significant tests over 720 after FDR correction, *i.e.* 9.4%) in the global dataset. However, more than half of the positive tests (38 out of 68) occurred in two populations with low clonal richness (ROD1, ROD2) and might just represent their low genetic diversity. All loci were polymorphic, with a total number of alleles ranging from eight (Lb01, Lb07) to 23 (Lb02) (mean \pm se = 14.4 ± 1.8). Some loci were monomorphic in several populations: Lb02 in population NCA1, Lb01 and Lb06 in ROD1 and ROD2, Lb05 in RUN2 and Lb10 in RUN4. Observed heterozygosities ranged from 0.331 to 0.572 (mean \pm se = 0.429 ± 0.019) for ROD2 and JUA4, respectively, and unbiased expected heterozygosities from 0.416 to 0.707 (mean \pm se = 0.543 ± 0.023) for RUN1 and JUA4, respectively. Mean allelic richness per locus ranged from 2.919 ± 0.142 (se) in RUN1 to 5.600 ± 0.510 (se) in JUA4 and mean number of private alleles per locus ranged from 0.009 ± 0.009 (se) in RUN1 to 1.691 ± 0.393 (se) in NCA1.

Multi-locus F_{IS} values were all significantly positive ($P < 0.05$; Table 1) and ranged between 0.051*** for ROD1 and 0.341*** for MAD3. Over all loci, significant heterozygote deficiencies were found in all populations (after FDR correction). For each locus in each population, presence of null alleles was checked. Null alleles were detected for several loci in several populations. However, as (1) LD was not constant among loci, (2) not a single locus was monomorphic over all populations, (3) the number of loci with null alleles was not constant between populations and (4) in several cases, the F_{IS} value was significantly positive without evidence of null alleles, we decided to keep the ten loci presenting less than 10% of missing data for further analyses.

Genetic clusters

Population structuring was evidenced by results of DAPC and STRUCTURE analyses that showed genotypes clustering according to their geographic origin. STRUCTURE outputs revealed that the plot of $\ln P(D)$ as a function of K showed a clear plateau starting at $K = 5$ (Figure 3A), the most likely number of clusters. Each cluster corresponded to a sampled island, except for MLGs from New Caledonia (Figure 3B). When analysing the genetic clustering of these MLGs alone (data not shown), they clustered according to their island of origin, *i.e.* Maré for NCA1 and west coast of Grande-Terre for NCA2 and NCA3.

DAPC also identified clusters corresponding to individuals sampled from the same island. In the successive values of K (number of cluster tested), the initial decline in BIC values slowed at $K = 5$ (Figure 4A). As the three first axes explained more than 95% of the variance, we decided to present the DAPC results in two ordination plots: (1) first and second axes and (2) second and third axes. DAPC separated the WIO from the TSP along the first PCA axis (75.5% of variance; eigenvalue = 19656.61). The second axis explained less variability (11.2% of variance; eigenvalue = 19656.61), slightly separating WIO colonies from the TSP and plotting colonies from Reunion Island distant from those of other populations in the WIO (Figure 4B). The third axis (8.62% of variance; eigenvalue = 2244.09) separated colonies sampled in Rodrigues from those originating from Juan de Nova and Madagascar (Figure 4C). As STRUCTURE, when analysing the clustering scheme of MLGs only from New Caledonia with DAPC, they clustered according to their island of origin.

Using provinces as groups and islands as populations, AMOVA revealed a highly significant genetic structuring among islands within provinces (WIO and TSP) and within islands ($P < 0.001$; Table 2), but not between provinces. The genetic variation explained by differences among islands

within provinces was higher than the genetic variation explained by differences among provinces (26.31%*** and 11.22%^{NS}, respectively) and the highest amount of genetic variation was found within islands (62.47%***).

Assessment of connectivity over different geographic scales and isolation by distance

All pairwise differentiation tests were significant but none of the exact Fisher's tests (after FDR correction). Pairwise F_{ST} values indicated a high and significant differentiation between populations (Table 4), ranging from 0.012*** to 0.560***. Concerning Reunion Island populations, the highest differentiation was found between RUN2 and RUN4 populations ($F_{ST} = 0.203$ ***). In the WIO, the lowest differentiation occurred between populations from the same island (minimum $F_{ST} = 0.013$ *, between JUA1 and JUA2) and the maximum F_{ST} values occurred between populations ROD2 (Rodrigues) and RUN4 (Reunion Island) ($F_{ST} = 0.560$ ***). Populations ROD1 and ROD2 were highly differentiated from all other populations of the WIO (F_{ST} ranged from 0.249*** to 0.560***). When comparing populations from the WIO with the population from New Caledonia, F_{ST} values were also high and highly significant, ranging from 0.297*** to 0.475*** (Table 4).

Table 3: *Lytocarpia brevirostris* α pairwise F_{ST} values (below diagonal) and geographic distance (in kilometres: above diagonal) for all pairs of populations. All values were highly significantly different from 0 ($P < 0.001$) after FDR correction, except for four values (in italics) with $0.001 < P < 0.05$.

Population	JUA1	JUA2	JUA3	JUA4	MAD1	MAD2	MAD3	RUN1	RUN2	RUN3	RUN4	ROD1	ROD2	NCA1
JUA1		15	9	13	748	696	719	1 388	1 389	1 398	1 434	2 202	2 190	12 704
JUA2	<i>0.013</i>		6	4	734	682	706	1 386	1 387	1 397	1 433	2 204	2 191	12 699
JUA3	0.019	0.025		6	470	688	712	1 388	1 388	1 397	1 433	2 204	2 191	12 701
JUA4	<i>0.020</i>	<i>0.014</i>	0.039		735	683	707	1 383	1 383	1 393	1 429	2 199	2 187	12 696
MAD1	0.216	0.202	0.218	0.203		53	29	1 234	1 230	1 234	1 258	2 097	2 081	12 208
MAD2	0.226	0.215	0.222	0.200	0.122		24	1 226	1 223	1 227	1 253	2 091	2 076	12 240
MAD3	0.238	0.212	0.226	0.215	0.039	0.075		1 225	1 221	1 225	1 250	2 089	2 073	12 221
RUN1	0.341	0.328	0.323	0.300	0.336	0.316	0.337		8	19	53	867	852	11 349
RUN2	0.300	0.288	0.291	0.265	0.328	0.327	0.332	0.073		13	49	870	855	11 347
RUN3	0.338	0.326	0.321	0.293	0.330	0.305	0.335	<i>0.016</i>	0.102		38	865	850	11 336
RUN4	0.361	0.343	0.343	0.322	0.337	0.322	0.344	0.112	0.203	0.117		839	823	11 298
ROD1	0.303	0.277	0.270	0.249	0.464	0.431	0.443	0.489	0.508	0.467	0.540		17	10 657
ROD2	0.322	0.296	0.298	0.272	0.486	0.450	0.482	0.516	0.537	0.489	0.560	0.162		10 665
NCA1	0.337	0.331	0.338	0.298	0.426	0.400	0.421	0.417	0.410	0.404	0.452	0.447	0.475	

Generally, population differentiation in *L. brevirostris* α was low between populations from the same island and high at every other scale. Mantel tests revealed a significant positive correlation between transformed F_{ST} values and the \log_{10} of the geographic distances among sites both within the WIO province ($n = 78$, $r = 0.624^{***}$, $R^2 = 0.390$) and within each of the two ecoregions, Western/Northern Madagascar and the Mascarene Islands ($n = 21$, $r = 0.977^{***}$, $R^2 = 0.955$ and $n = 15$, $r = 0.957^{***}$, $R^2 = 0.917$, respectively). Similarly, a strong pattern of IBD was evidenced at the Indo-Pacific scale ($n = 91$, $r = 0.639^{***}$, $R^2 = 0.408$; Figure 5).

Discussion

We investigated the genetic structure and connectivity among populations of a widely distributed hydrozoan species, *Lytocarpia brevirostris* α (*sensu* Postaire et al. 2016b) across three spatial scales in the Indo-Pacific region using a set of 16 microsatellite loci. The study revealed that populations of this brooding hydrozoan were characterized by low connectivity even at the smallest spatial scale (a few kilometres), presented an IBD pattern and that the detected genetic clusters correspond to the sampled islands. Our results are congruent with those of the only other molecular study on marine hydrozoans, based on two markers, where high genetic differentiation was observed among populations (Schuchert 2014, but see also Schuchert 2005). Such pattern of low population connectivity may be typical for other marine species with similar life cycles.

