



Thèse de doctorat

Aix Marseille Université

École doctorale 251 : Sciences de l'Environnement

Spécialité "Océanographie"

présentée et soutenue publiquement par

Marine PRATLONG

le 05 décembre 2016

Le défi évolutif du changement climatique, processus adaptatifs chez le corail rouge (*Corallium rubrum*)

Membres du jury :

Nicolas BIERNE (rapporteur)

Brigitte CROUAU-ROY (rapporteur)

Nicolas GALTIER (examineur)

Jean-Christophe POGGIALE (examineur)

Ève TOULZA (examineur)

Didier AURELLE (directeur de thèse)

Pierre PONTAROTTI (directeur de thèse)



Crédit photographie de couverture : Frédéric Zuberer

Table des matières

1	Chapitre 1 : Introduction	11
1.1	Qu'est-ce que l'adaptation locale?	11
1.1.1	Origine du concept d'adaptation	11
1.1.2	Adaptation locale	14
1.1.3	Impact des autres forces évolutives sur l'adaptation locale	15
1.2	Comment détecter l'adaptation locale?	18
1.2.1	Études expérimentales	18
1.2.2	La génétique de l'adaptation locale	19
1.3	Exemples en milieu marin	23
1.3.1	Études expérimentales	23
1.3.2	Approches de génomique des populations	30
1.4	Le corail rouge <i>Corallium rubrum</i>	33
1.4.1	Description de l'espèce	33
1.4.2	La génétique du corail rouge	35
1.4.3	Thermotolérance chez le corail rouge	36
1.5	Objectifs	40
2	Recherche de différences transcriptomiques à l'état basal entre individus issus d'environnements contrastés	44
2.1	Introduction	48
2.2	Material and Methods	50
2.2.1	Sampling and RNA extraction	50
2.2.2	Library preparation and sequencing	51
2.2.3	Assembly	51
2.2.4	Meta-assembly	52
2.2.5	Comparative analysis	52
2.2.6	Transcriptome annotation	53

2.2.7	Differentially expressed genes	53
2.2.8	Detection and analysis of SNPs	55
2.3	Results	56
2.3.1	Transcriptome description	56
2.3.2	Comparative analysis	57
2.3.3	Analysis of depth specific expression patterns	57
2.3.4	Polymorphism analysis	58
2.4	Discussion	61
2.4.1	Transcriptome comparative analysis	62
2.4.2	Biological functions involved in coral adaptation to thermal stress	62
2.4.3	Phenotypic plasticity or genetic adaptation?	64
2.4.4	Polymorphism pattern	64
2.5	Conclusion	65
2.6	Acknowledgements	66
2.7	Appendices	66
3	Identification de locus d'adaptation locale dans un contexte de forte structure génétique	68
3.1	Introduction	72
3.2	Material and methods	76
3.2.1	Sampling and DNA extraction	76
3.2.2	RAD-Sequencing	77
3.2.3	Diversity and neutral genetic structure	78
3.2.4	Detection of local adaptation	79
3.2.5	Functional annotation and enrichment tests	80
3.3	Results	80
3.3.1	RAD-Sequencing and genotyping	80
3.3.2	Genetic diversity	81
3.3.3	Population structure analysis	81
3.3.4	SNPs outliers	85
3.3.5	Functional annotation	86
3.4	Discussion	86
3.4.1	Genetic diversity and structure	86
3.4.2	Potential biases in the search of outlier loci	88
3.4.3	Local adaptation to depth in the red coral	89
3.5	Conclusion	91

3.6	Acknowledgements	92
4	Mise en évidence d'un déterminisme génétique du sexe chez le corail rouge	93
4.1	Results and Discussion	96
4.1.1	Morphological identification of sex	96
4.1.2	Identification of sex-linked loci	97
4.1.3	Identification of sex-specific loci	98
4.1.4	Cross-validation	102
4.2	Conclusion	102
4.3	Experimental procedures	103
4.3.1	Sampling and DNA extraction	103
4.3.2	Morphological sex identification	104
4.3.3	RAD-Sequencing	104
4.3.4	Identification of sex-linked loci	105
4.3.5	Identification of sex-specific loci	105
4.3.6	Real-time PCR	106
4.4	Authors contributions	106
4.5	Acknowledgments	106
5	Discussion générale	108
5.1	Résumé des résultats obtenus	108
5.2	Pertinence de l'interprétation fonctionnelle des gènes candidats pour l'adaptation locale	110
5.3	Quelle(s) origine(s) aux différences d'expression de gènes observée(s) entre profondeurs ?	111
5.3.1	Stratégies de résistance au stress thermique	112
5.3.2	Rôle des modifications épigénétiques	112
5.4	Obstacles méthodologiques à l'étude de l'adaptation locale chez le corail rouge	114
5.5	Déterminisme génétique du sexe	115
5.6	Perspectives	116
5.6.1	Liens avec le phénotype	116
5.6.2	Apport du génome du corail rouge	117
5.6.3	Les autres gorgones méditerranéennes comme modèle d'étude ?	119
5.7	Conservation	120

5.8 Conclusion	123
A Phylogénie des anthozoaires (Cnidaires) : comment expliquer les différences entre phylogénies nucléaires et mitochondriales	I
A.1 Introduction	III
A.2 Material and Methods	V
A.2.1 Markers selection	V
A.2.2 Phylogenetic analyses	VIII
A.2.3 Substitution saturation analysis	VIII
A.3 Results and discussion	X
A.3.1 Nuclear phylogeny of Cnidaria	X
A.3.2 Possible origin of the discrepancy between nuclear and mitochondrial phylogenies	XII
A.4 Conclusion	XV
A.5 Acknowledgements	XV
A.6 Appendices	XVI
B Annexes du Chapitre 2	XXI
C Annexes du Chapitre 3	XXVII
C.1 Figures et Tables Supplémentaires	XXVII
C.2 Aspect phylogéographique - corail rouge d'Atlantique	XLII
C.3 Analyses exploratoires réalisées non présentées dans l'article	XLIII
C.3.1 Recherche statistique d'un effet "profondeur" dans les données de polymorphisme issues de RAD-Séquençage	XLIII
C.3.2 Autres tests de recherche de marqueurs d'adaptation locale	XLVIII
D Annexes du Chapitre 4	L
Références	XCIII

Remerciements

Tout d'abord, je souhaite remercier mes deux directeurs de thèse, Didier Aurelle et Pierre Pontarotti. Vous m'avez fait confiance et avez été toujours présents et à l'écoute pendant ma thèse, j'ai beaucoup appris à vos côtés. Vous êtes, par de nombreux aspects, très complémentaires dans vos approches, et cela a été très agréable de travailler avec vous. Didier, merci de m'avoir donné une chance une première fois pour mon stage volontaire en Master 1, puis en Master 2 et en thèse. On sous-estime souvent l'impact des enseignements et des enseignants à l'université, dans mon cas il est évident que je ne serais jamais arrivée là sans ma découverte de la génétique des populations grâce à toi en Master 1. Malgré ton emploi du temps souvent chargé, tu as toujours pris le temps de répondre à mes questions et de discuter tout au long des trois années. Pierre, tu es toujours disponible et j'ai beaucoup apprécié nos discussions sur ma thèse, mais aussi sur le fonctionnement du monde de la recherche en général. Ton expérience et ta façon de faire fonctionner le labo m'ont beaucoup ouvert l'esprit sur la recherche en dehors de mon domaine d'étude. Tu as toujours été très pragmatique dans ta façon d'aborder mon travail, de bien poser les questions scientifiques et d'identifier les réponses que l'on pouvait apporter avec nos données. Cette vision m'a souvent aidé à répondre à mes inquiétudes dans les moments où j'étais un peu perdu. Encore merci à vous deux.

J'ai eu l'opportunité de réaliser cette thèse grâce à un financement OT-Med, je tiens donc à remercier les membres du labex et particulièrement Joël Guiot pour m'avoir fait confiance et permis de raccrocher ma thèse à des problématiques plus globales. Un grand merci à Katarzyna Tarnowska, Sophie Bonifay-Pekar et Morgane Soler pour leur dynamisme, leur investissement et leur gentillesse.

Mes remerciements vont aussi à Anne Haguénauer et Olivier Chabrol, mes mentors de la biomol et de la bioinfo, sans qui, il faut le dire, je ne serais pas allée bien loin dans ma thèse. Anne, tu m'as permis de repartir dans les manip les maintes fois où j'étais découragée, toujours à tenter de nouvelles choses, à persévérer, à trouver des solutions. Tu m'as accompagnée pendant les longues séances de sonication et de bioanalyser. Sans la reine de la biomol à mes côtés, je ne serais jamais arrivée au bout de mes préparations de bibliothèques. Et que dire de son homologue Sherlock qui, on ne sait pour quelle raison, s'est intéressé au sexe du corail rouge, et nous a conduit à une belle découverte ! Olivier, cher confrère cévenol, tu m'as fait découvrir le merveilleux monde de la bioinfo. Merci d'avoir pris le temps de m'apprendre alors que cela t'aurait pris un temps infiniment plus court de le faire à ma place (au passage, merci de m'avoir laissée ramer pendant des mois sur Perl et de m'avoir dit après que Python c'était mieux). Grâce à toi, j'ai compris une

partie de la logique informatique, même si je ne doute pas que tu rigoleras pendant encore longtemps en lisant mes scripts. N'oublie pas pour les temps difficiles : «All we have to decide is what to do with the time that is given us.» (J.R.R. Tolkien).

J'ai bénéficié au cours de ma thèse du travail et de l'investissement de beaucoup de personnes de la Station Marine d'Endoume. Je remercie donc Caroline Rocher et Marjorie Selva du SCBM pour l'organisation du laboratoire de biologie moléculaire, ainsi que Frédéric Zubérier et les plongeurs du service plongée de l'OSU pytheas pour l'échantillonnage de corail rouge sans lequel ma thèse n'aurait pas existé. Un grand merci supplémentaire à Frédéric qui a fait de magnifiques photos de corail rouge qui ont illustré cette thèse, de nombreuses présentations et posters mais aussi la couverture de *Molecular Ecology Resources*. Merci à Sandrine dont le remarquable travail nous a permis de faire le lien entre génétique et morphologie dans l'identification du sexe du corail rouge. Je remercie aussi Joëlle qui a toujours été là pour tenter de faciliter les démarches administratives avec beaucoup de patience ainsi que Michel et Kamel toujours disponibles pour régler nos petits soucis matériels. Un grand merci à Alrick et Julien pour avoir réglé le moindre de nos petits soucis informatiques avec une grande efficacité. Je remercie également les membres de l'équipe IDEA, jamais avares de conseils, suggestions et idées à la croisée de deux couloirs ou autour d'un café.

Je remercie les membres de mon comité de thèse, Pierre-Alexandre Gagnaire et Guillaume Mitta qui m'ont aidée à faire le point à mi-parcours et qui m'ont donné de précieux conseils pour avancer et trouver des solutions dans mes problèmes de préparation de bibliothèques. Merci à tous les membres de mon jury, Nicolas Bierne, Nicolas Galtier, Eve Toulza, Brigitte Crouau-Roy, Jean-Christophe Poggiale qui ont accepté de prendre du temps pour relire et évaluer mon travail. Un grand merci à Manuela Carenzi qui a pris beaucoup de son temps pour m'aider en statistiques. Je tiens aussi à remercier tous ceux qui participent à l'ANR Adacni, les réunions de travail et échanges m'ont beaucoup aidés pour ancrer ma thèse et orienter ma problématique. Un merci particulier à Cédric Mallien et Kelly Brenner pour les discussions et partages sur nos sujets de thèse respectifs très liés.

L'enseignement a accompagné ma thèse pendant ces trois années, et je tiens à remercier Didier Aurelle et Emmanuelle Renard pour m'avoir fait partager leur expérience, ainsi que Carole Borchellini pour avoir accepté une de mes surveillances d'examen, sans même que je puisse lui renvoyer l'ascenseur.

Mes remerciements vont aussi à Hélène Magalon et Baptiste Postaire qui m'ont fait confiance pendant mon stage de Master 2 et chaleureusement accueillis à la Réunion. A vos côtés, j'ai beaucoup appris sur la rigueur scientifique que ce soit en biologie moléculaire ou

dans l'analyse de données. Ce fut un plaisir de te retrouver Bautisse cette dernière année dans le bureau d'à côté, et de t'entendre crier «café!» à travers le mur.