***High genetic differentiation of L. brevirostris* α populations**

Pairwise F_{ST} values (Table 3) revealed the high isolation of all sampled populations and highlighted the extreme differentiation of populations from Rodrigues and New Caledonia from all the others, underscoring the isolated position of Rodrigues in the WIO. Using microsatellites also allowed the identification of strong and significant genetic structuring at smaller geographic scales, indicating that gene flow is low even at distances less than 10 km (*e.g.* between JUA2 and JUA4 populations from Juan de Nova Island). Private alleles were found in several populations, with the population from New Caledonia presenting the highest mean number of private alleles. The finding that individuals from several populations were difficult to amplify for certain loci (probably because of existing null alleles) supports the inference of high divergence among populations from the WIO and the TSP. Bayesian clustering and PCA analyses further confirmed the high isolation of *L. brevirostris* α populations across the Indo-Pacific as they identified sampled colonies from each island as putative populations, while the AMOVA showed that a high and significant proportion of

the global differentiation occurred among islands. Furthermore, Mantel tests over several geographic scales revealed that population differentiation is related to geographic distance: the IBD pattern is detected at scales ranging from 100 to > 1,000 km, but absent at the local scale. Supporting this general pattern of population isolation, genetic indices of diversity (mean allelic richness and heterozygosity) were slightly different between marine ecoregions: overall, populations from the Mascarene Islands present a slightly lower allelic richness and lower heterozygosity (Table 1). These dissimilarities might reflect differences in current selective pressures among islands, but might also attest to past climatic and geological events (sea level change) that modified the dynamics of the sampled populations, e.g. population bottlenecks. This aspect merits further study.

Extremely high genetic differentiation between populations of a single species is unusual although similar levels have been documented over large geographic scales in sponges (Chaves-Fonnegra et al. 2015), coastal sharks (Ashe et al. 2015) and marine mammals (Fruet et al. 2014), but also at distances less than 100 km in terrestrial animals (Sethuraman et al. 2013) or freshwater diatoms (Vanormelingen et al. 2015). Our results indicate that the populations of *L. brevirostris* α present clear geographic boundaries and are consistent with two of the characteristics of a metapopulation model (Grimm et al. 2003): (1) local populations have their own dynamics and (2) they are connected by limited dispersal.

Potential barriers to gene flow and limited dispersal

Our results clearly indicate that expanses of deep ocean waters represent a barrier to dispersal for *L. brevirostris* α , similar to some other coastal organisms with low dispersal capacities (e.g. Ragionieri et al. 2010; Aurelle et al. 2011). First of all, each MLG was restricted to a single population. Moreover, within sampled populations, individuals close to each other (*i.e.* presenting small difference in field numbers) presented a higher probability to share the same MLG, forming a colony (or genet). As colonies of *L. brevirostris* α can grow through stolonial expansion, sampled individuals sharing the same MLG were either connected through their stolon or represent distinct fragments of an ancient extended colony. Thus colonies are spatially restricted, spanning a few centimetres or decimetres. This clonal range is quite narrow compared to some other clonal marine species, such as scleractinian corals [several kilometres (Baums et al. 2006; Pinzón et al. 2012; Japaud et al. 2015)]. Nevertheless, when restricting the data to one individual per colony, all populations showed a significant deviation from HWE due to heterozygote deficiency.

Heterozygote deficiency is relatively common in marine colonial organisms, such as scleractinian corals (Ayre and Hughes 2000; Baums et al. 2005; Underwood et al. 2007; Ridgway et al. 2008), but the exact mechanism driving this effect cannot always be determined (inbreeding or spatial/temporal Wahlund effect). Considering the reproductive strategy of *L. brevirostris* α (absence of medusa stage, internal fertilization of eggs, larviparity), this pattern might be explained by restricted dispersal of gametes and/or larvae. Indeed, observations of larval behaviour in hydrozoans suggest that dispersal after planulae release is low: they tend to settle nearby their mother colony (Sommer 1990), thus favouring fertilization between related individuals. Furthermore, observations on hydrozoans in aquaria indicate that the life span of male gametes in the water column is only a few hours (Yund 1990), limiting long distance dispersal and thus gene flow among distant colonies.

The isolation between WIO and TSP populations was quite expected as these two regions are > 10,000 km distant (Figure 1). Indeed, several studies found such an Indian Ocean-Western Pacific disjunction (Kochzius and Nuryanto 2008, Yasuda et al. 2009, Richards et al. 2016). For westward migration, *L. brevirostris* α propagules originating from New Caledonia would have to survive in the plankton for several months to disperse into the WIO. In the opposite direction, propagules from the WIO should be able to survive a pelagic environment a long time but also withstand unfavourable oceanic conditions, as they would migrate along the temperate southern coast of Australia to reach New Caledonia (Schott et al. 2009). Given the reproductive mode of *L. brevirostris* α , such migrations would be extremely rare events. At smaller geographic scales, within the WIO, previous population genetics studies revealed that oceanic gyres isolate Juan de Nova Island from the Southern part of the Mozambique Channel (Bourjea et al. 2006; Krishna et al. 2006; Muths et al. 2011); our results support these findings. In the Mascarene Islands, the Southeast Madagascar current (Schott et al. 2009) could allow connectivity between Rodrigues and Reunion Island populations, but our results indicate that such gene flow is absent or extremely low. Marine circulation models at even smaller scales (i.e. around islands, coastal areas) are still under development for most of the sampled islands. Nevertheless, a model of oceanic circulation around Reunion Island indicates high heterogeneity in the direction of currents across seasons and years (Pous et al. 2014), partly explaining the absence of IBD at this geographic scale because reproduction has been observed throughout the year (BP, HM and CAF Bourmaud, pers. obs.). To our knowledge, no detailed population genetic study of a marine hydrozoan species presenting biphasic life cycle (i.e. medusa and fixed colonial stages) using microsatellites exists yet to compare our findings. However Schuchert (2005), using a mitochondrial marker, found that the

genetic diversity of a hydrozoan species with a medusa stage [*Coryne eximia* Allman, 1859] was modest over wide geographic distances (several oceans) when compared to monophasic species of the same genus (*C. muscoides* (Linnaeus, 1761) and *C. pintneri* Schneider, 1897), supporting the importance of a long-lived planktonic medusa stage in the dispersal capacities of hydrozoans. This aspect needs to be further explored by modelling gene flows in correlation with ocean circulation models, i.e. seascape genetics, an approach that has already shown merit to explain patterns of connectivity in marine organisms, such as bivalves, crustaceans and scleractinians (Tremblay et al. 2007; Ragionieri et al. 2010; Thomas et al. 2012).

Our results suggest that ocean circulation plays a minor role in determining spatial patterns of genetic differentiation in *L. brevirostris* α populations of the WIO and the TSP. Instead, short-distance exchange of gametes, larval brooding and restricted movements of larvae seem to be mostly responsible of the observed pattern, as they tend to favour small-scale genetic differentiation (within populations or islands). The wide Indo-Pacific distribution of *L. brevirostris* α is better explained by rafting of adult colonies fixed on floating objects. According to Thiel and Gutow (2005), Aglaopheniidae present several life traits enhancing their capacity of travelling through rafting: they can cling on various substrata and notably on natural or artificial floating items (BP and HM, pers. obs.) and adult individuals feed on plankton, a common pelagic resource. In addition, their capacity of clonal growth facilitates the colonization of new suitable habitats. Thus, even if rafting events are rare, they might occur at a sufficient rate to colonize new islands and explain the repartition of the clade formed by *L. brevirostris* α . This assumption is supported by the presence of both cryptic species of *L. brevirostris* α and β in the WIO and the TSP (Postaire et al. 2016b). The importance of rafting in the marine environment, above all in species presenting direct development, is increasingly recognized (DeVantier 1992; Johnson et al. 2001; Thiel and Gutow 2005; Thiel and Haye 2006; Rocha et al. 2006).

Implications for hydrozoan taxonomy and marine conservation

Our study also provides some taxonomic clues. Indeed, the extensive population subdivision in *L. brevirostris* α is concordant with the phylogeny based on mitochondrial and nuclear markers of this species (Postaire et al. 2016b) and reveals that populations inhabiting the same marine ecoregions (as defined by Spalding et al. 2007) represent independent evolutionary units or even species when considering the genealogical species concept (Baum and Shaw 1995). As they match several metapopulation characteristics, these groups of populations might actually represent

lineages engaged in a speciation process but situated in the 'grey zone' (De Queiroz 1998, 2005), *i.e.* the evolutionary time during which two lineages are definitively diverging but where criteria commonly used for identifying divergence might not be applicable or in agreement (Pante et al. 2015). As both *L. breviostris* cryptic species (α and β) present an Indo-Pacific distribution (Millard 1975; Postaire et al. 2016b), the observed extensive population differentiation within *L. breviostris* α may reflect very limited gene flow between groups of populations but sufficient to prevent allopatric speciation. On the contrary, the extensive geographic distribution might be the testimony of the antiquity of these clades, each being composed of several species (Le Gac et al. 2004). Indeed, almost all clusters found in this study correspond to monophyletic groups (Postaire et al. 2016b), underlining the low dispersal of *L. breviostris* α . Controlled crosses are necessary to assess whether these clusters correspond to biological species or interfertile groups with disjunct distribution ranges leading to effective absence of gene flow. Thus, given the high population differentiation and the importance of IBD over several geographic scales in *L. breviostris* α , a taxonomic revision of *L. breviostris* might consider individuals from each sampled island as a potential new species (Pante et al. 2014, 2015). The apparent absence of morphological clues, a criterion already known to poorly describe the diversity of Aglaopheniidae (Leclère et al. 2007; Moura et al. 2012; Postaire et al. 2016a; Postaire et al. 2016b), can be explained by the maintenance of similar selective pressures and ecologies (Le Gac et al. 2004).

Conclusions

This study is one of the few presenting data from such an extended geographic scale, notably including the eastern margin of the WIO province (Rodrigues). Our study particularly underlines the population isolation in this brooding species, each island potentially representing an independent metapopulation with high dependence on local recruitment for their maintenance. We believe that our approach provides valuable information for the management and creation of marine protected areas in these particular regions. Indeed, the design of marine reserve networks requires an understanding of effective dispersal (*i.e.* transport, recruitment and reproduction) over several scales to apprehend whether populations in reserves are open or self-recruiting and whether reserve networks can exchange recruits (Briggs 2005; Jones et al. 2009; Christie et al. 2010). Considering our results, connectivity of populations between islands or even between reefs on the same island might be extremely low for species without long planktonic life phase, underlining the need for multiple protected areas to preserve evolutionary dynamics in these species (Briggs 2003, 2005; Obura 2012a, 2012b). In this view, the conservation of the coral reefs

of Rodrigues seems particularly important in view of their genetic isolation from other populations of the WIO and the TSP. Future studies on other key benthic marine species, such as scleractinians, molluscs or echinoderms may add support to these findings from hydrozoans.

Acknowledgements

We gratefully acknowledge the Laboratoire d'Excellence CORAIL for financial support. Hydrozoan sampling in New Caledonia (HM) was carried out during Cobelo (doi: <http://dx.doi.org/10.17600/14003700>) and Bibelot (doi: <http://dx.doi.org/10.17600/13100100>) oceanographic campaigns on board of RV Alis (IRD). Sampling in Reunion Island (HM, BP) was supported by program HYDROSOOI (Labex CORAIL fund). Sampling in Madagascar (HM) and Rodrigues (HM) was supported by project Biodiversity (POCT FEDER fund) and sampling in Juan de Nova (HM) was supported by program Biorecie (financial supports from INEE, INSU, IRD, AAMP, FRB, TAAF, and the foundation Veolia Environnement). We gratefully acknowledge the Plateforme Gentyane of the Institut National de la Recherche Agronomique (INRA, Clermont-Ferrand, France) for guidance and technical support. The first author was financially supported by a Ph.D. contract from the STS Doctoral School of the Université de La Réunion.

Conflict of interest:

The authors declare no conflict of interest.

Data archiving:

Data have been submitted to Dryad :

References

- Arnaud-Haond S, Belkhir K (2007) GENCLONE: a computer program to analyse genotypic data, test for clonality and describe spatial clonal organization. *Molecular Ecology Notes* 7:15–17
- Ashe JL, Feldheim KA, Fields AT, Reyier EA, Brooks EJ, O'Connell MT, Skomal G, Gruber SH, Chapman DD (2015) Local population structure and context-dependent isolation by distance in a large coastal shark. *Marine Ecology Progress Series*, 520:203–216
- Aurelle D, Ledoux JB, Rocher C, Borsa P, Chenuil A (2011) Phylogeography of the red coral (*Corallium rubrum*): inferences on the evolutionary history of a temperate gorgonian. *Genetica* 139:855–869
- Ayre DJ, Hughes TP (2000) Genotypic diversity and gene flow in brooding and spawning corals along the Great Barrier Reef, Australia. *Evolution* 54:1590–1605
- Ayre DJ, Minchinton TE, Perrin C (2009) Does life history predict past and current connectivity for rocky intertidal invertebrates across a marine biogeographic barrier? *Molecular Ecology* 18:1887–1903

- Baum DA, Shaw KL (1995) Genealogical perspectives on the species problem. *Experimental and Molecular Approaches to Plant Biosystematics* 53:289–303
- Baums IB, Hughes CR, Hellberg ME (2005) Mendelian microsatellite loci for the Caribbean coral *Acropora palmata*. *Marine Ecology Progress Series* 288:115-127
- Baums IB, Miller MW, Hellberg ME (2006) Geographic variation in clonal structure in a reef-building caribbean coral, *Acropora palmata*. *Ecological Monographs* 76:503-519
- Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate - a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B, Statistical methodology* 57:289–300
- Boero F, Bouillon J, Piraino S (1992) On the origins and evolution of hydromedusan life cycles (Cnidaria, Hydrozoa). *Sex Origin and Evolution* 6:59–68
- Boissin E, Féral J-P, Chenuil A (2008) Defining reproductively isolated units in a cryptic and syntopic species complex using mitochondrial and nuclear markers: the brooding brittle star, *Amphipholis squamata* (Ophiuroidea). *Molecular Ecology* 17:1732–1744
- Boissin E, Hoareau TB, Berrebi P (2011) Effects of current and historic habitat fragmentation on the genetic structure of the sand goby *Pomatoschistus minutus* (Osteichthys, Gobiidae). *Biological Journal of the Linnean Society* 102:175-198
- Bouillon J, Gravili C, Pagès F, Gili J-M, Boero F (2006) An Introduction to Hydrozoa. (P. Grandcolas, J.-M. Betsch, P. Bouchet, & C. Erard, Eds.) (Vol. 194). Paris: Mémoires du Muséum d'Histoire Naturelle.
- Bourjea J, Lapègue S, Gagnevin L, Broderick D, Mortimer JA, Ciccione S, Roos D, Taquet C, Grizel H (2006). Phylogeography of the green turtle, *Chelonia mydas*, in the Southwest Indian Ocean. *Molecular Ecology* 16(1):175–186
- Bowen BW, Rocha LA, Toonen RJ, Karl SA, The Tobo Laboratory (2013) The origins of tropical marine biodiversity. *Trends in Ecology & Evolution* 28:359–366
- Briggs JC (2003) Marine centres of origin as evolutionary engines. *Journal of Biogeography* 30:1–18
- Briggs JC (2005) Coral reefs: Conserving the evolutionary sources. 126:297–305
- Busk G (1852) An account of the Polyzoa and sertularian zoophytes collected in the voyage of the 'rattlesnake' on the coast of Australia and the Louisiade Archipelago. Narrative of the voyage of H.M.S. Rattlesnake, commanded by the late Captain Owen Stanley, R.N., F.R.S. & c. during the years 1846-50 1:343–402
- Di Camillo CG, Puce S, Bavestrello G (2011) *Lytocarpia* and *Cladocarpus* (Cnidaria: Hydrozoa, Aglaopheniidae) from the Bunaken National Marine Park (North Sulawesi, Indonesia). *Marine Biodiversity* 41:517–536
- Chaves-Fonnegra A, Feldheim KA, Secord J, Lopez JV (2015) Population structure and dispersal of the coral-excavating sponge *Cliona delitrix*. *Molecular Ecology*, 24(7):1447–1466
- Christie MR, Tissot BN, Albins MA, Beets JP, Jia Y, Ortiz DM, Thompson SE, Hixon MA (2010) Larval connectivity in an effective network of marine protected areas. *PLoS ONE* 5:e15715
- Cowen RK, Gawarkiewicz G, Pineda J, Thorrold SR, Werner FR (2007) Population connectivity in marine systems. *Oceanography* 20:14–21
- Darling JA, Folino-Rorem NC (2009) Genetic analysis across different spatial scales reveals multiple dispersal mechanisms for the invasive hydrozoan *Cordylophora* in the Great Lakes. *Molecular Ecology* 18:4827–4840
- Dawson M (2001) Phylogeography in coastal marine animals: a solution from California? *Journal of Biogeography* 28:723–736
- DeVantier LM (1992) Rafting of tropical marine organisms on buoyant coralla. *Marine Ecology Progress Series* 86:301–302

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources* 4:359–361

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software structure: a simulation study. *Molecular Ecology* 14:2611–2620

Excoffier L, Laval G, Schneider S (2005) Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary bioinformatics online* 1:47–50

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

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

Fruet PF, Secchi ER, Daura-Jorge F, Vermeulen E, Flores PAC, Simões-Lopes PC, Genoves RC, Laporta P, Di Tullio JC, Freitas TRO, Dalla Rosa L, Valiati VH, Beheregaray LB, Möller LM (2014) Remarkably low genetic diversity and strong population structure in common bottlenose dolphins (*Tursiops truncatus*) from coastal waters of the Southwestern Atlantic Ocean. *Conservation Genetics* 15(4):879–895

Le Gac M, Féral J-P, Poulin E, Veyret M (2004) Identification of allopatric clades in the cosmopolitan ophiuroid species complex *Amphipholis squamata* (Echinodermata). The end of a paradox? *Marine Ecology* 278:171–178

Gili J-M, Hughes RG (1995) The ecology of marine benthic hydroids. *Oceanography and Marine Biology: An Annual Review* 33:351–426

Gravier-Bonnet, N. (2006). Hydroids of New Caledonia from literature study. In C. Payri, B. Richer de Forges, & F. Colin (Eds.), *Compendium of marine species from New Caledonia* (pp. 117–123. (Documents Scientifiques et Techniques – 2.7)).