Une thèse c'est aussi des moments de partages avec les copains doctorants. Les premiers ce sont Quentin, Alex et Laura qui m'ont accueillie dès mon premier stage à la Station Marine, qui m'ont fait partager leur expérience et leur bonne humeur. Quentin, ta présence au cours de mes premières années de thèse a été indispensable pour moi. Ce n'aurait pas été la même chose sans les pauses thé/coca, sans tes sauvetages quand mes manip ne marchaient pas. J'ai toujours pu compter sur toi et nos expériences de botanique, excursions à Truffaut et baignade de novembre vont particulièrement me manquer. Alex, tu as passé des années à essayer de m'endurcir, je crois malheureusement que ce fut un échec ! Merci pour ta sincérité qui m'a permis d'en apprendre un peu plus sur moi et sur les autres. Merci aussi à toi Laura pour ta générosité, ta spontanéité et ta bonne humeur tout au long de nos thèses, finalement on fera un grand final ensemble en décembre ! Ensuite est arrivée la seconde vague de doctorants : Aurélien et Laure. Cher Aurélien, est-ce que j'ai eu raison ou tort de t'adresser la parole en premier lieu, nul ne le saura jamais. Ce qui est sûr c'est qu'on a bien rigolé, et ça a été un plaisir de côtoyer un esprit brillant comme l'on en voit rarement. Sur ce, j'tembrasse gros tu m'as gonflé ! Laure, ton enthousiasme débordant n'est bridé que par ton emploi du temps plein à craquer. Pas toujours disponible, pas toujours ponctuelle, tu es pourtant la personne sur qui j'ai pu le plus compter ces dernier mois. Merci aux membres du groupe Happy SME (Pauline, Sandrine, Laure, Aurélien, Dorian, Marie, Abi, Bautisse, Selmane, César, Giulia) pour les discussions, les partages de bons plans, les entraides scientifiques et les moments de délire général. Finalement, merci à Abi pour nous avoir fait partager ton haut niveau scientifique et ta culture générale sans faille (et pour avoir corrigé l'anglais de mon article)... on n'aura vraiment eu aucune chance au Trivial Pursuit ! Ça a été un plaisir de partager des soirées et petits week end avec toi. Je souhaite aussi remercier tous mes collègues doctorants et post-doctorants de Saint-Charles grâce auxquels j'ai pu ouvrir mon esprit sur d'autres domaines que le marin et la génétique des populations, et qui m'ont aussi fait découvrir leur culture. Un merci tout particulier à Sandrine dont le dynamisme et l'investissement n'ont jamais eu d'égal et qui a grandement participé à mon intégration dans l'équipe.

Enfin, je remercie ma famille qui a toujours été là pour me soutenir et me supporter, et qui est à l'origine de ma passion pour le milieu marin. Les pauses cévenoles et méditerranéennes, les retours aux sources m'ont permis d'arriver au bout de ces trois années de thèse.

Avant propos

Cette thèse est une thèse sur articles qui comporte quatre articles dont deux publiés et deux en attente de soumission. Les articles présentés sont en anglais et précédé d'un résumé en français. L'introduction et la discussion générales sont en français. Les articles présentés dans le texte principal de cette thèse ont été intégrés de la façon la plus harmonieuse possible afin de faciliter la lecture. Les figures et tables de ces articles sont donc intégrés au texte, leur numérotation étant constituée du numéro du chapitre correspondant suivi du numéro de la figure ou de la table à l'intérieur de ce chapitre (exemple : la figure 2.3 est la figure 3 du Chapitre 2). Chaque chapitre présente une annexe à la fin du document contenant les figures et tables supplémentaires des articles et éventuellement des analyses qui n'ont pas été présentées dans l'article mais qui ont été réalisées. Les annexes sont dans le même ordre que les chapitres et identifiées par des lettres. La numérotation des figures et tables en annexe suit le même principe que celle du texte principal (exemple : la figure C.3 est la figure 3 de l'Annexe C, qui est l'annexe correspondant au Chapitre 3).

Chapitre 1

Chapitre 1 : Introduction

1.1 Qu'est-ce que l'adaptation locale ?

1.1.1 Origine du concept d'adaptation

Dans le milieu naturel, il est parfois intéressant de se rendre compte non seulement de la diversité des formes de vie, mais aussi de leur ajustement parfois remarquable avec leurs conditions de vie. Pour expliquer ces ajustements, on doit envisager des modifications morphologiques, physiologiques ou comportementales des organismes en relation avec leur milieu de vie, qui lui aussi évolue au cours du temps. Charles Darwin et Jean-Baptiste Lamarck ont tous deux formalisé, de manière différente, leur théorie et défini les concepts d'évolution des espèces et d'adaptation au XIX^e siècle. Dans les deux cas, l'adaptation au cours du temps des individus à leur environnement est un des concepts fondamentaux dans l'évolution des espèces.

Lamarckisme

Lamarck a posé les bases du transformisme. Selon lui, les espèces se transforment lentement au cours du temps, en gravissant peu à peu des niveaux de complexité [180]. Lamarck considère qu'un individu peut acquérir un trait avantageux, ou perdre un trait neutre au cours de sa vie, simplement par l'usage ou le non usage d'un organe (première loi de Lamarck) [180, 60]. Ce sont ces acquisitions par les individus d'une espèce et leur transmission à leur descendance (deuxième

loi de Lamarck) qui vont être le moteur de l'évolution de cette espèce. L'exemple le plus repris, et peut être le plus parlant que cite Lamarck est celui de la girafe [180]¹. Selon lui, les individus ont pu acquérir au cours de leur vie, un cou plus long en essayant d'atteindre des feuilles plus hautes, le sol étant trop aride. Le fait que tous les individus de l'espèce aient cette habitude, et que le trait acquis se transmette au moins en partie à la génération suivante explique qu'aujourd'hui toutes les girafes aient un long cou. Il y a dans la théorie de Lamarck une certaine idée de volonté de l'individu : si le cou d'une girafe s'allonge, c'est parce qu'elle s'efforce d'atteindre les feuilles hautes. Cette transmission des caractères acquis, bien qu'ayant eu moins de succès que la théorie de la sélection naturelle de Darwin, est récemment réinterprétée à la lumière de l'étude de l'hérédité non-génétique [86].

Darwinisme

Pour Darwin, le concept d'adaptation locale est indissociable de celui de sélection naturelle. Il a défini ce dernier dans *L'Origine des espèces* comme le processus qui va permettre aux variations de traits phénotypiques les plus adaptées à un environnement donné d'être conservées au cours du temps, alors que les variations délétères seront éliminées [88]. Cette sélection des variations avantageuses peut avoir lieu uniquement si des variations d'un trait sont pré-existantes chez l'espèce, si ces variations confèrent aux individus des différences dans leur succès de survie ou de reproduction dans un environnement donné, et si ces variations sont héréditaires. Dans ce cas uniquement, les individus les plus adaptés à leur environnement produiront plus de descendants jusqu'à ce que la variation avantageuse d'un trait ait envahi la population. La sélection naturelle est donc le processus de sélection des variants avantageux au cours du temps². Notons ici que Alfred Rus-

1. "Relativement aux habitudes, il est curieux d'en observer le produit dans la forme particulière et la taille de la girafe (camelo-pardalis) : on sait que cet animal, le plus grand des mammifères, habite l'intérieur de l'Afrique, et qu'il vit dans des lieux où la terre, presque toujours aride et sans herbage, l'oblige de brouter le feuillage des arbres, et de s'efforcer continuellement d'y atteindre. Il est résulté de cette habitude, soutenue, depuis longtemps, dans tous les individus de sa race, que ses jambes de devant sont devenues plus longues que celles de derrière, et que son col s'est tellement allongé, que la girafe, sans se dresser sur les jambes de derrière, élève sa tête et atteint à six mètres de hauteur (près de vingt pieds)." [180]

2. "J'ai donné le nom de sélection naturelle à cette conservation des différences et des variations individuelles favorables et à cette élimination des variations nuisibles." [88]. "This preser-

sel Wallace, découvrit le rôle de la sélection naturelle et aboutit à des conclusions presque identiques à celles de Darwin en 1858 [328, 327, 60]. Il n’y a pas dans la théorie de Darwin de notion de volonté, mais au contraire l’idée que la présence dans la population de variations avantageuses soit le fruit du hasard. Si on reprend ici l’exemple de la girafe, il y aurait pour Darwin dans la population des girafes ayant des cous plus ou moins longs (variation pré-existante). Les girafes ayant un cou plus long, peuvent accéder à des feuilles inaccessibles aux herbivores, donc être mieux nourries et produire potentiellement plus de descendants qui auront à leur tour un cou plus grand que les autres individus si ce caractère est héritable.

Darwin établit une distinction entre un trait adaptatif qui est pour lui inné et un trait acquis au cours de la vie d’un individu qui participe au fait que cet individu soit adapté à son environnement, il parle alors d’acclimatation. Ce trait acquis peut être réversible si l’environnement change. Pour Darwin, une espèce peut être adaptée à un environnement et être présente dans une plus large gamme d’environnement grâce à une acclimatation des individus. Finalement, Darwin souligne le fait qu’il peut être difficile voire impossible de séparer les effets de l’adaptation des effets de l’acclimatation dans l’adéquation que l’on observe entre une population et son environnement.

Que ce soit dans le sens de Darwin ou de Lamarck, l’adaptation peut désigner à la fois le processus au cours duquel une espèce devient adaptée à son environnement et le résultat de ce processus. Reprenons ici l’exemple de la girafe, l’adaptation désigne à la fois le processus au cours duquel le cou s’est allongé, et le résultat, c’est à dire le long cou. Afin d’éviter toute confusion, le terme “adaptation” sera utilisé dans ce manuscrit pour décrire le processus, celui de “signal d’adaptation” pour le résultat de ce processus. Nous emploierons le terme “acclimatation” pour parler d’un effet acquis à l’échelle de la vie d’un individu, réversible à plus ou moins long terme et non héritable. Toute référence à la transmission de caractère acquis sera faite explicitement.

vation of favourable variations and the rejection of injurious variations, I call Natural Selection.”

[88]

La théorie synthétique de l'évolution

A partir de la théorie de Darwin, l'adaptation sera définie comme un caractère nouveau apparu chez un organisme et maintenu par la sélection naturelle [146]. Au début du XX^e siècle, les travaux de Thomas Hunt Morgan proposent les bases de la théorie chromosomique de l'hérédité. Dans les années 1930, ces avancées sont intégrées aux approches populationnelles et donnent naissance à la génétique des populations. Les mécanismes responsables de la sélection naturelle et de l'adaptation vont être décrits et définis à la lumière de la génétique. Les variations préexistantes des traits phénotypiques deviennent les différents allèles d'un gène, l'apparition aléatoire d'un nouvel allèle étant le résultat d'une mutation. Selon cette nouvelle définition, la sélection naturelle correspond à l'augmentation dans une population de la fréquence d'un allèle avantageux sous l'action de la sélection et l'adaptation le processus au cours duquel une population devient génétiquement adaptée à son environnement (notons ici que l'adaptation locale à travers l'invasion dans une population d'un trait adaptatif peut conduire à l'extinction de cette population [159]). Au contraire, on considérera que l'acclimatation permet de produire différents phénotypes à partir du même génotype, selon l'environnement considéré, et n'a donc aucune base génétique.

1.1.2 Adaptation locale

Définition

La génétique des populations a permis d'appliquer le principe de l'adaptation à l'échelle des populations. Dans ce cas, on considère une espèce dont les populations sont distribuées dans des environnements contrastés. En l'absence d'autres forces évolutives, chaque population devrait évoluer vers des traits lui conférant un avantage dans son environnement local [166]. Ainsi, chaque génotype résident pourrait avoir, *in fine*, une capacité à transmettre ses gènes plus élevée dans son environnement d'origine que dans n'importe quel autre environnement de l'aire de répartition de cette espèce.

Il peut y avoir adaptation locale si la direction de la sélection change pour un allèle entre les habitats (effet environnemental antagoniste ou pléiotropie antagoniste), ou encore si l'intensité de la sélection sur plusieurs locus maintenus à un

certain niveau de polymorphie covarie négativement entre habitats [187, 166].

Résultats possibles de l'évolution en environnements contrastés

Lorsque les populations d'une espèce sont réparties dans des environnements contrastés, plusieurs processus peuvent conduire à une adaptation à ces environnements [166] :

- Un seul phénotype généraliste présentant un degré d'adaptation similaire dans tous les environnements.
- Un seul phénotype spécialiste adapté de manière optimale à un habitat et peu adapté aux autres habitats.
- Plusieurs phénotypes spécialistes, chacun ayant une fitness maximale dans un type d'habitat.