Gravier-Bonnet N, Bourmaud CA-F (2006a) Hydroids (Cnidaria, Hydrozoa) of Coral Reefs: Preliminary Results on Community Structure, Species Distribution and Reproductive Biology in Juan de Nova Island (Southwest Indian Ocean). *Western Indian Ocean Journal of Marine Science* 5:123–132

Gravier-Bonnet N, Bourmaud CA-F (2006b) Hydroids (Cnidaria, Hydrozoa) of coral reefs: preliminary results on community structure, species distribution and reproductive biology in the île Glorieuses (southwest Indian Ocean). *Proceedings of 10th International Coral Reef Symposium* 188–196

Gravier-Bonnet N, Bourmaud CA-F (2012) Hydroids (Cnidaria, Hydrozoa) of Baa Atoll (Indian Ocean, Maldives Archipelago). *Atoll Research Bulletin* 590:85–123

Grimm V, Reise K, Strasser M (2003) Marine metapopulations: a useful concept? *Helgoland Marine Research* 56:222–228

Hohenlohe PA (2004) Limits to gene flow in marine animals with planktonic larvae: models of *Littorina* species around Point Conception, California. *Biological Journal of the Linnean Society* 82:169–187

Jakobssen M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics* 23:1801–1806

Japaud A, Bouchon C, Manceau JL, Fauvelot C (2015) High clonality in *Acropora palmata* and *Acropora cervicornis* populations of Guadeloupe, French Lesser Antilles. *Marine and Freshwater Research* 66(9):847–851

Johnson MP, Allcock AL, Pye SE, Chambers SJ, Fitton DM (2001) The effects of dispersal mode on the spatial distribution patterns of intertidal molluscs. *Journal of Animal Ecology* 70:641–649

- 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 Genetics* 11:94
- Jones GP, Almany GR, Russ GR, Sale PF, Steneck RS, van Oppen MJH, Willis BL (2009) Larval retention and connectivity among populations of corals and reef fishes: history, advances and challenges. *Coral reefs* 28:307–325
- Kimura M, Weiss GH (1964) The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics* 49:561–576
- Kochzius M, Nuryanto A (2008) Strong genetic population structure in the boring giant clam, *Tridacna corcea*, across the Indo-Malay Archipelago: implications related to evolutionary processes and connectivity. *Molecular Ecology* 17: 3775–3787
- Krishna G, Tolley KA, Groeneveld JC, Matthee CA (2006) Mitochondrial DNA variation in spiny lobster *Palinurus delagoae* suggests genetically structured populations in the southwestern Indian Ocean. *Marine Ecology Progress Series* 319:191–198
- Kyle CJ, Boulding EG (2000) Comparative population genetic structure of marine gastropods (*Littorina* spp.) with and without pelagic larval dispersal. *Marine Biology* 137:835–845
- Leclère L, Schuchert P, Manuel M (2007) Phylogeny of the Plumularioidea (Hydrozoa, Leptothecata): evolution of colonial organisation and life cycle. *Zoologica Scripta* 36:371–394
- Malay MCMD, Paulay G (2010) pPeripatric speciation drives diversification and distributional pattern of reef hermit crabs (Decapoda: Diogenidae: *Calcinus*). *Evolution* 64:634–662
- Mantel N (1967) The detection of disease clustering and a generalized regression approach. *Cancer research* 27:209–220
- Marktanner-Turneretscher G (1890) Hydroiden des K. & K.Naturhistorischen Hofmuseums. *Annalen des K. K. Naturhistorischen Hofmuseums* 5:195–286
- Meirmans PG, van Tienderen PH (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* 4:792–794
- Millard NAH (1975) Monograph on the Hydroida of Southern Africa. *Annals of the South African Museum* 68:513
- Mokhtar-Jamāi K, Pascual M, Ledoux JB, Coma R, Féral J-P, Garrabou J, Aurelle D (2011) From global to local genetic structuring in the red gorgonian *Paramuricea clavata*: the interplay between oceanographic conditions and limited larval dispersal. *Molecular Ecology* 20:3291–3305
- Moura CJ, Cunha MR, Porteiro FM, Rogers AD (2012) A molecular phylogenetic appraisal of the systematics of the Aglaopheniidae (Cnidaria: Hydrozoa, Leptothecata) from the north-east Atlantic and west Mediterranean. *Zoological Journal of the Linnean Society* 164:717–727
- Muths D, Tessier E, Gouws G, Craig MT, Mwale M, Mwaluma J, Bourjea J (2011) Restricted dispersal of the reef fish *Myripristis berndti* at the scale of the SW Indian Ocean. *Marine Ecology Progress Series* 443:167–180
- Obura DO (2012a) The diversity and biogeography of Western Indian Ocean reef-building corals. *PLoS ONE* 7(9):e45013
- Obura DO (2012b) Evolutionary mechanisms and diversity in a western Indian Ocean center of diversity. *Proceedings of the 12th International Coral Reef Symposium, Cairns, Australia, 9-13 July 2012*
- Pous S, Lazure P, André G, Dumas F, Halo I, Penven P (2014) Circulation around La Réunion and Mauritius islands in the south-western Indian Ocean: A modeling perspective. *Journal of Geophysical Research: Oceans* 119(3):1957–1976
- Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. *Molecular Ecology Notes* 4:535–538

Palumbi SR (1992) Marine speciation on a small planet. *Trends in Ecology & Evolution* 7:114–118

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. *Molecular Ecology* 24:525–544

Pante E, Schoelinck C, Puillandre N (2014) From integrative taxonomy to species description: one step beyond. *Systematic Biology* 64:152–160

Paulay G, Meyer CP (2002) Diversification in the tropical Pacific: comparisons between marine and terrestrial systems and the importance of founder speciation. *Integrative and Comparative Biology* 42:922–934

Pinzón JH, Reyes-Bonilla H, Baums IB, LaJeunesse TC (2012) Contrasting clonal structure among *Pocillopora* (Scleractinia) communities at two environmentally distinct sites in the Gulf of California. *Coral Reefs* 31:765–777

Postaire B, Aurelle D, Bourmaud CA-F, Bruggemann JH, Magalon H (2015a) Isolation and characterisation of 16 microsatellite loci from a widespread tropical hydrozoan, *Lytocarpia breviostris* (Busk, 1852). *Conservation Genetics Resources* 7:505–507

Postaire B, Aurelle D, Bourmaud CA-F, Bruggemann JH, Magalon H (2015b) Isolation and characterisation of 26 microsatellite loci from a widespread tropical hydrozoan, *Macrorhynchia phoenicea* (Leptothecata, Aglaopheniidae), and cross-amplification in closely related species. *Biochemical Systematics and Ecology*, 62(C), 137–141.

Postaire B, Magalon H, Bourmaud CAF, Gravier-Bonnet N, Bruggemann JH (2016a) Phylogenetic relationships within Aglaopheniidae (Cnidaria, Hydrozoa) reveal unexpected generic diversity. *Zoologica Scripta* 45:103–114

Postaire B, Magalon H, Bourmaud CAF, Bruggemann JH (2016b) Molecular species delimitation methods and population genetics data reveal extensive lineage diversity and cryptic species in Aglaopheniidae (Hydrozoa). *Molecular Phylogenetics and Evolution* (105): 36–49

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

De Queiroz K (1998) The general concept of species, species criteria, and the process of speciation: a conceptual unification and terminological recommendations. *Endless Forms: Species and Speciation* 57–75

De Queiroz K (2005) Ernst Mayr and the Modern Concept of Species. *Proceedings of the National Academy of Sciences of the United States of America* 102:6600–6607

R Development Core Team (2004) R: A language and environment for statistical computing.

Ragionieri L, Cannicci S, Schubart CD, Fratini S (2010) Gene flow and demographic history of the mangrove crab *Neosarmatium meinerti*: A case study from the western Indian Ocean. *Estuarine, Coastal and Shelf Science* 86:179–188

Richards ZT, Berry O, van Oppen MJH (2016) Cryptic genetic divergence within threatened species of *Acropora* coral from the Indian and Pacific Oceans. *Conservation Genetics* 17(3): 577–591

Ridgway T, Riginos C, Davis J, Hoegh-Guldberg O (2008) Genetic connectivity patterns of *Pocillopora verrucosa* in southern African Marine Protected Areas. *Marine Ecology Progress Series* 354:161–168

Rocha S, Carretero MA, Vences M, Glaw F, James Harris D (2006) Deciphering patterns of transoceanic dispersal: the evolutionary origin and biogeography of coastal lizards (*Cryptoblepharus*) in the Western Indian Ocean region. *Journal of Biogeography* 33:13–22

Rosenberg NA (2004) Distruct: a program for the graphical display of population structure. *Molecular Ecology Notes* 4:137–138

- Rousset F (1997) Genetic differentiation and estimation of gene flow from F-statistics under isolation by distance. *Genetics* 145:1219–1228
- Sale PF, Cowen RK, Danilowicz BS, Jones GP, Kritzer J, Lindeman K, Planes S, Polunin N, Russ GR, Sadovy Y (2005) Critical science gaps impede use of no-take fishery reserves. *Trends in Ecology & Evolution* 20:74–80
- Samadi, S, Barberousse A (2006) The tree, the network, and the species. *Biological Journal of the Linnean Society* 89:509–521.
- Schiavina M, Marino IAM, Zane L, Melià P (2014) Matching oceanography and genetics at the basin scale. Seascape connectivity of the Mediterranean shore crab in the Adriatic Sea. *Molecular Ecology* 23:5496–5507
- Schott FA, Xie S-P, McCreary JP Jr (2009) Indian Ocean circulation and climate variability. *Reviews of Geophysics* 47(1): RG1002–46.
- Schuchert P (2005) Species boundaries in the hydrozoan genus *Coryne*. *Molecular Phylogenetics and Evolution* 36(1): 194–199
- Schuchert P (2014) High genetic diversity in the hydroid *Plumularia setacea*: A multitude of cryptic species or extensive population subdivision? *Molecular Phylogenetics and Evolution* 76:1–9
- Sethuraman A, McGaugh SE, Becker ML, Chandler CH, Christiansen JL, Hayden S, Leclère A, Monson-Miller J, Myers EM, Paitz RT, Refsnider JM, VanDeWalle TJ, Janzen FJ (2013) Population genetics of Blanding's turtle (*Emys blandingii*) in the midwestern United States. *Conservation Genetics*, 15(1):61–73.
- Shanks AL (2009) Pelagic larval duration and dispersal distance revisited. *Biological Bulletin* 216:373–385
- Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution* 47:264–279
- Sommer C (1990) Post-embryonic larval development and metamorphosis of the hydroid *Eudendrium racemosum* (Cavolini) (Hydrozoa, Cnidaria). *Helgoländer Meeresuntersuchungen* 44:425–444
- Spalding M, Fox HE, Allen GR, Davidson N, Ferdaña ZA, Finlayson M, Halpern BS, Jorge MA, Lombana AL, Lourie SA, Martin KD, McManus E, Molnar J, Recchia CA, Robertson J (2007) Marine ecoregions of the world: a bioregionalization of coastal and shelf areas. *BioScience* 57:573–583
- Szpiech ZA, Jakobsson M, Rosenberg NA (2008). ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics* 24(21):2498–2504
- Thiel M, Gutow L (2005) The ecology of rafting in the marine environment. II. The rafting organisms and community. *Oceanography and Marine Biology: An Annual Review* 43:279–418
- Thiel M, Haye PA (2006) The ecology of rafting in the marine environment. III. Biogeographical and evolutionary consequences. *Oceanography and Marine Biology: An Annual Review* 44:323–429
- Thomas Y, Le Gendre R, Garen P, Dumas F, Andréfouët S (2012) Bivalve larvae transport and connectivity within the Ahe atoll lagoon (Tuamotu Archipelago), with application to pearl oyster aquaculture management. *Marine Pollution Bulletin* 65:441–452
- Treml EA, Halpin PN, Urban DL, Pratson LF (2007) Modeling population connectivity by ocean currents, a graph-theoretic approach for marine conservation. *Landscape Ecology* 23:19–36
- Underwood JN, Smith LD, van Oppen MJH, Gilmour JP (2007) Multiple scales of genetic connectivity in a brooding coral on isolated reefs following catastrophic bleaching. *Molecular Ecology* 16:771–784
- Vanormelingen P, Evans KM, Mann DG, Lance S, Debeer A-E, D'hondt S, Verstraete T, De Meester L, Vyverman W (2015) Genotypic diversity and differentiation among populations of two benthic freshwater diatoms as revealed by microsatellites. *Molecular Ecology*, 24(17):4433–4448

Weersing K, Toonen RJ (2009) Population genetics, larval dispersal, and connectivity in marine systems. *Marine Ecology Progress Series* 393:1–12

White C, Selkoe KA, Watson J, Siegel DA, Zacherl DC, Toonen RJ (2010) Ocean currents help explain population genetic structure. *Philosophical Transactions of the Royal Society, B: Biological Sciences* 277:1685–1694

Yasuda N, Nagai S, Hamaguchi M, Okaji K, Gérard K, Nadaoka K (2009) Gene flow of *Acanthaster planci* (L.) in relation to ocean currents revealed by microsatellite analysis. *Molecular Ecology* 18: 1574-1590.

Yund PO (1990) An *in situ* measurement of sperm dispersal in a colonial marine hydroid. *Journal of Experimental Zoology* 253:102–106

Titles and legends to figures

Figure 1: Sampling locations of *Lytocarpia brevirostris* α in the Western Indian Ocean and the Tropical Southwestern Pacific with population names and the number of individuals sampled (in parentheses).

Figure 2: Distribution of the difference in field numbers (sequentially numbered sampling bags) between pairs of individuals presenting the same multi-locus genotype.

Figure 3: *Lytocarpia brevirostris* α . Assignment probabilities of multi-locus genotypes (MLG) to putative clusters using an admixture model. (A) Mean $\text{LnP}(K)$ values. (B) Average probability of membership (y-axis) of MLGs ($N = 470$, x-axis) in $K = 5$ clusters as identified by STRUCTURE. G.-T.: Grande-Terre.

Figure 4: *Lytocarpia brevirostris* α . Discriminant analysis of principal components (DAPC) of multi-locus genotypes (MLG) sampled in the Western Indian Ocean and the Tropical Southwestern Pacific. (A) The Bayesian information criterion. (B) Scatter plots of the MLGs using the first and second components and (C) the second and third components.

Figure 5: *Lytocarpia brevirostris* α . Correlation between genetic distances computed as $F_{ST}/(1 - F_{ST})$ and the \log_{10} of geographical distances (in kilometres) between site pairs at the in the Western Indian Ocean and the Tropical Southwestern Pacific.

Figure 1.