C'est ce dernier point qui est attendu dans un cas d'adaptation locale au sens strict. Notons ici que dans certains cas particuliers, la sélection peut agir contre l'adaptation locale. C'est le cas si une population est soumise à des conditions environnementales fluctuantes, pour lesquelles la variation temporelle de la pression de sélection peut favoriser les phénotypes généralistes [175]. Enfin, chez les espèces présentes en environnement hétérogène, l'évolution vers une **plasticité phénotypique adaptative** peut être favorisée. Dans ce cas, on s'attend à observer la présence d'un génotype qui produit dans chaque habitat un phénotype localement optimal et qui serait fixé dans toutes les populations (en absence de coûts de la plasticité) [166]. On observe dans ce cas une différenciation phénotypique entre populations présentes dans des environnements différents, mais pas de différenciation génétique pour des locus qui pourraient être concernés par l'adaptation.

1.1.3 Impact des autres forces évolutives sur l'adaptation locale

La sélection divergente (c'est à dire ayant des directions différentes selon les environnements) est la force directrice de l'adaptation locale. Cependant, le résultat observé dépend de l'interaction entre la sélection divergente et les autres forces évolutives (Fig 1.1).

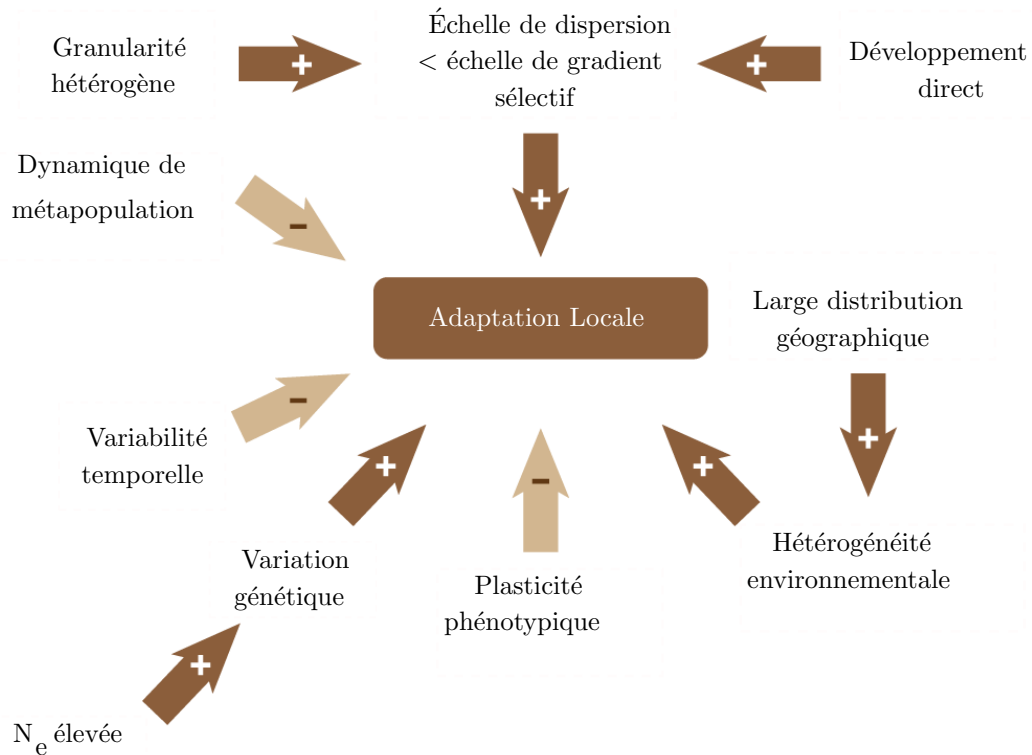


FIGURE 1.1 – Facteurs connus pour faciliter (flèches foncées) ou contraindre (flèches claires) l’adaptation locale. N_e = taille efficace de la population. Adapté de [276].

Flux de gènes

Pour estimer l’effet du flux de gènes sur l’adaptation locale, il faut tenir compte de l’interaction entre les effets positifs et négatifs de la dispersion [187, 129]. En mettant à disposition un pool de variations génétiques, le flux de gènes peut augmenter le potentiel évolutif d’une population [139, 3]. De plus, le flux de gènes dans de nouveaux environnements permet à la fois de purger des allèles mal adaptés et d’augmenter le taux auquel des allèles favorables sont partagés entre des sous-populations [181, 139]. Au contraire, l’apport dans une population d’allèles mal adaptés change les fréquences alléliques dans une direction opposée à la sélection naturelle, conduisant à une diminution de l’adaptation locale de cette population (diminution de la fitness moyenne de la population). On appelle ce phénomène le **fardeau de migration** [187, 35]. De plus, un flux de gènes élevé favorise la perte de polymorphismes favorables dans l’environnement local [166].

On peut ici indiquer deux cas particuliers pouvant influencer l’adaptation locale

à travers le flux de gènes :

- Un flux de gène asymétrique a lieu lorsque la probabilité d'échange de gènes entre populations est plus ou moins élevée selon le sens des échanges. Ce type de patron peut conduire à une structure de type source-puits. Dans ce cas, le fardeau de migration pourrait être trop élevé dans les populations puits pour que l'adaptation locale se mette en place. Au contraire, le potentiel adaptatif des populations sources pourrait être limité par l'absence d'apports de nouveaux variants génétiques [166].
- Des variations génétiques pour le choix de l'habitat peuvent provoquer un flux de gènes préférentiel entre des habitats proches environnementalement parlant, quelle que soit leur distance géographique. Ce type de choix d'habitats est très favorable à la mise en place d'adaptation locale [166].

Finalement, on considère souvent qu'un flux de gènes réduit est un facteur favorable pour la mise en place de l'adaptation locale [166, 187].

Dérive génétique

Dans une population de taille finie, on peut observer des changements de fréquences alléliques au cours du temps, simplement par échantillonnage aléatoire à chaque génération des individus qui se reproduisent. La dérive génétique peut conduire à la fois à une perte de variabilité génétique diminuant le potentiel adaptatif d'une population et à la fixation de génotypes mal adaptés [3]. Il a été décrit que dans certaines situations, un flux de gène modéré peut permettre de restaurer la variabilité génétique perdue par dérive et d'augmenter la fitness moyenne d'une population soumise à une forte dérive génétique [3]. Lorsque le produit du coefficient de sélection s et de la taille efficace N_e [342] est inférieur à 1, alors la dérive génétique va contraindre les réponses adaptatives et peut ainsi freiner ou empêcher l'évolution de l'adaptation locale [263]. Enfin, comme on le verra plus loin, la dérive génétique peut aussi gêner la détection de l'adaptation locale.

Variation polygénique

Théoriquement, le maintien du polymorphisme adaptatif entre deux populations issues d'environnements divergents est plus favorable si le locus en question a un effet fort. Ce type de locus devrait aussi présenter une plus forte différenciation

des fréquences alléliques dans une situation de sélection divergente [147]. Enfin, un allèle à effet fort est moins susceptible d'être perdu par dérive génétique [82].

Bien que ce type de locus soit le plus facile à détecter par les approches moléculaires, de nombreux caractères liés à la fitness et susceptibles de jouer un rôle dans l'adaptation locale présentent une variation polygénique, notamment dans le cas d'adaptation à des conditions environnementales complexes [346, 20, 38, 231].

1.2 Comment détecter l'adaptation locale ?

1.2.1 Études expérimentales

La présence d'adaptation locale peut être détectée en testant la relation génotype - environnement. Ceci peut être réalisé en mesurant la fitness d'un génotype donné dans son environnement d'origine A et dans un autre environnement B, et d'un génotype dans son environnement d'origine B et dans l'environnement A. Cette approche prend généralement deux formes : **transplants réciproques** ou **expériences en jardin commun** dans lesquelles les conditions environnementales étudiées sont artificiellement créées.

Kawecki & Ebert (2004) [166] insistent sur le fait qu'une différence de fitness entre son environnement local (ou sympatrique) et un environnement étranger (ou allopatrique) ne suffit pas pour conclure à de l'adaptation locale (Fig. 1.2). Deux critères ont été proposés pour identifier des cas d'adaptation locale :

- Le critère local vs étranger : les individus issus d'un environnement donné ont une fitness plus élevée dans leur environnement d'origine que des individus issus d'un autre environnement (Fig. 1.2A).
- Le critère proche vs éloigné : Les individus issus d'un environnement donné ont une fitness plus élevée dans leur environnement d'origine que dans un environnement étranger (Fig. 1.2B).

Cependant, les effets de l'acclimatation (ou effets non génétiques, incluant des effets morphologiques, développementaux, maternels, physiologiques ...) à son environnement d'origine sur la fitness d'un individu peuvent être forts, et maintenus à plus ou moins long terme [276, 263, 166]. Si c'est le cas, même après un temps long de transplantation ou en jardin commun, il peut être difficile de séparer les ef-

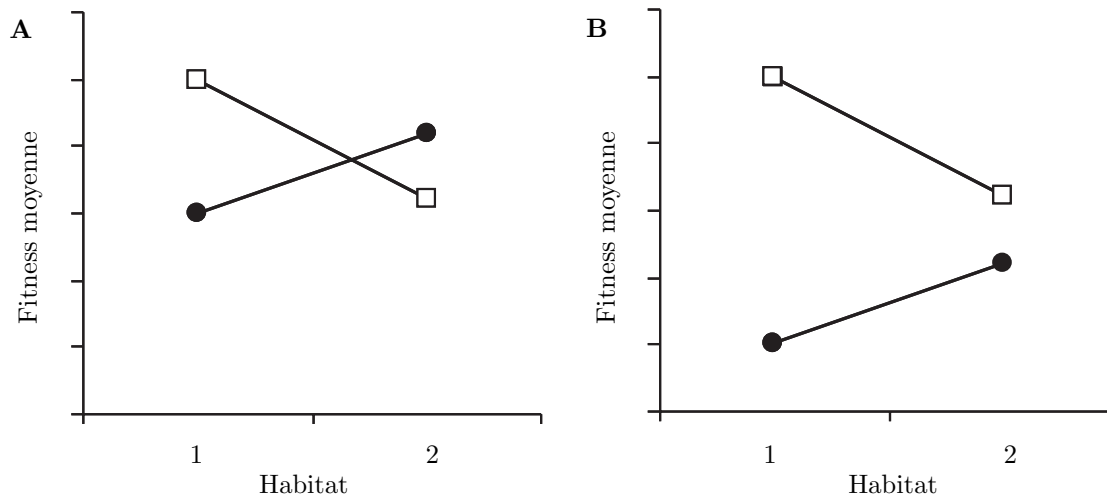


FIGURE 1.2 – Patron d’interaction population - environnement pour la fitness. **A** satisfait le critère local vs étranger. **B** satisfait le critère proche vs éloigné. Adapté de [166].

fets de l’acclimatation de ceux de l’adaptation génétique. Il peut être recommandé de mesurer la fitness après plusieurs générations en jardin commun, mais même dans ce cas, on ne peut souvent pas exclure la possibilité que des effets à long terme, incluant des effets épigénétiques transmis entre générations confondent un éventuellement effet d’adaptation locale. Ceci a été décrit en milieu marin chez des huîtres et oursins exposés à une acidification [230, 101], ainsi que chez des poissons tropicaux acclimatés à la température [273].

1.2.2 La génétique de l’adaptation locale

Le mécanisme d’acclimatation aux conditions environnementales locales complexifie la démonstration du rôle de l’adaptation locale à partir de différences phénotypiques observées dans des populations issues d’environnements contrastés. L’outil génétique permet de s’affranchir des effets de l’acclimatation pour chercher directement dans le génome des traces d’adaptation locale. Le développement des approches génétiques et génomiques sur des espèces non modèles a ainsi fortement enrichi le domaine de l’écologie moléculaire en ouvrant de nouvelles visions de l’adaptation en milieu naturel. De nombreuses méthodes ont pour point de départ

le phénotype avec pour objectif de rechercher des signaux génétiques associés. C’est le cas des études de Quantitative Trait Locus (QTL), qui peuvent être associées à des études expérimentales telles que le jardin commun ou le transplant réciproque, dans la mesure où les descendants de croisements entre populations sont inclus dans l’expérimentation [284]. D’autres méthodes visent à estimer le Q_{ST} (proportion de la variance d’un trait quantitatif donné qui peut être attribuée à des différences génétiques entre populations) à travers des expériences de croisements dans des environnements contrôlés, et à comparer cette valeur avec le F_{ST} afin de savoir si la divergence du trait étudié dans la population peut être expliquée par l’action seule de processus évolutifs neutres ou si la sélection naturelle doit être considérée (i.e. approches de comparaison F_{ST}/Q_{ST} [186, 334]). Dans certains cas, ce type d’approches peut être utile pour préciser le rôle de l’adaptation locale et de l’acclimatation.

Ici, nous traiterons uniquement les méthodes de génétique des populations. Dans ce cas, il s’agit de chercher des signaux d’adaptation locale directement au niveau de l’ADN [284, 200]. De nombreuses méthodes ont été développées pour détecter des signaux de sélection naturelle sur les séquences en cherchant des écarts au modèle neutre (par exemple le D de Tajima [307]) ou sont basées sur la comparaison entre polymorphisme et divergence (comme le test de MacDonald–Kreitman [206]), mais ne sont pas dédiées à détecter directement l’adaptation locale [284]. Ici, nous traiterons uniquement des deux principaux types de méthodes dédiées à la détection de l’adaptation locale.

Différenciation des populations à partir de la distribution de l’indice de fixation de Wright (F_{ST})

Cette approche, basée sur le travail de Lewontin & Krakauer (1973) [188] de détection de locus sous sélection, permet notamment de distinguer des forces évolutives ayant un effet sur le génome entier (dérive génétique, flux de gènes, consanguinité) des effets des forces influençant un ou des locus particulier(s) (sélection naturelle, recombinaison) [196]. Il s’agit d’utiliser un grand nombre de locus pour estimer une distribution neutre de F_{ST} , puis de tester la présence de locus ayant un comportement aberrant par rapport à cette distribution neutre (par la suite, ces locus seront appelés “locus outliers”). En effet, on s’attend à ce que tous les locus

du génome répondent de façon plus ou moins similaire à l’histoire démographique et neutre des populations (avec une variance associée) alors que seulement quelques locus vont mettre en évidence un patron de différenciation génétique ayant pour origine un événement sélectif [37]. Du fait de la nécessité d’estimer cette distribution neutre de F_{ST} , le développement de ces méthodes a bénéficié de l’émergence de la génomique des populations qui permet de génotyper aléatoirement de nombreux locus chez plusieurs individus et populations [33]. Ces développements permettent ainsi de rechercher des signaux d’adaptation locale sans à priori, y compris sur les traits impliqués dans l’adaptation ou sur les facteurs sélectifs.

La principale difficulté liée à l’identification d’outliers par rapport à une distribution neutre réside dans le fait que les processus neutres affectent chaque locus différemment, et que cette distribution neutre est extrêmement dépendante de la structure et de la démographie des populations étudiées. Ainsi, sans que des effets sélectifs soient impliqués, des locus outliers peuvent être détectés simplement du fait de l’importance de la variance des F_{ST} [31, 149]. Ceci est d’autant plus vrai quand le niveau moyen de différenciation est élevé, la variance des F_{ST} augmentant avec le F_{ST} moyen [149]. Ce problème de détection de locus outliers qui ne sont pas sous sélection (erreur de type I), était particulièrement important dans le cas du test original de Lewontin & Krakauer (1973) [188]. En effet, cette méthode est basée sur une structure spatiale des populations suivant un modèle en île de Wright, qui n’est pas appropriée dans certains cas de populations réelles dont la structure s’écarte de ce modèle, la corrélation entre les fréquences alléliques des sous-populations conduisant à l’augmentation de la variance neutre des F_{ST} [31]. De nouvelles méthodes statistiques tenant compte de ce problème existent aujourd’hui et peuvent être appliquées dans plusieurs cas de structures spatiales des populations en limitant ce nombre de faux positifs [31, 21, 116]. Ces méthodes diffèrent par le modèle démographique sous-jacent (modèle en îles ou hiérarchique), l’approche statistique qu’elles adoptent et si la sélection est explicitement incluse ou non [284, 212].

Il est important de garder à l’esprit que des locus peuvent être détectés comme outliers par ce type de méthodes sans être pour autant des marqueurs de l’adaptation locale :

- Le locus détecté, sans être soumis lui même à l’adaptation locale, peut être

lié à un locus qui l'est.

- Bien que les nouvelles méthodes de détection de locus outliers prennent en compte des structures spatiales des populations s'écartant du modèle en îles de Wright, l'importance de la variance des valeurs de F_{ST} pose toujours problème dans certaines situations précises de structure génétique ou de démographie (forte structure génétique moyenne, structure spatiale hiérarchique ou auto-corrélée ; [31, 149]). Certaines méthodes ont été développées pour traiter ce type de données [109], mais ne résolvent pas complètement le problème des faux positifs.
- Des mécanismes autres que l'adaptation locale peuvent conduire à des F_{ST} élevés : sélection contre des mutations délétères, balayages sélectifs à l'échelle de l'espèce, zones hybrides cryptiques, effets aléatoires dans l'expansion des populations [32].

Corrélation entre fréquences alléliques et variables environnementales

Dans cette seconde catégorie de méthodes, on ne cherche plus à identifier des locus présentant une différenciation anormalement élevée entre populations, mais des locus présentant des fréquences alléliques corrélées avec une ou plusieurs variables environnementales. Dans ce cas, ces locus peuvent être sous sélection dirigée par ces variables environnementales, ou des variables environnementales qui leur sont corrélées [142, 199]. Les méthodes de corrélation entre fréquences alléliques et variables environnementales reposent sur une estimation correcte des fréquences alléliques dans les populations. Il est donc indispensable de tenir compte de la structure génétique neutre afin de limiter le nombre de faux positifs et faux négatifs [73, 246, 140, 122, 261]. Dans tous les cas, ces méthodes, tout comme celles basées sur les locus outliers, sont connues pour conduire à un taux élevé de faux positifs particulièrement dans certains cas précis présentant une structure génétique hiérarchique, une expansion d'aire de répartition ou une sélection polygénique par exemple [195, 91, 92, 123, 212, 261]. Ils peuvent également être trompeurs dans les cas où un gradient environnemental se superpose à une zone de contact entre deux entités génétiquement différenciées [32].

1.3 Exemples en milieu marin

Les études d'adaptation locale se sont longtemps focalisées sur les environnements terrestres, dont les gradients environnementaux forts et persistants propices à sa mise en place sont bien caractérisés. Pourtant, des gradients parfois forts de température, pH, salinité, nutriments et autres caractéristiques physiques du milieu existent en milieu marin parfois de façon persistante, mais pouvant aussi être soumis à des fluctuations importantes dans certains cas [276]. L'existence de tels gradients en milieu marin, souvent à petite échelle et associée à la dominance d'espèces présentant une phase dispersive plus ou moins importante en fait un milieu particulièrement intéressant pour l'étude de l'adaptation locale.

1.3.1 Études expérimentales

Les coraux constituent un groupe polyphylétique d'espèces de cnidaires caractérisées par un stade de vie adulte sessile et présentant un squelette calcaire dur. Ce mode d'organisation se retrouve chez les hexacoralliaires, les octocoralliaires et les hydrozoaires. Ce sont des organismes immobiles et longévifs qui ne peuvent généralement pas migrer vers leur optimum environnemental après leur installation larvaire. Leur présence dans des milieux où ils expérimentent des stress environnementaux persistants plus ou moins forts en font donc des cibles privilégiées et des modèles d'intérêts pour l'étude de l'adaptation locale en milieu marin.

La plupart de ces espèces, supports de biodiversité et souvent à la base des ressources marines côtières, sont d'un grand intérêt à la fois économique et écologique. Le cycle de vie de nombreux coraux tropicaux peu profonds est basé sur leur endosymbiose avec un dinoflagellé photosynthétique du genre *Symbiodinium*. Chez ces organismes, la première réponse à un stress thermique est le blanchissement, qui se traduit par une perte du symbionte [335], et peut conduire à une augmentation du taux de mortalité du corail. Certains coraux ne présentent pas de symbiose avec une espèce du genre *Symbiodinium* et sont aussi sensibles au stress thermique qui se traduit alors par une nécrose tissulaire (c'est le cas de plusieurs espèces d'octocoralliaires de Méditerranée). Du fait de cette sensibilité au stress thermique, mais aussi de l'acidification de leur milieu et des conséquences de leur

proximité avec des densités de populations élevées, les événements de blanchissement et de nécrose chez les coraux augmentent à la fois en fréquence et en intensité ces dernières années à l'échelle mondiale, conduisant à une dégradation progressive des écosystèmes coralliens et compromettant leur évolution future dans un scénario de changement climatique [157, 151]. Ainsi, les études tentant de comprendre les capacités de survie future des coraux dans ce contexte de changement climatique à travers l'étude de l'adaptation et de l'acclimatation locale via des expérimentations de transplants réciproques et de jardin commun se sont multipliées ces dernières années. Dans ce type d'études, l'enjeu de conservation est double : les capacités d'adaptation et/ou d'acclimatation des populations sont-elles suffisantes pour garantir leur maintien au cours du changement climatique ? Peut-on envisager de restaurer certaines populations par transplants d'individus issus d'autres populations ?

Dans la plupart des études, des coraux de différentes origines (surface vs profond, côte vs large) sont transplantés réciproquement [247, 106, 47, 294, 227], ou placés en jardin commun [102, 17, 223, 106] où ils sont exposés à un stress (thermique ou irradiance), différents paramètres de survie, croissance, santé, teneur en symbiontes étant mesurés au cours du temps afin d'évaluer le degré de spécialisation des individus à leur environnement d'origine (Table 1.1). L'utilisation de coraux dans ce type d'études ajoute la possibilité de transplanter dans différents environnements des clones issus de la fragmentation d'une même colonie, et d'explorer la plasticité phénotypique d'un génotype donné. On contrôle ainsi les effets possibles de la variabilité génétique entre individus. Parmi ces études, des génotypes plutôt généralistes ont été observés chez certaines espèces, comme chez *Porites lobata* dont le taux de croissance du squelette ainsi que le taux de calcification sont plus élevés, et la densité est plus faible chez les individus transplantés sur le récif du côté lagon par rapport aux individus situés sur le récif côté océan, quelle que soit l'origine des individus [294]. Chez d'autres espèces, des effets fixés de l'environnement d'origine des individus ont été mis en évidence. C'est le cas chez *Pocillopora damicornis* chez qui des différences de thermotolérance importantes ont été mises en évidence entre des individus originaires ou non d'une zone de d'upwelling [102]. Ce type d'étude ne permet pas de savoir si cette spécialisation apparente des individus à leur environnement d'origine est le résultat d'adapta-

tion locale ou de maintien à moyen terme de différences d’acclimatation. Notons ici que la spécialisation relative mise en évidence par ce type d’études de transplantations croisées et de jardin commun dépend à la fois du schéma expérimental et du trait phénotypique mesuré, et ne peut pas être généralisée. Par exemple, bien que certains génotypes généralistes de *Porites lobata* aient été mis en évidence dans une étude présentée ici, une étude similaire chez cette espèce a montré une surexpression fixée d’une protéine chez les individus originaires du récif côté lagon par rapport à ceux venant de l’extérieur de lagon, quel que soit leur lieu de transplant [18].

Finalement, il semble que l’acclimatation joue un rôle chez la majorité des espèces étudiées dans l’augmentation de la thermotolérance des individus au cours des stress thermiques rencontrés, avec des cas d’acclimatation d’une partie de colonie devenant plus thermotolérante que la partie non acclimatée [46] (Table 1.1). Pourtant, la réponse de ces organismes à un stress thermique ne peut pas être généralisée comme cela a été montré chez des espèces de gorgones méditerranéennes étudiées sur des sites et profondeurs similaires [242, 239] : chez *Eunicella cavolini* les individus les moins profonds semblent plus thermotolérants que les individus profonds [242], le patron inverse a été observé chez *Eunicella singularis* [239], un effet potentiellement lié à la présence (*E. singularis*) ou non (*E. cavolini*) de *Symbiodinium*. De plus, la mise en évidence de différences de thermotolérances entre adultes et juvéniles ainsi qu’entre mâles et femelles chez *Paramuricea clavata* pourrait indiquer une importante complexité intra-populationnelle dans la réponse au stress thermique [9].