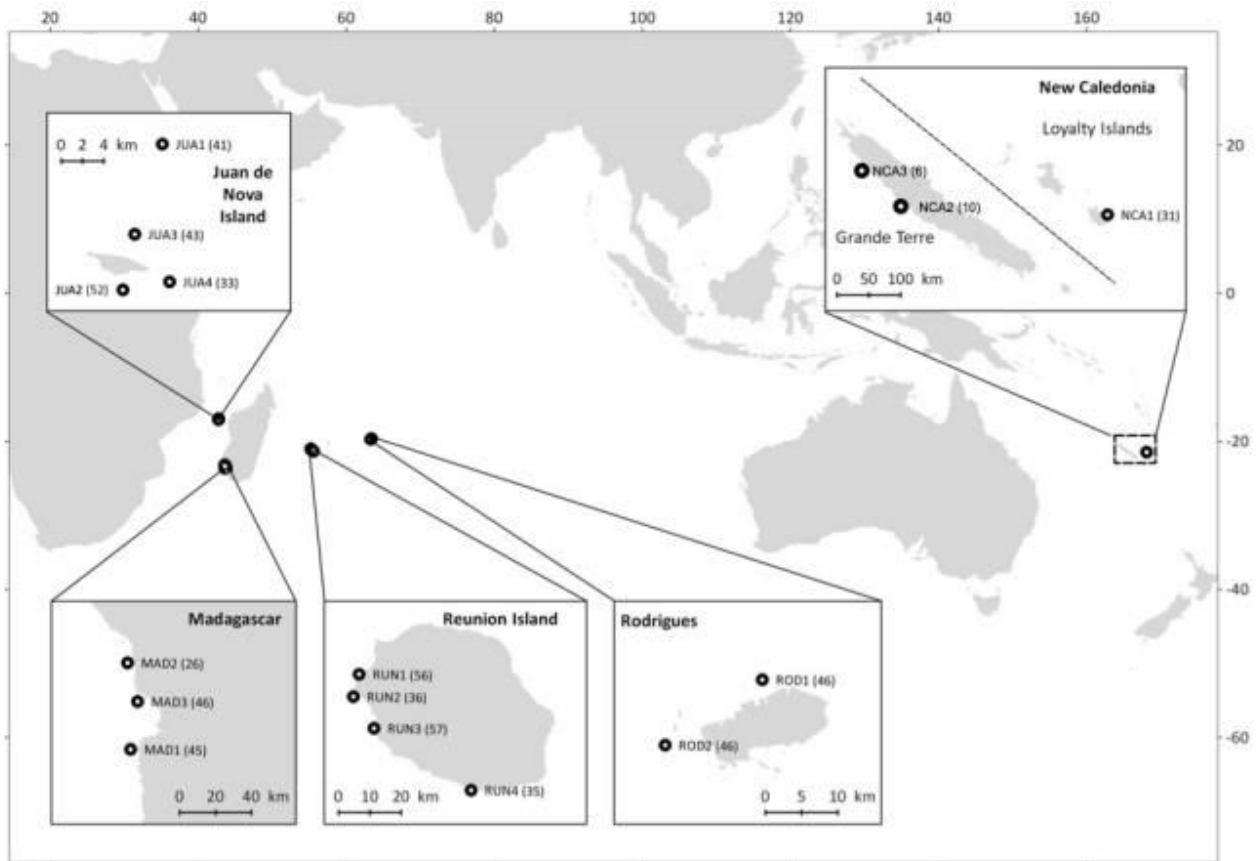


Figure 2.

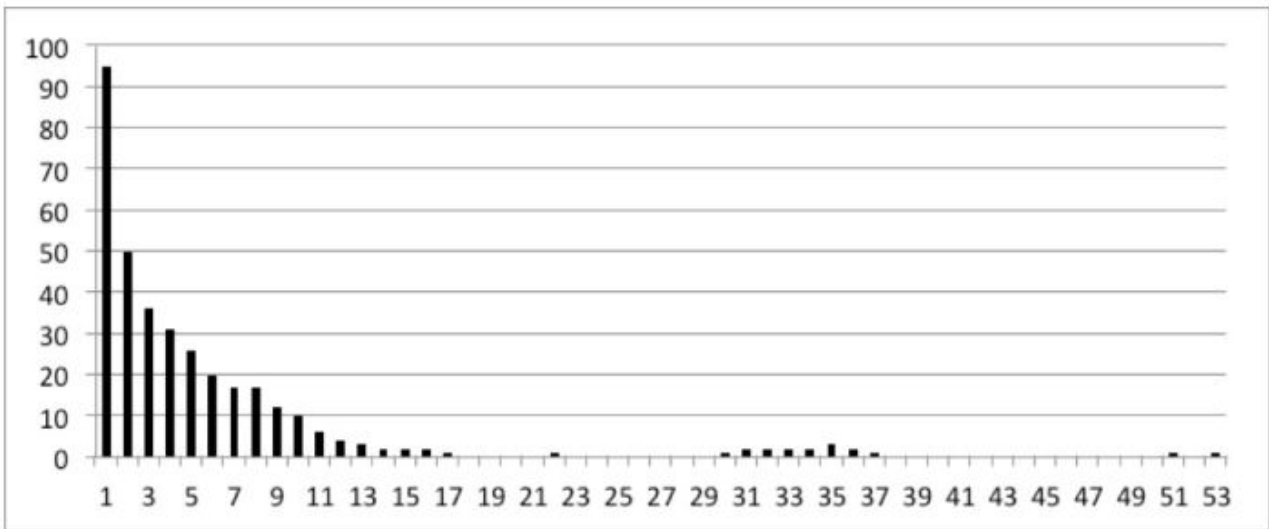
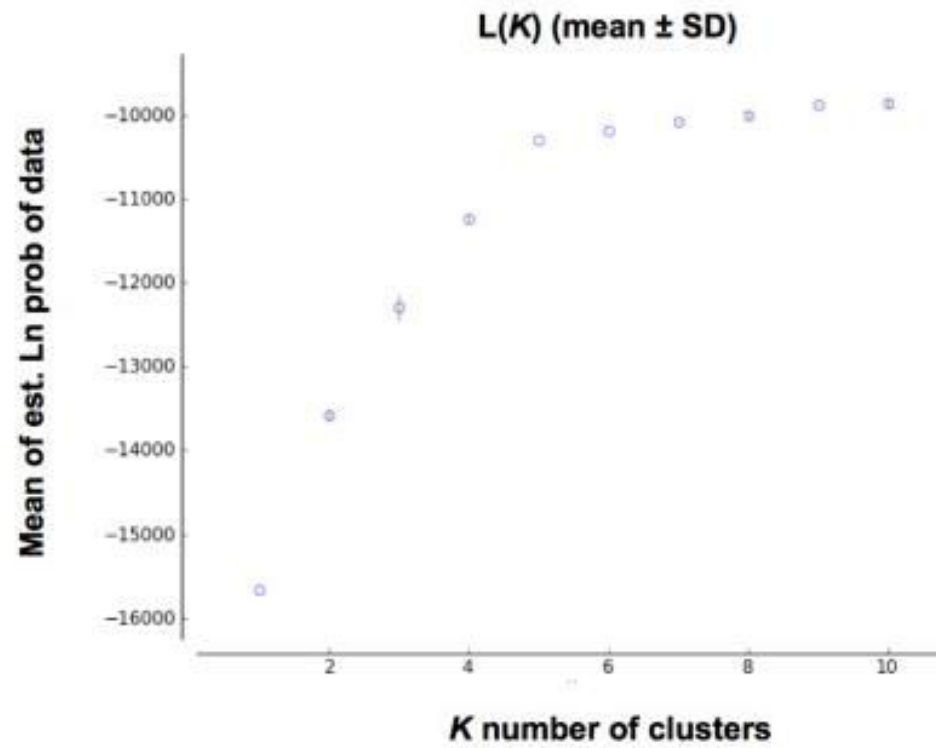


Figure 3.

A



B

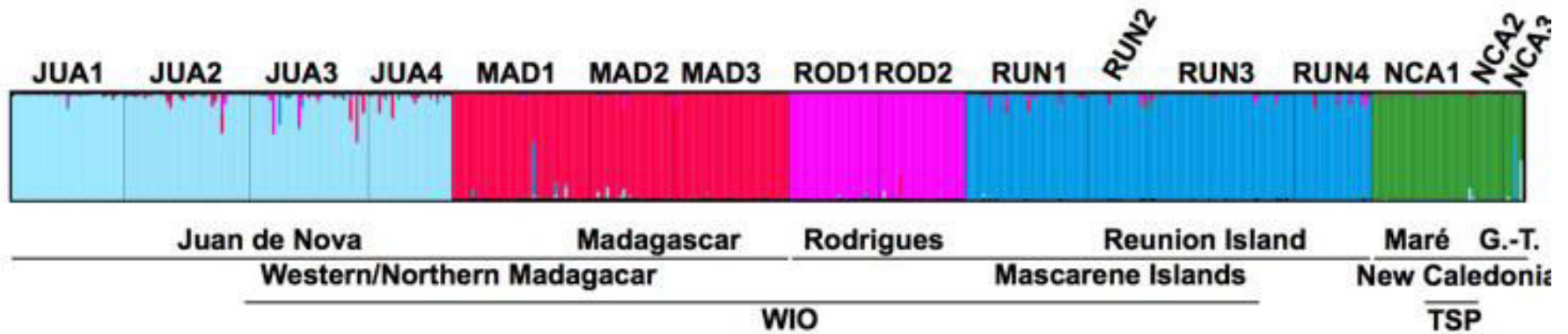


Figure 4.