Ces études expérimentales semblent indiquer que les coraux présentent bien, dans la grande majorité des cas, des différences phénotypiques de thermotolérance selon les conditions thermiques de leur environnement d’origine [336]. Dans certains cas la vie dans des environnements différents semble pouvoir correspondre à des lignées génétiques bien différenciées comme cela a été montré pour les gorgones *Eunicea* en milieu tropical [249]. Pourtant dans la plupart des cas, les bases de ces différences de thermotolérance sont encore toujours mal comprises. Peu d’études sont dédiées à estimer la part de la thermolérance qui est due à de l’adaptation génétique et celle due à l’acclimatation individuelle [18, 227], et aucune n’étudie la thermotolérance des générations suivant la transplantation afin de supprimer les

effets d'acclimatation au cours de la vie de l'individu avant sa période de transplant [166, 276].

TABLE 1.1 – Études de recherche de signaux d'adaptation locale chez des espèces de coraux via des méthodes expérimentales.

Référence	Espèce	Expérimentation	Résultat
Potts <i>et al.</i> (1984) [247]	<i>Acropora palifera</i> <i>Acropora cuneata</i>	Transplants Clones	Dans chaque environnement les phénotypes les plus vulnérables sont éliminés par sélection.
Oliver <i>et al.</i> (2011) [223]	<i>Acropora hyacinthus</i>	Stress thermique en aquarium Environnement variable vs Environnement stable	Exposition passée à un environnement variable augmente la thermotolérance.
Barshis <i>et al.</i> (2013) [17]	<i>Acropora hyacinthus</i>	Stress thermique en aquarium Thermosensibles vs Thermotolérants	L'expression de milliers de gènes change durant le stress chez les thermosensibles et les thermorésistants. Chez les thermorésistants, certains gènes ont une expression élevée à l'état basal et pourraient participer au maintien d'une résilience physiologique durant les stress environnementaux.
Palumbi <i>et al.</i> (2014) [227]	<i>Acropora hyacinthus</i>	Transplants réciproques Très variables (HV) vs modérément variable (MV)	Les individus issus d'un environnement MV transplantés dans un environnement HV augmentent leur thermotolérance sans atteindre celles des individus natifs. Au cours de la transplantation, certains gènes a effet d'acclimatation fort ont une expression qui change complètement, d'autres à effet fixé maintiennent leur expression d'origine.
Ayre <i>et al.</i> (1995) [13]	<i>Actinia tenebrosa</i>	Transplants réciproques à grande et courte distance	Différences de fitness plus importantes quand la distance de transplant est plus grande.
Sherman <i>et al.</i> (2008) [288]	<i>Actinia tenebrosa</i>	Transplants réciproques	Meilleures performances des individus dans leur environnement d'origine que dans un environnement étranger.
Kersting <i>et al.</i> (2013) [172]	<i>Cladocora caespitosa</i>	Stress thermique en présence ou pas d'algue invasive Populations ayant des histoires thermiques et de nécroses contrastées	Pas de différences de survie.

Haguenauer <i>et al.</i> (2013) [143]	<i>Corallium rubrum</i>	Stress thermique en aquarium Surface vs Profondeur	Surexpression de HSP70 par les individus de surface.
Ledoux <i>et al.</i> (2015) [183]	<i>Corallium rubrum</i>	Transplants réciproques Jardin Commun	Transplants réciproques : fitness plus élevée dans son environnement d'origine dans un cas sur deux.
Pratlong <i>et al.</i> (2015) [250]	<i>Corallium rubrum</i>	Surface vs Profondeur Jardin Commun	Jardin commun : survie plus élevée des individus de surface.
Pivotto <i>et al.</i> (2015) [242]	<i>Eunicella cavolini</i>	Surface vs Profondeur	Différences d'expression entre les deux profondeurs en situation basale.
Bruno <i>et al.</i> (1997) [47]	<i>Madracis mirabilis</i>	Stress thermique en aquarium Surface vs Profondeur	Les individus de surface sont plus thermotolérants que les individus profonds.
Walsh <i>et al.</i> (1981) [329]	<i>Metridium senile</i>	Transplants dans différents environnements Clones Jardin commun	Plasticité phénotypique élevée.
		Individus issus de latitudes différentes	Différences physiologiques selon la latitude d'origine.
Ulstrup <i>et al.</i> (2003) [322]	<i>Pocillopora damicornis</i> <i>Turbinaria reniformis</i>	Stress thermique en aquarium Individus le long d'un gradient latitudinal	La sensibilité au blanchissement différent selon le clade du <i>Symbiodinium</i> . Chez <i>P. damicornis</i> , la sensibilité au blanchissement est corrélée à la latitude.
D'Croz <i>et al.</i> (2004) [102]	<i>Pocillopora damicornis</i>	Stress thermique en aquarium génotype issus de zones de upwelling et de zones hors upwelling	Les génotypes issus des zones de upwelling sont plus sensibles au stress thermique.
Ziegler <i>et al.</i> (2014) [348]	<i>Pocillopora verrucosa</i>	Transplants réciproques Surface vs profondeur	Potentiel limité d'acclimatation hétérotrophe des colonies de surface.
Edmunds <i>et al.</i> (1994) [106]	<i>Porites porites</i>	Stress thermique en aquarium Génotypes différents issus du même environnement	Certains génotypes sont plus sensibles que d'autres au stress thermique ce qui conduit à une variabilité intra-spécifique du blanchissement.

Smith <i>et al.</i> (2007) [294]	<i>Porites lobata</i>	Transplants réciproques Avant récif vs Arrière récif	Changements des caractéristiques squelettiques par plasticité. Certaines caractéristiques restent influencées par la population source.
Barshis <i>et al.</i> (2010) [18]	<i>Porites lobata</i>	Transplant réciproque Fluctuant vs Stable	Effet fixé de l'expression d'un conjugué de l'ubiquitine, un marqueur de la réponse physiologique au stress.
Bongaerts <i>et al.</i> (2011) [36]	<i>Seriatopora hystrix</i>	Transplants réciproques	Taux de survie plus élevé dans l'environnement d'origine. Les assemblages hôte-symbiontes sont restés stables pendant l'expérimentation.
Mayfield <i>et al.</i> (2012) [205]	<i>Seriatopora hystrix</i>	Stress thermique en aquarium génotype issus de zones de upwelling et de zones hors upwelling	Les individus exposés à un régime de températures variable augmentent leur capacité de photosynthèse.

1.3.2 Approches de génomique des populations

L'étude de l'adaptation locale en milieu marin est largement dominée par les études expérimentales. Pourtant, les études de génétique des populations offrent de grandes possibilités d'exploration de l'adaptation locale, à la fois à travers des approches de gènes candidats mais aussi de balayage du génome, et permettent de mieux comprendre les mécanismes impliqués. Les premières études sont basées sur des approches de gènes candidats pour l'adaptation à des conditions environnementales particulières. Il s'agit dans ce cas de rechercher une relation entre les fréquences alléliques du gène en question et une variable environnementale donnée. Ce type d'études associent souvent des données de génétique des populations avec des expériences en aquarium de recherche de relation entre un génotype et l'activité d'une enzyme et/ou des variations de fitness. C'est le cas du travail de Hilbish & Koehn (1985) [148] sur *Mytilus edulis* qui a mis en évidence un cline dans les fréquences alléliques de l'enzyme *Lap1* (leucine amino peptidase), impliquée dans l'osmorégulation) corrélé à un gradient de salinité le long d'une côte aux États-Unis. Dans ce cas, il ne s'agit non pas d'une adaptation locale au sens strict au sein des populations considérées, mais d'une contre-sélection de certains génotypes après l'installation des juvéniles dans certaines populations. Ces mortalités de juvéniles ont pu être associées à des différences des capacités d'osmorégulation associées à l'activité de l'enzyme [148]. Chez le copépode *Trigriopus californicus*, une corrélation entre un allèle du gène GPT et la salinité a été démontrée entre des populations expérimentant des niveaux de salinité contrastés [50]. La présence de cet allèle est associée à une augmentation de l'activité de l'enzyme et de la fitness en aquarium [51]. Power & Schulte (1998) [248] ont démontré l'action de la sélection par la température sur le gène *LDH* (lactate déshydrogénase) chez *Fundulus heteroclitus*, et des différences d'activités des différents allèles ont été mises en évidence. Enfin, Planes & Romans (2004) [243] ont montré que les individus de *Diplodus sargus* possédant un allèle du gène *PGM* (phosphoglutamatase) présentaient des taux de croissance plus faibles que les autres individus. La diminution de la fréquence de cet allèle dans des cohortes au cours du temps en environnement naturel a pu être associée à une prédation plus forte sur les petits individus.

Ces approches de gènes candidats recherchent des associations entre les fréquences

alléliques d'un gène dont on connaît en général bien la fonction et un gradient environnemental. Le développement plus récent des approches de type balayage du génome permet de rechercher des candidats sans à priori sur leur identité ou leur fonction. Cependant l'interprétation fonctionnelle reste, dans ce cas, très difficile. Récemment, diverses études d'adaptation locale basées sur des analyses de SNPs issus de séquençage haut débit ont été conduites en milieu marin (Table 1.2). C'est par exemple le cas chez le hareng *Clupea harengus*, espèce très mobile à larve planctonique pour laquelle des candidats à l'adaptation à la salinité et à la température ont été mis en évidence [141], ou chez le crabe vert *Carcinus maenas*, espèce invasive à capacités de dispersion élevées pour laquelle des signaux d'adaptation à la température ont été observés [312]. Finalement, les études présentées ici qui ont identifié des locus outliers en milieu marin, concernent généralement des espèces à structure génétique faible ou absente sur les grandes échelles géographiques étudiées (plusieurs centaines de kilomètres; Table 1.2). Chez ce type d'espèces à capacités de dispersion élevées, même si le flux de gène est largement suffisant pour maintenir une homogénéité génétique entre deux populations pour des marqueurs nucléaires neutres, ce n'est pas forcément le cas pour des marqueurs sous sélection, une différenciation pour des locus sous sélection pouvant se mettre en place entre des populations apparemment non structurées génétiquement sur des locus neutres [276]. Ici, les locus outliers détectés sont les locus pour lesquels la force de la sélection a été suffisante pour contre-balancer l'effet d'homogénéisation de la migration, donc des locus potentiellement à effet forts [263]. De plus, lorsque la majorité du génome présente des niveaux de différenciation très faibles, des locus très différenciés sont plus facilement détectés, et avec plus de confiance par les méthodes habituelles de recherche de locus sous sélection. Dans le cas d'espèces à forte structure génétique au contraire, la présence de locus neutres présentant des niveaux de différenciation élevés contribuent à augmenter la difficulté de détection d'outliers résultats de processus adaptatifs. L'augmentation de la variance des valeurs de F_{ST} avec leur valeur moyenne contribue à augmenter le nombre de faux positifs détectés et donc l'identification des locus sous sélection, dans les cas de structure génétique élevée (Paragraphe 1.2.2, [149]).

TABLE 1.2 – Études de recherche de signaux d’adaptation locale en milieu marin via des méthodes non expérimentales. La colonne structure indique la distance géographique séparant les populations présentant le F_{ST} le plus faible observé.

Référence	Espèce	Expérimentation	Variable	Structure
Bay & Palumbi (2014) [20]	<i>Acropora hyacinthus</i>	Génotypage de SNPs	Température	Pas de structure significative
Tepolt <i>et al.</i> (2015) [312]	<i>Carcinus maenas</i>	SNPs issus de séquençage de transcriptomes	Température	$F_{ST} = 0,015^*$ > 4 000 km
Gleason <i>et al.</i> (2016)[137]	<i>Chlorostoma funebris</i>	RAD-Séquençage	Conditions environnementales locales	$F_{ST} = 0,0042$ en moyenne
Limborg <i>et al.</i> (2012) [193]	<i>Clupea harengus</i>	SNPs issus de séquençage de transcriptomes	Température, salinité	Pas de structure significative 1 300 km
Guo <i>et al.</i> (2016) [141]	<i>Clupea harengus</i>	RAD-Séquençage	Température, salinité	-
Bernardi <i>et al.</i> (2016) [27]	<i>Fistularia commersonii</i>	RAD-Séquençage	Salinité	Pas de structure significative 2 000 km
Bradbury <i>et al.</i> (2010) [40]	<i>Gadus morhua</i>	SNPs issus de séquençage de transcriptomes	Température	-
Berg <i>et al.</i> (2015) [24]	<i>Gadus morhua</i>	Génotypage par SNP-chip	Salinité	$F_{ST} = 0,00962^*$ 400 km
De Wit <i>et al.</i> (2013) [93]	<i>Haliotis rufescens</i>	SNPs issus de séquençage de transcriptomes	Géographie	Pas de structure significative 600 km
Benestan <i>et al.</i> (2016) [22]	<i>Homarus americanus</i>	RAD-Séquençage	Température de surface	$F_{ST} = 0,0018$ en moyenne
Wang <i>et al.</i> (2016) [331]	<i>Lateolabrax maculatus</i>	RAD-Séquençage	Conditions environnementales locales	$F_{ST} = 0,0347^*$ 2 500 km
Milano <i>et al.</i> (2014) [211]	<i>Merluccius merluccius</i>	Génotypage de SNPs	Température de surface, salinité	$F_{ST} = 0,038^*$ 1 700 km
Araneda <i>et al.</i> (2016) [7]	<i>Mytilus chilensis</i>	RAD-Séquençage	Conditions extrêmes vs intérieur de baie vs extérieur de baie	$F_{ST} = 0,012^*$ 1 000 km
Chu <i>et al.</i> (2014) [66]	<i>Nucella lapillus</i>	RAD-Séquençage	Température	$F_{ST} = 0,0254$ 300 km

1.4 Le corail rouge *Corallium rubrum*

1.4.1 Description de l'espèce

Le corail rouge est une espèce de cnidaire octocoralliaire de la famille des Coralliidae (Fig. 1.3). Il appartient à la classe des anthozoaires, dont la monophylie a été débattue mais récemment clairement établie ([61, 347], Annexe A). Il s'agit d'une espèce benthique suspensivore passive [320]. Le corail rouge est une espèce longévive (des individus de plusieurs centaines d'années ont été observés), à croissance lente (entre 0,91 et 1,57 mm en diamètre par an estimés par les méthodes pétrographiques [277, 130, 278] et environ 0,35 mm par an par la méthode de Marshall *et al.* (2014) [202]). Le corail rouge atteint sa maturité sexuelle entre 2 et 7/10 ans (selon les méthodes d'estimation utilisées et selon les populations considérées [315, 279]).



FIGURE 1.3 – Photographie d'une colonie de corail rouge (*Corallium rubrum*).
Crédit photographie : Frédéric Zuberer.

Le cycle de vie du corail rouge a été décrit précisément par Lacaze-Duthiers en 1864 [90]. Bien que Lacaze-Duthiers (1864) parle d'individus hermaphrodites à l'échelle de la colonie et du polype, aucun individu hermaphrodite n'a été décrit depuis (Fig. 1.4). Il s'agit donc d'une espèce gonochorique incubante dont le développement des gonades culmine en une ponte annuelle en été [319]. La fécondation

est interne, les larves sont incubées puis relâchées à la fin de l'été [281, 333]. La durée de vie larvaire, seule étape de dispersion dans le cycle de vie de cette espèce, a été décrite comme très courte en aquarium (entre 4 et 12 jours [333]), mais selon une étude récente, les larves de corail rouge seraient caractérisées par une durée de vie libre (toujours en aquarium) de 16 à 42 jours, avec des périodes de nage active [204].

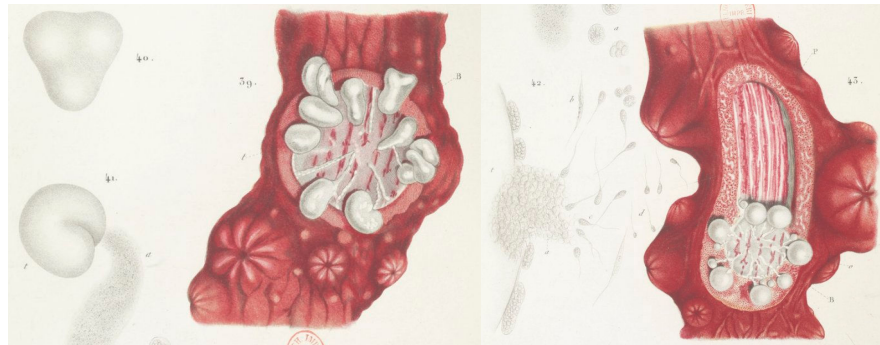


FIGURE 1.4 – Planche illustrant les gonades mâles et femelles chez le corail rouge (*Corallium rubrum*). Gauche : polype mâle ouvert présentant les gonades mâles. Droite : polype femelle ouvert présentant des œufs à divers états de développement. Issue de [90].

Le corail rouge a été observé entre 5 m et 800 m de profondeur dans le bassin méditerranéen et les côtes adjacentes d'Atlantique, où il a été récemment caractérisé génétiquement [79, 34] (Fig. 1.5). Utilisé par l'homme depuis l'antiquité, le corail rouge est encore aujourd'hui une espèce emblématique de la Méditerranée. Son exploitation pour l'industrie de la bijouterie depuis des décennies a fortement contribué à des changements démographiques dramatiques dans certaines régions [194, 321, 320, 270, 132].

En dehors de l'exploitation, le corail rouge est potentiellement menacé par les effets du changement climatique : réchauffement et acidification [42]. Finalement, l'impact du gastropode *Pseudosimnia carnea*, prédateur exclusif du corail rouge, a récemment été caractérisé dans des populations profondes (entre 85 m et 90 m) ([252], Fig. 1.6), bien que cette relation entre *Pseudosimnia carnea* et le corail rouge ne semble pas être récente [81]. Ce prédateur affecte trois fois plus les femelles que les mâles et conduit à une réduction de la fécondité pouvant atteindre 81 % sans que d'effets à la long terme sur la survie des populations aient été mis en évidence



FIGURE 1.5 – Aire de répartition du corail rouge. Données FAO (Organisation des Nations Unies pour l'alimentation et l'agriculture).

[252].



FIGURE 1.6 – Photographie de la prédation par *Pseudosimnia carnea* sur une colonie de corail rouge. Une zone sphérique broutée où les polypes sont absents est visible sur la branche en bas en droite. Issue de [252].

1.4.2 La génétique du corail rouge

La structure génétique des populations de corail rouge a été analysée à plusieurs échelles, d'abord par Abbiati *et al.* (1993) [2] qui ont mis en évidence une structure génétique entre des populations séparées par seulement 10 m en mer tyrrhénienne grâce à l'étude d'allozymes, puis par Del Gaudio *et al.* (2004) [94] et l'utilisation de fragments AFLP (Table 1.3). Dans la recherche de nouveaux marqueurs génétiques qui permettraient de caractériser plus finement la génétique des populations de corail rouge, les marqueurs microsatellites ont par la suite été

largement privilégiés aux marqueurs mitochondriaux, dont la variabilité chez les gorgonaires (et plus généralement chez les anthozoaires [287]) ne permettait d'identifier assez précisément ni la structure des populations ni la connectivité entre les populations [52]. Cependant l'identification récente d'une potentielle région de contrôle a permis de mettre en évidence un peu de polymorphisme mitochondrial et de l'utiliser en phylogéographie [74, 34]. Le séquençage d'un intron nucléaire a confirmé la structure génétique marquée de cette espèce mais l'analyse n'a pas montré de rupture profonde de type phylogéographique. L'ensemble des études de génétique des populations conduites chez le corail rouge s'accordent donc sur la structuration génétique très élevée observée chez cette espèce, quelle que soit l'échelle considérée (Table 1.3, [185, 184, 2, 94, 77, 76, 12]). De tels niveaux de structure génétique suggèrent des capacités de dispersions extrêmement réduites chez le corail rouge, en contradiction avec les durées de vie et comportements larvaires récemment étudiés chez cette espèce (de 16 à 42 jours, Paragraphe 1.4.1, [204]). La réponse se trouve peut-être dans la distribution des distances de dispersion avec une valeur moyenne très faible et la possibilité d'événements plus rares de dispersion à longue distance. Dans le nord-ouest méditerranéen, quatre groupes génétiques cohérents avec la géographie ont été identifiés [185] (Fig. 1.7) : bassin liguro-provençal, Corse, Catalogne et Îles Baléares (à l'exception de la population Corse de Porto se regroupant avec le cluster Catalan). Récemment, des études menées sur des populations mésophotiques ont confirmé le flux de gènes limité à petite et grande échelle spatiale observés chez des populations moins profondes [74, 75, 53], et suggèrent une diminution de la variabilité génétique avec la profondeur chez cette espèce [79] (résultat qui n'est pas retrouvé par Cannas *et al.* (2015) [53]). Finalement, la caractérisation génétique récente d'individus d'Atlantique suggère une distinction nette entre un cluster génétique méditerranéen et un cluster atlantique, avec la possibilité d'une introgression de l'Atlantique vers la Méditerranée [34].

1.4.3 Thermotolérance chez le corail rouge

Du fait de sa large aire de répartition verticale, les individus de corail rouge vivant à différentes profondeurs sont soumis à des conditions environnementales extrêmement contrastées. Les populations les plus proches de la surface, situées

TABLE 1.3 – Études de génétique des populations sur le corail rouge (*Corallium rubrum*). ¹ ITS1 (*internal transcribed spacer one*) est une région non codante de l'ADN ribosomique nucléaire. ² *mtMSH* est un gène mitochondrial. ³ EF1 est un intron du gène *EF1* (*elongation factor 1*). ⁴ mtC correspond à la région de contrôle mitochondriale.

Reference	Région géographique	Marqueurs	Profondeurs (m)
Abbiati <i>et al.</i> (1993) [2]	Mer tyrrhénienne	Allozymes	-
Del Gaudio <i>et al.</i> (2004) [94]	Espagne et Italie	Fragments AFLP	-
Costantini <i>et al.</i> (2007) [76]	Mer de Ligurie	Microsatellites	25
Costantini <i>et al.</i> (2007) [77]	Nord-Ouest méditerranéen	Microsatellites & ITS1 ¹	25 à 30
Costantini <i>et al.</i> (2010) [80]	Malte et Linosa	Microsatellites, <i>mtMSH</i> ² & ITS1 ¹	607 à 819
Ledoux <i>et al.</i> (2010) [184]	Marseille	Microsatellites	15
Ledoux <i>et al.</i> (2010) [185]	Nord-Ouest méditerranéen	Microsatellites	14 à 60
Aurelle <i>et al.</i> (2011) [12]	Nord-Ouest méditerranéen & Mer adriatique	Microsatellites & EF1 ³	14 à 60
Costantini <i>et al.</i> (2011) [79]	Nord-Ouest méditerranéen	Microsatellites	20 à 70
Costantini <i>et al.</i> (2013) [75]	Mer tyrrhénienne	Microsatellites, <i>mtMSH</i> ² & mtC ⁴	58 à 118
Cannas <i>et al.</i> (2015) [53]	Sardaigne	Microsatellites	87 à 107
Costantini <i>et al.</i> (2015) [74]	Mer de Ligurie & Mer Ionienne	Microsatellites & mtC ⁴	55 à 120
Boavida <i>et al.</i> (2016) [34]	Portugal	mtC ⁴	60 à 100

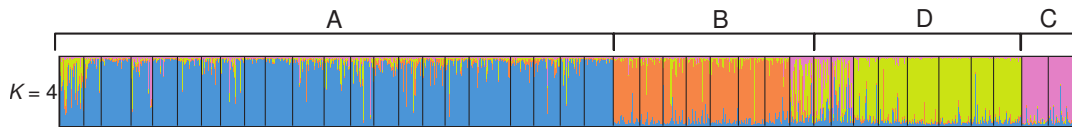


FIGURE 1.7 – Structure des populations de corail rouge (*Corallium rubrum*) en Méditerranée occidentale estimée par le logiciel STRUCTURE (adapté de [185]). A. Bassin liguro-provençal. B. Nord-Ouest Corse. C. Îles Baléares. D. Corse.

au dessus de la thermocline saisonnière sont soumises à des maximum thermiques élevés ainsi qu'à des fluctuations de températures intenses et fréquentes en été, suivant les conditions atmosphériques (température et intensité du vent notamment). Plus la profondeur augmente, plus les conditions environnementales et notamment thermiques sont tamponnées, les populations les plus profondes expérimentant des températures stables tout au long de l'année. En hiver, l'homogénéisation de la colonne d'eau conduit à diminuer les différences environnementales entre les populations des différentes profondeurs (température autour de 13 °C).