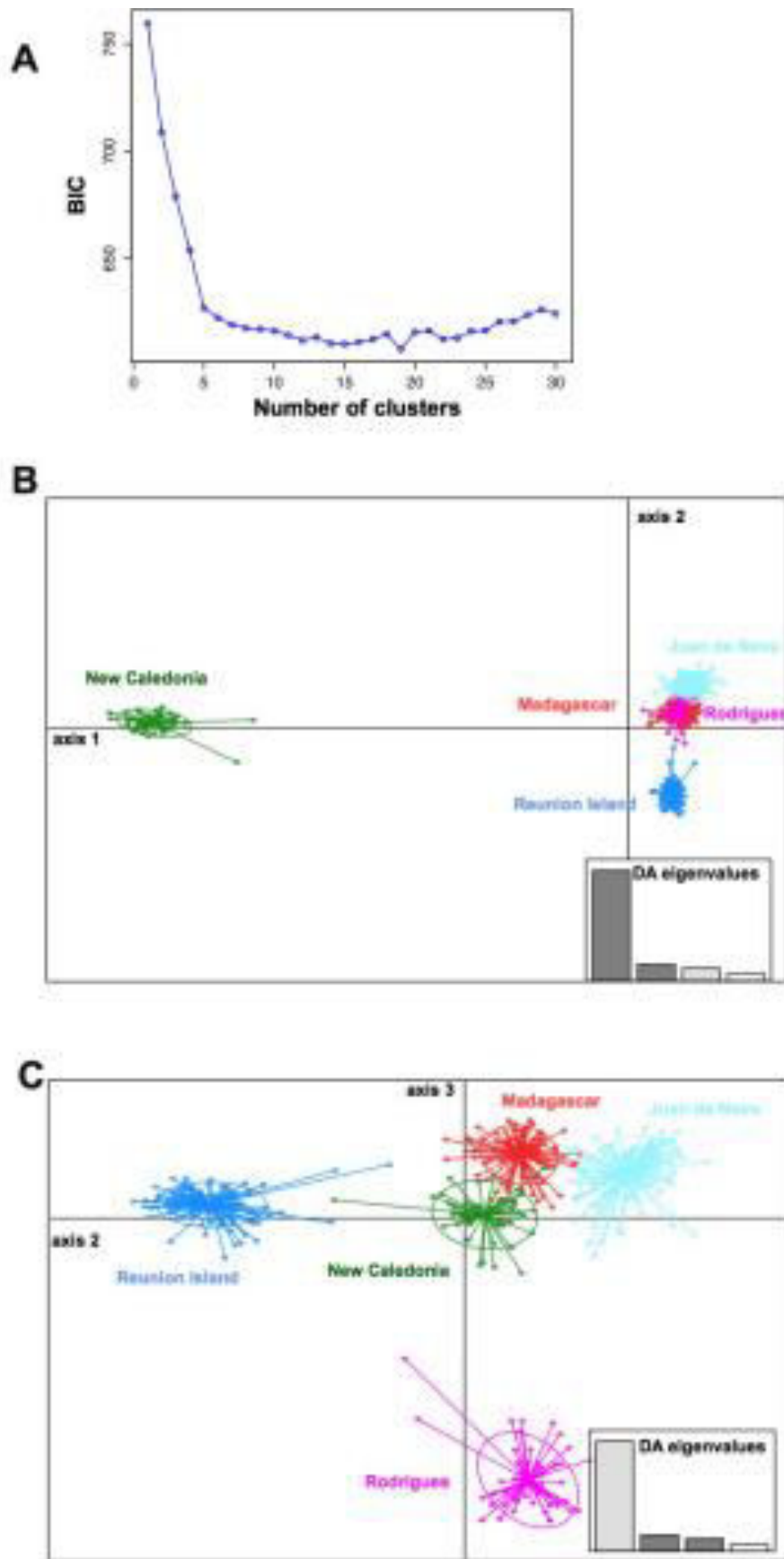
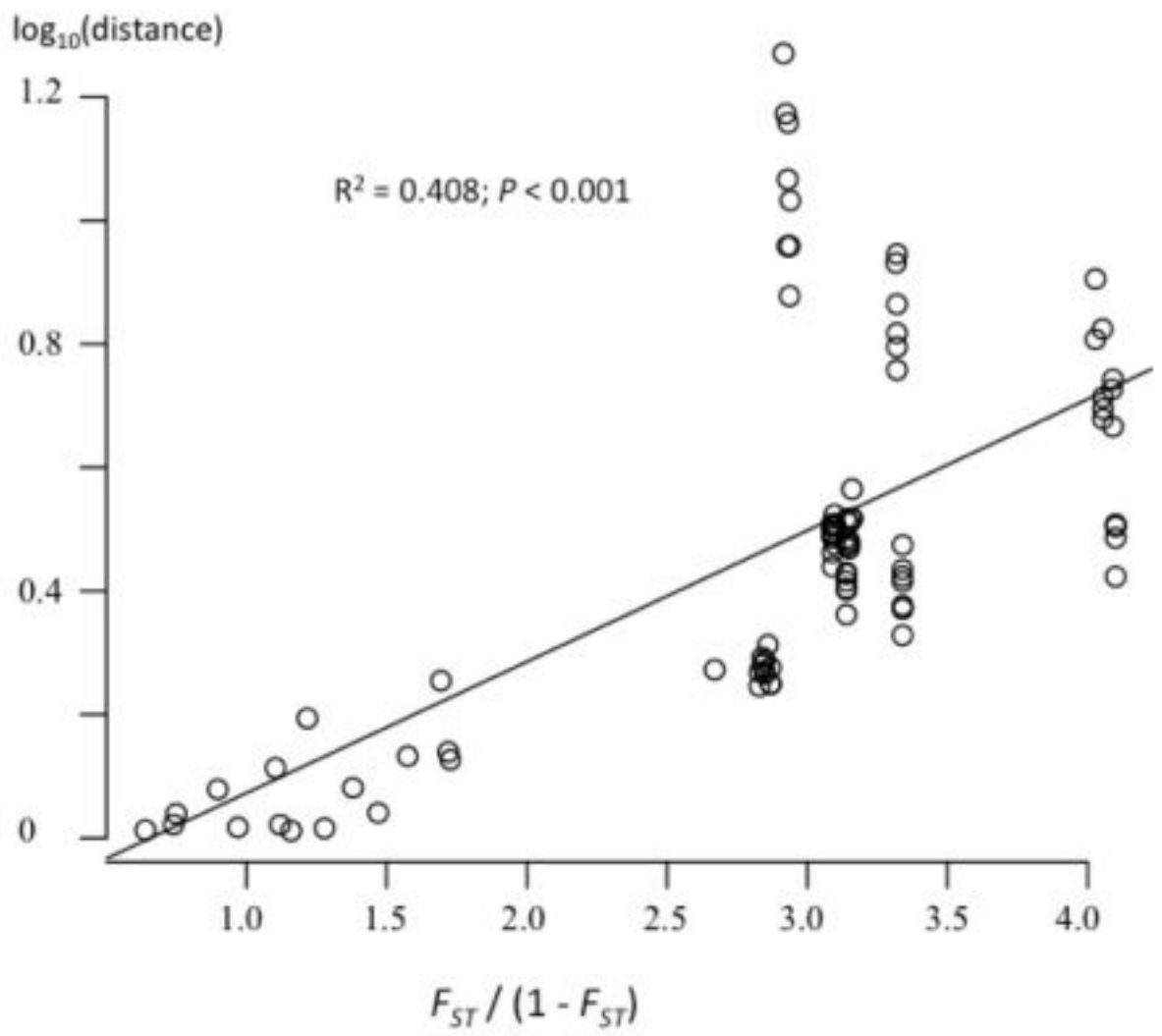


Figure 5.

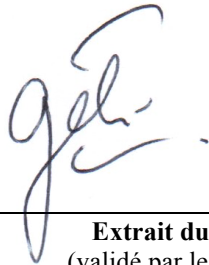


LETTRE D'ENGAGEMENT DE NON-PLAGIAT

Je, soussigné(e) **Pauline GELIN**,
en ma qualité de doctorant(e) de l'Université de La Réunion, déclare être conscient(e) que le
plagiat est un acte délictueux passible de sanctions disciplinaires. Aussi, dans le respect de la
propriété intellectuelle et du droit d'auteur, je m'engage à systématiquement citer mes
sources, quelle qu'en soit la forme (textes, images, audiovisuel, internet), dans le cadre de la
rédaction de ma thèse et de toute autre production scientifique, sachant que l'établissement est
susceptible de soumettre le texte de ma thèse à un logiciel anti-plagiat.

Fait à **Saint Denis**, le (date) **18 novembre 2016**

Signature :



Extrait du Règlement intérieur de l'Université de La Réunion
(validé par le Conseil d'Administration en date du 11 décembre 2014)

Article 9. Protection de la propriété intellectuelle – Faux et usage de faux, contrefaçon, plagiat

L'utilisation des ressources informatiques de l'Université implique le respect de ses droits de propriété intellectuelle ainsi que ceux de ses partenaires et plus généralement, de tous tiers titulaires de tels droits.