Dans la région de Marseille, la thermocline saisonnière se situe autour de 20/30 m de profondeur. Lors des périodes d'absence de vent en été, la stabilité hydrostatique conduit à une augmentation de la température de surface, qui peut atteindre des valeurs extrêmement élevées (jusqu'à 26,9 °C enregistré en août 2012 sur le site de Figuier 5 m dans la région de Marseille). Lors des épisodes de Mistral estivaux (vent de nord-ouest), la mise en place d'un phénomène d'upwelling conduit à des variations de température pouvant atteindre 7 °C en quelques heures seulement [144, 316].

La thermotolérance chez le corail rouge n'a commencé à intéresser la communauté scientifique que relativement récemment, après l'observation d'épisodes de mortalités massives chez plusieurs espèces benthiques dont le corail rouge dans le bassin ouest-méditerranéen durant les étés 1999 et 2003 [133, 233, 131]. Bien que les raisons de ces mortalités ne soient pas complètement comprises, ces événements ont eu lieu après une période de stabilité hydrostatique associée à des anomalies thermiques positives sur plusieurs semaines. De plus, la diminution de l'impact des mortalités avec la profondeur, suivant l'intensité de l'anomalie thermique confirme le rôle de la température dans ces événements [316]. Une grande variabilité a été observée dans l'intensité des mortalités chez le corail rouge à l'échelle individuelle,

inter et intra-populationnelle [233, 131]. Chez le corail rouge en particulier et les gorgonaires en général, ces mortalités se traduisent par une nécrose partielle ou totale des tissus pouvant aller jusqu'à la mort de la colonie (Fig. 1.8). Aujourd'hui encore, les facteurs déclencheurs de cette nécrose sont inconnus. Les études du stress thermique chez le corail rouge mentionnent l'éventualité d'un dysfonctionnement métabolique ou d'une ré-allocation énergétique suite à l'action directe de la température [316, 143]. Chez une autre espèce de gorgone méditerranéenne (*Paramuricea clavata*), une bactérie pathogène thermodépendante (*Vibrio coralliilyticus*) a été identifiée sur des individus nécrosés [15]. Depuis, cette bactérie a été identifiée chez d'autres espèces de cnidaires, associée à un stress thermique, et pourrait jouer un rôle dans le phénomène de nécrose chez le corail rouge [325, 174].



FIGURE 1.8 – Photographie d'une colonie de corail rouge (*Corallium rubrum*) présentant des zones de nécrose. La nécrose est visible par l'apparition du squelette nu strié et sa colonisation par des organismes épiphytes. Crédit photographie : Frédéric Zuberer.

A la suite de ces événements, des expériences en aquarium impliquant des individus prélevés au dessus (11 m) et en dessous (40 m) de la thermocline saisonnière dans la région de Marseille ont mis en évidence une thermotolérance plus élevée chez les individus de surface que chez les individus plus profonds (activité des polypes plus importante, taux de calcification plus élevé, apparition des nécroses retardées). Ces résultats ont été par la suite confirmés par Haguenauer *et al.* (2013)[143], qui ont associé les différences de thermotolérance observées à une induction de l'expression de la protéine HSP70 plus élevée chez les individus les

moins profonds (comparaison 5/20/40 m). Finalement, Ledoux *et al.* (2015) [183] ont conduit des transplants réciproques entre des individus issus de 20 m et de 40 m à Marseille et en Corse. Cette expérience a permis de mettre en évidence un signal d’adaptation locale de type “local vs étranger” (voir paragraphe 1.2.1, [166], Fig. 1.9A) à Marseille, mais pas de résultats significatifs en Corse (Fig. 1.9B). Ces expériences ont été couplées à une comparaison P_{ST} - F_{ST} qui a mis en évidence que les variations phénotypiques observées sur l’expérience conduite à Marseille ne pouvaient pas être le résultat de processus neutres uniquement, contrairement à l’expérience conduite en Corse (dans certains cas d’organismes non modèles pour lesquels des croisements en laboratoires ne peuvent pas être mis en place, la divergence d’un trait quantitatif ou Q_{ST} peut être approximé par la divergence d’un trait phénotypique ou P_{ST} , [45]). Les auteurs de cette étude ont également conduit une expérience en jardin commun en transplantant des individus issus de populations de 20 m et de 40 m à 5 m à Marseille, et ont montré que les individus de 40 m étaient significativement plus affectés par des nécroses que ceux de 20 m après 95 jours d’expérimentation. Enfin, des mesures de fécondité in situ à Marseille n’ont pu mettre en évidence aucune différence entre des individus de 15-22 m et des individus de 39-42 m [314], ce qui pourrait suggérer l’existence d’un compromis entre la survie/croissance et la fécondité chez les individus des populations de Marseille. Les différences physiologiques entre profondeurs mises en évidence par ces études chez le corail rouge s’accompagnent de différences génétiques sur des marqueurs neutres [185, 183, 143]. Un patron similaire a été observé chez la gorgone *Paramuricea clavata* [213], mais des différences de thermotolérance entre des populations génétiquement non différenciées ont aussi été mis en évidence chez *Eunicella cavolini* et *Eunicella singularis*. Finalement, Prada *et al.* (2007) [249] ont décrit un cas extrême pour lequel des populations issues de profondeurs différentes présentant des différences morphologiques chez la gorgone *Eunicea flexuosa* correspondaient à des lignées génétiques différentes, et probablement à des espèces différentes.

1.5 Objectifs

Cette thèse s’inscrit dans le cadre plus général du projet ANR Adacni (<http://adacni.imbe.fr>). L’objectif de ce projet est la compréhension de l’évolution de la

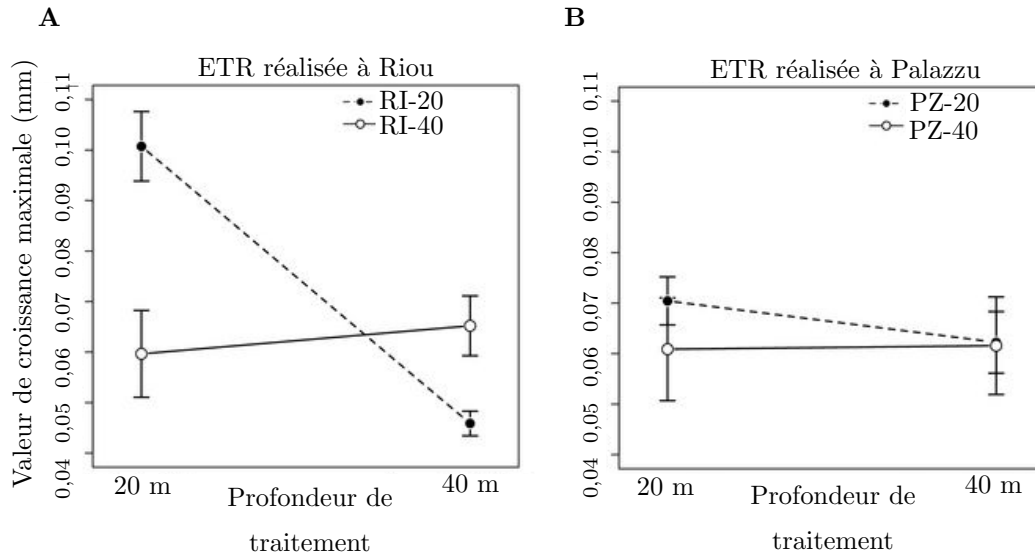


FIGURE 1.9 – Résultats de l’expérience de transplants réciproques sur le maximum moyen de la croissance en diamètre réalisée par Ledoux *et al.* (2015) [183] **A.** à Marseille, **B.** en Corse (barres : erreur standard ; $SE = SD/\sqrt{n}$ avec SD = déviation standard et n = nombre d’échantillons). Adapté de [183].

thermotolérance chez les cnidaires, avec trois modèles d’étude : l’anémone *Anemonia viridis*, le sclérectiniaire *Pocillopora damicornis* et le corail rouge *Corallium rubrum*. Sur ces trois espèces seul le corail rouge ne possède pas de *Symbiodinium*. Ces espèces présentent différents niveaux de thermotolérance et au sein de chaque espèce l’objectif était de comparer des populations issues de régimes thermiques plus ou moins variables afin de tester l’importance relative de l’acclimatation et de l’adaptation génétique. Pour cela deux approches complémentaires sont utilisées à chaque fois : une étude transcriptomique de la réponse au stress thermique en conditions expérimentales et une étude de génomique des populations avec recherche de signaux d’adaptation locale. Cette thèse se concentre sur l’étude de l’adaptation locale chez le corail rouge à partir d’approches populationnelles. Les aspects transcriptomiques sont réalisés dans le cadre de la thèse de Kelly Brener à l’Université de Perpignan.

L’ensemble des études menées sur le corail rouge confirment des différences de thermotolérance entre les individus issus de profondeurs différentes dans la région de Marseille (5 m, 20 m et 40 m). Ces différences peuvent être basées sur une réponse individuelle ou populationnelle au stress thermique, qui diffèrent notam-

ment quant à leur échelle de temps et leur stabilité évolutive. Des différences liées à l'histoire individuelle à travers l'acclimatation fournissent une réponse à court terme mais avec un héritage transgénérationnel rare ou instable à travers des processus épigénétiques. Des différences génétiques sont quant à elles plus stables au cours des générations et pourrait garantir une adaptation à l'échelle des populations stable à long terme. Les précédentes études de la thermotolérance chez le corail rouge avaient plus pour objectif l'identification de différences phénotypiques de thermotolérance que la compréhension de leur origine. Ces études ne suffisent donc pas pour faire la part entre adaptation locale et acclimatation individuelle qui nécessite un schéma expérimental dédié. De plus, bien que de nombreux travaux explorent les différences de thermotolérance chez des coraux tropicaux, eux aussi soumis à des épisodes de blanchissement massifs, une seule étude a identifié précisément des marqueurs potentiels d'adaptation locale [20]. Celle-ci, bien qu'envisagée théoriquement, reste un mécanisme largement sous-exploré d'un point de vue empirique chez les coraux. La problématique générale de cette thèse est donc de **comprendre les bases génétiques des différences de thermotolérance observées chez le corail rouge**.

Trois questions ont été abordées pour répondre à cette problématique :

1. Les différences d'expression de gènes chez des individus issus de profondeurs différentes en situation de stress thermique sont-elles maintenues en situation basale ?

Les précédentes études de la thermotolérance chez le corail rouge sont majoritairement basées sur des expériences de stress thermique en aquarium. Cependant, l'identification de différences entre individus issus d'environnements différents en situation de stress n'est pas un bon indicateur pour faire la part entre acclimatation individuelle et adaptation génétique. Il s'agira donc dans le Chapitre 2 de cette thèse de confirmer que des différences phénotypiques stables au cours du temps existent entre des individus issus d'environnements différents. L'absence de différences phénotypiques stables au cours du temps serait un argument en faveur d'une réponse dominée par un mécanisme d'acclimatation individuelle. Cette partie repose sur une approche transcriptomique. Les données obtenues ont aussi pu être utilisées pour tester l'hypothèse phylogénétique de la monophylie des antho-

zoaires, une question d'actualité car remise en cause par des travaux récents. Les résultats de cette étude sont présentés dans l'Annexe A.

2. Les différences de thermotolérance observées entre individus issus de profondeurs différentes dans la région de Marseille sont-elles liées à un mécanisme d'adaptation locale des populations ?

L'adaptation locale peut être basée sur des associations entre génétique et environnement pour certains locus présentant certains allèles en fréquences plus élevées dans un environnement donné en réponse à la sélection naturelle. La mise en évidence d'un tel patron serait cruciale dans la compréhension de l'adaptation à la profondeur chez le corail rouge. L'objectif du Chapitre 3 sera donc de rechercher des variations génétiques associées à une adaptation à des environnements différents. Cet objectif requiert une description précise préalable de la structure génétique des populations étudiées, notamment des relations génétiques neutres existant entre les profondeurs étudiées, qui seront donc analysées dans ce chapitre. Par ailleurs dans le cas d'une espèce à forte structure génétique comme le corail rouge, il est nécessaire de bien évaluer l'efficacité des différentes méthodes d'étude de l'adaptation locale au niveau moléculaire.

3. La réponse à une situation thermique similaire est-elle basée sur des variations génétiques similaires dans des régions génétiquement indépendantes ?