En conséquence, chaque utilisateur doit :

- utiliser les logiciels dans les conditions de licences souscrites ;
- ne pas reproduire, copier, diffuser, modifier ou utiliser des logiciels, bases de données, pages Web, textes, images, photographies ou autres créations protégées par le droit d'auteur ou un droit privatif, sans avoir obtenu préalablement l'autorisation des titulaires de ces droits.

La contrefaçon et le faux

Conformément aux dispositions du code de la propriété intellectuelle, toute représentation ou reproduction intégrale ou partielle d'une œuvre de l'esprit faite sans le consentement de son auteur est illicite et constitue un délit pénal.

L'article 444-1 du code pénal dispose : « Constitue un faux toute altération frauduleuse de la vérité, de nature à causer un préjudice et accomplie par quelque moyen que ce soit, dans un écrit ou tout autre support d'expression de la pensée qui a pour objet ou qui peut avoir pour effet d'établir la preuve d'un droit ou d'un fait ayant des conséquences juridiques ».

L'article L335_3 du code de la propriété intellectuelle précise que : « Est également un délit de contrefaçon toute reproduction, représentation ou diffusion, par quelque moyen que ce soit, d'une œuvre de l'esprit en violation des droits de l'auteur, tels qu'ils sont définis et réglementés par la loi. Est également un délit de contrefaçon la violation de l'un des droits de l'auteur d'un logiciel (...) ».

Le plagiat est constitué par la copie, totale ou partielle d'un travail réalisé par autrui, lorsque la source empruntée n'est pas citée, quel que soit le moyen utilisé. Le plagiat constitue une violation du droit d'auteur (au sens des articles L 335-2 et L 335-3 du code de la propriété intellectuelle). Il peut être assimilé à un délit de contrefaçon. C'est aussi une faute disciplinaire, susceptible d'entraîner une sanction.

Les sources et les références utilisées dans le cadre des travaux (préparations, devoirs, mémoires, thèses, rapports de stage...) doivent être clairement citées. Des citations intégrales peuvent figurer dans les documents rendus, si elles sont assorties de leur référence (nom d'auteur, publication, date, éditeur...) et identifiées comme telles par des guillemets ou des italiques.

Les délits de contrefaçon, de plagiat et d'usage de faux peuvent donner lieu à une sanction disciplinaire indépendante de la mise en œuvre de poursuites pénales.

Résumé

Dans un monde changeant, la conservation et la gestion de la biodiversité sur le long terme nécessitent une connaissance de la répartition des taxons, mais également une estimation précise de leur diversité spécifique ainsi que des processus de spéciation ayant permis leur formation et des liens qui unissent différentes populations d'un même taxon (connectivité). En génétique des populations, ces liens sont les flux de gènes entre différentes populations. Il est crucial de connaître le degré de connectivité des populations, et ce pour différentes espèces marines, afin de déterminer les patrons de dispersion à différentes échelles spatiales et temporelles. La connaissance de ces patrons permet de dessiner des réseaux d'aires marines protégées de façon plus efficace afin de préserver et gérer la biodiversité.

Ce travail de thèse porte sur la connectivité des populations de coraux du genre *Pocillopora* dans le Sud-Ouest de l'océan Indien et l'océan Pacifique tropical. Ces coraux sont répartis sur toute la frange tropicale des océans Indien et Pacifique. Traditionnellement, les espèces étaient identifiées sur la base critères morphologies [17 espèces décrites dans Veron (2000)]. Différentes études utilisant des données génétiques ont révélé que la délimitation des espèces était parfois floue chez ces coraux. Ainsi, au cours de ce travail, l'utilisation de méthodes de délimitation d'espèces à partir d'ADN mitochondrial (ABGD, GMYC, PTP) et nucléaire (haplowebs) 16 hypothèses primaires d'espèces (PSH) ont été identifiées. Ces PSH ont ensuite été confrontées à des tests d'assignement à partir de marqueurs microsatellites, révélant un minimum de 18 hypothèses d'espèces secondaires (SSH).

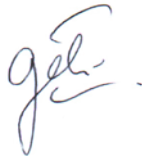
Une fois que les hypothèses d'espèces sont définies, il est possible de réaliser des études de connectivité. Au cours de ce travail, deux hypothèses d'espèces présentant des écologies différentes ont été choisies pour mener ces analyses. La première, *Pocillopora damicornis* type β (SSH05) a été échantillonnée dans les lagons et la seconde, *Pocillopora eydouxi* (SSH09) a, quant à elle, été échantillonnée sur la pente externe. L'estimation de la structure génétique des populations a permis d'estimer les modes de reproduction (sexuée ou asexuée) chez ces deux hypothèses d'espèces et les analyses de connectivité ont révélé des patterns de structuration complexes pour chacune des SSHs.

Mots-clés : génétique des populations, hypothèse d'espèces, méthodes de délimitation d'espèces, tests d'assignement, scléactiniaires, Sud-Ouest de l'océan Indien, Sud-Ouest de l'océan Pacifique, microsatellites, séquences

LETTRE D'ENGAGEMENT DE NON-PLAGIAT

Je, soussigné(e) Pauline GELIN, en ma qualité de doctorant(e) de l'Université de La Réunion, déclare être conscient(e) que le plagiat est un acte délictueux passible de sanctions disciplinaires. Aussi, dans le respect de la propriété intellectuelle et du droit d'auteur, je m'engage à systématiquement citer mes sources, quelle qu'en soit la forme (textes, images, audiovisuel, internet), dans le cadre de la rédaction de ma thèse et de toute autre production scientifique, sachant que l'établissement est susceptible de soumettre le texte de ma thèse à un logiciel anti-plagiat.

Fait à Saint Denis , le 22/11/2016



Extrait du Règlement intérieur de l'Université de La Réunion

(validé par le Conseil d'Administration en date du 11 décembre 2014)

Article 9. Protection de la propriété intellectuelle – Faux et usage de faux, contrefaçon, plagiat

L'utilisation des ressources informatiques de l'Université implique le respect de ses droits de propriété intellectuelle ainsi que ceux de ses partenaires et plus généralement, de tous tiers titulaires de ces droits.

En conséquence, chaque utilisateur doit :

- utiliser les logiciels dans les conditions de licences souscrites ;
- ne pas reproduire, copier, diffuser, modifier ou utiliser des logiciels, bases de données, pages Web, textes, images, photographies ou autres créations protégées par le droit d'auteur ou un droit privatif, sans avoir obtenu préalablement l'autorisation des titulaires de ces droits.

La contrefaçon et le faux

Conformément aux dispositions du code de la propriété intellectuelle, toute représentation ou reproduction intégrale ou partielle d'une œuvre de l'esprit faite sans le consentement de son auteur est illicite et constitue un délit pénal.

L'article 444-1 du code pénal dispose : « Constitue un faux toute altération frauduleuse de la vérité, de nature à causer un préjudice et accomplie par quelque moyen que ce soit, dans un écrit ou tout autre support d'expression de la pensée qui a pour objet ou qui peut avoir pour effet d'établir la preuve d'un droit ou d'un fait ayant des conséquences juridiques ».

L'article L335_3 du code de la propriété intellectuelle précise que : « Est également un délit de contrefaçon toute reproduction, représentation ou diffusion, par quelque moyen que ce soit, d'une œuvre de l'esprit en violation des droits de l'auteur, tels qu'ils sont définis et réglementés par la loi. Est également un délit de contrefaçon la violation de l'un des droits de l'auteur d'un logiciel (...) ».

Le plagiat est constitué par la copie, totale ou partielle d'un travail réalisé par autrui, lorsque la source empruntée n'est pas citée, quel que soit le moyen utilisé. Le plagiat constitue une violation du droit d'auteur (au sens des articles L 335-2 et L 335-3 du code de la propriété intellectuelle). Il peut être assimilé à un délit de contrefaçon. C'est aussi une faute disciplinaire, susceptible d'entraîner une sanction.

Les sources et les références utilisées dans le cadre des travaux (préparations, devoirs, mémoires, thèses, rapports de stage...) doivent être clairement citées. Des citations intégrales peuvent figurer dans les documents rendus, si elles sont assorties de leur référence (nom d'auteur, publication, date, éditeur...) et identifiées comme telles par des guillemets ou des italiques.

Les délits de contrefaçon, de plagiat et d'usage de faux peuvent donner lieu à une sanction disciplinaire indépendante de la mise en œuvre de poursuites pénales.