Bien que les observations de différences de thermotolérance entre populations de surface et plus profondes chez le corail rouge aient été réalisées à Marseille, l'existence de différences similaires dans d'autres régions géographiques présentant des contrastes de températures similaires entre profondeurs est envisageable. Il est donc pertinent de comparer les éventuels signaux génétiques d'adaptation locale qui seront explorés dans la région de Marseille avec des résultats obtenus dans des groupes génétiques indépendants mais environnementalement similaires. Cet aspect sera lui aussi abordé dans le Chapitre 3.

Finalement, la recherche de signaux d'adaptation locale chez le corail rouge nous a conduit à observer des marqueurs génétiques du sexe chez cette espèce. Le Chapitre 4 traitera de l'identification du déterminisme génétique du sexe chez le corail rouge.

Chapitre 2

Recherche de différences
transcriptomiques à l'état basal
entre individus issus
d'environnements contrastés

Chapitre 2 : Recherche de différences transcriptomiques à l'état basal entre individus issus d'environnements contrastés

Ce chapitre correspond à un article publié dans *Molecular Ecology Resources* (2015).

Objectifs

La majorité des études portant sur la thermotolérance chez le corail rouge est basée sur des stress en aquarium d'individus issus d'environnements contrastés. Ces études sont construites pour mettre en évidence des différences de réponses entre ces individus dans des conditions dans lesquelles elles sont le plus propice à se manifester. L'objectif est, dans notre cas, de tenter de faire la part entre une réponse due à l'acclimatation, au cours de la vie d'un individu, à ses conditions environnementales locales, et une réponse de type adaptation génétique de la population dont est originaire cet individu à ces conditions. Pour cela, une expérimentation en jardin commun a été mise en place dans le but de diminuer au maximum l'acclimatation des individus à leur environnement d'origine, de les placer dans un état physiologique basal, non stressé, et d'identifier la présence ou l'absence de différences phénotypiques maintenues à moyen terme. Bien que l'acclimation aux conditions de vie d'un individu puisse persister à plus ou moins long terme, l'identification de différences d'expression de gènes entre des individus issus de profondeurs différentes après plusieurs mois en jardin commun pourrait constituer un premier indice en faveur de l'existence d'un signal d'adaptation locale dans ces populations. Finalement, la génération et l'annotation du transcriptome du corail rouge permet de mettre à disposition un nouvel outil pour les étapes suivantes de cette thèse et plus largement pour le projet Adacni, ainsi que pour les futures études sur le corail rouge.

Résultats

Nous mettons ici en évidence des différences d'expression de gènes maintenues entre les individus issus de deux profondeurs après une période en jardin commun. Certains des gènes présentant ce patron de différences d'expressions correspondent à des fonctions ayant déjà été décrites pour la réponse au stress chez des cnidaires

et sont de bons candidats pour l'étude de l'adaptation locale chez le corail rouge, ainsi que de forts arguments en faveur de la conservation de leur fonction de réponse au stress chez les cnidaires. On peut par exemple citer cinq gènes de type *tumor necrosis factor receptors associated-factors* (TRAFs) surexprimés chez les individus de faible profondeur et qui ont déjà été mis en évidence chez le corail *Acropora hyacinthus* comme surexprimés de façon basale chez les individus issus d'un environnement plus variable [17]. Plusieurs SNPs présentant des allèles différentiellement fixés entre les deux profondeurs étudiées sont mis en évidence ici et pourraient aussi constituer de bons candidats pour l'étude de l'adaptation locale si ces fréquences alléliques étaient confirmées à l'échelle populationnelle. Nous fournissons ici de nouvelles ressources pour l'étude de l'adaptation chez le corail rouge, mais aussi des données de polymorphisme qui pourront servir pour construire de nouveaux marqueurs utilisables en génétique des populations chez cette espèce.

The red coral (*Corallium rubrum*) transcriptome: a new resource for population genetics and local adaptation studies

Pratlong M.^{1,2}, Haguenauer A.¹, Chabrol O.², Klopp C.³, Pontarotti P.² & Aurelle D.¹

1. Aix Marseille Université, CNRS, IRD, Avignon Université, IMBE UMR 7263, 13397, Marseille, France
2. Aix Marseille Université, CNRS, Centrale Marseille, I2M UMR 7373, 13453, Equipe Evolution Biologique et Modélisation, Marseille, France
3. Plateforme Bioinformatique Toulouse Midi-Pyrénées, UR 875 UMIAT, INRA, Auzeville Castanet-Tolosan, France

Abstract

The question of species survival and evolution in heterogeneous environments has long been a subject for study. Indeed, it is often difficult to identify the molecular basis of adaptation to contrasted environments, and non-genetic effects increase the difficulty to disentangle fixed effects, such as genetic adaptation, from variable effects, such as individual phenotypic plasticity, in adaptation. Nevertheless, this question is also of great importance for understanding the evolution of species in a context of climate change. The red coral (*Corallium rubrum*) lives in the Mediterranean Sea, where, at depths ranging from 5 to 600 m it meets very contrasted thermal conditions. The shallowest populations of this species suffered from mortality events linked with thermal anomalies that have highlighted thermotolerance differences between individuals. We provide here a new transcriptomic resource, as well as candidate markers for the study of local adaptation. We sequenced the transcriptome of six individuals from 5 m and six individuals from 40 m depth at the same site of the Marseilles bay, after a period of common garden acclimatization. We found differential expression maintained between the two depths even after common garden acclimatization, and we analysed the polymorphism pattern of these samples. We highlighted contigs po-

tentially implicated in the response to thermal stress, and which could be good candidates for the study of thermal adaptation for the red coral. Some of these genes are also involved in the response to thermal stress in other corals. Our method enables the identification of candidate loci of local adaptation useful for other non-model organisms.

2.1 Introduction

Marine species usually deal with more or less pronounced gradients of environmental conditions (temperature, salinity...) along their natural range. These species might have found adaptive responses to maximize their fitness in all environmental conditions they encounter. In this context, two types of mechanisms could occur. First, in the absence of other evolutionary pressures (migration, genetic drift), and if the environmental gradient is persistent for an extended period of time, each local population submitted to local selection could become genetically adapted to the corresponding local environmental conditions [166]. On the other hand, during their life, individuals can develop physiological adaptations to their local environmental conditions. This acclimatization (or phenotypic plasticity), is often reversible at short term, but can be in some cases, maintained at medium to long-term [237]. In this context, we are interested in understanding the ability of individuals from contrasted environments to revert to a common basal state and in knowing if they developed adaptive mechanisms stable at medium or long-term. Two experimental approaches are mainly used for the study of local adaptation : reciprocal transplant and common garden experiments. Recent experimental studies of local adaptation in marine environments generally focus on the identification of differentially expressed genes between individuals from different environments and submitted to the same stress [237, 17, 19, 143]. In these cases, few candidates loci are commonly used [19, 143]. Nevertheless, transcriptomic patterns of individuals from different environments but acclimatized to a basal state in common garden conditions remain poorly researched. If two individuals from different environmental conditions are kept in common garden conditions, expression differences due to phenotypic plasticity should tend to decrease in time, while fixed effects stemming from genetic adaptation, epigenetic or developmental modifications with medium

to long-term stability will be maintained [17, 227]. Furthermore, most studies treated separately the question of differential expression and sequence polymorphism, and these two aspects are rarely studied in the same framework [93]. Differential expression analysis and expressed sequences polymorphism have different molecular and evolutionary origins. Indeed, these two aspects may bear traces of adaptation (broad sense) and taking them into account may contribute to our global understanding of this complex mechanism. Here we propose to study local adaptation and acclimatization along an environmental gradient based on transcriptomic data for the red coral (*Corallium rubrum*). This species is an octocoral, an ecologically important clade but less studied than hexacorals. It lives in the Mediterranean Sea and Eastern Atlantic between 5 m and 800 m depth [79] where it meets very contrasted thermal conditions. Shallow populations are exposed to high maximum temperatures and to frequent and intense temperature fluctuations. These two parameters tend to decrease with increasing depth. Previous studies also showed that the red coral is an especially low disperser, with populations separated by fewer than 10 m detected as genetically different [185, 184]. Furthermore, experiments have shown that individuals from different depths (20 – 40 m) of the same site have contrasted responses to thermal stress, in terms of respiration, growth, calcification and necrosis [316, 185, 184]. The study of candidate genes in experimental conditions of thermal stress revealed the differential expression of HSP70 between individuals from different depths of the same site [143]. The present study is a logical continuation of this previous research concerning the local adaptation of the red coral. As such, it is now necessary to use new molecular approaches to try to identify genes that could be good candidates for the study of local adaptation, because of their sequence polymorphisms or expression pattern, without preconceived ideas concerning candidate genes. To this end, we sequenced the transcriptome of individuals from two depths of the same site in the bay of Marseilles, after three months of in situ common thermal conditions and a one month aquarium acclimatization. Our objective here was to evidence underlying expression differences stable at medium term between individuals from different depths of the same site. We also studied the patterns of sequence polymorphism according to depth. The general goal was to use these new transcriptomic resources to highlight candidate loci for the study of local adaptation in heterogeneous environment. This method

could be extended to other non-model organisms. Additionally, this sequence database will be useful for the study of the response of other octocoral species to thermal stress.

2.2 Material and Methods

2.2.1 Sampling and RNA extraction

Red coral (*Corallium rubrum*) colonies were collected from two areas near the city of Marseilles, the 5 m population of Figuier cave (FIG; 43 ° 12.330'N, 5 ° 26.790'E) and the 40 m population of Cap Morgiou (MOR; 43 ° 12.060'N, 5 ° 27.060'E) separated by 622 m of horizontal distance, with six individuals sampled at each depth. These two sampling depths are characterized by contrasted thermal conditions from April to the end of September [143]. The population from Figuier is the shallowest red coral population near Marseilles. These individuals are submitted to a large range of thermal conditions in spring and summer (from 15 to 26 ° C), to fast and daily increase or decrease of temperature during this period, and are exposed to more days with temperatures above 20 ° C than the deeper population from Morgiou [143]. Individuals from Morgiou are submitted to a narrower thermal range (from 14 to 20 ° C), to less frequent thermal variations, and the temperature stays generally under 20 ° C. From October to March, the temperature and others physical and chemical parameters are similar between these two depths because of a seasonal homogenization of the water column [100]. During this period, the thermal homogenization reaches a common minimal value of around 13 ° C [143]. Furthermore, because the red coral does not bear photosynthetic symbionts, individuals from 5 m and 40 m should not be differentially impacted by light exposition differences. Sampling took place in December, when colonies from 5 m and 40 m had spent three months in common thermal conditions. Six individuals were sampled from each site and maintained during one additional month in common conditions in a semi-closed water circulation system with a temperature between 14 and 15 ° C. After this acclimatization period, total RNA was extracted.

2.2.2 Library preparation and sequencing

Total RNA from each sample was purified as previously described in Haguenaer *et al.* (2013) [143]. Residual DNA was digested using TurboDNase (Ambion) following the manufacturer’s instructions. RNA samples were pooled in two duplicates for each depth (each duplicate containing the same quantity of RNA from three different individuals coming from the same depth). RNA-Seq libraries were generated using the TruSeq RNA-Seq Sample Prep kit v2 according to the manufacturer’s protocol. During this preparation, libraries were individually tagged in order to pool them before sequencing. Library sizes were controlled on analysis with a Bioanalyzer 2100 (Agilent Technologies) on High Sensitivity chips. The four libraries were quantified by qPCR following the manufacturer’s protocol. Libraries were finally pooled before sequencing on one lane on the Illumina HiSeq2000 (Illumina Inc., San Diego, CA) as paired-end reads of length 100 bp with the TruSeq PE kit. Library preparation and sequencing were performed at the GenoToul platform (<http://get.genotoul.fr/>).

2.2.3 Assembly

One assembly was run for each condition (4 assemblies were performed in total). Raw sequence data were first filtered to remove unknown nucleotides. If a read contained Ns, it was split in subsequences without Ns and the longest of these subsequences was kept if its length exceeded half of the initial read length (a subsequence is a part of a read with no unknown nucleotides). Subsequently, because raw data produced are not filtered for low quality reads, we applied the `fastq_illumina_filter` (v.0.1 available at http://cancan.cshl.edu/labmembers/gordon/fastq_illumina_filter/). Finally, duplicated reads were discarded and only unique reads were kept to perform assembly. The transcriptome de novo assembly was performed using Oases (v.0.2.06; [285]). Nine assemblies using nine different k-mers (25, 31, 37, 43, 49, 55, 61, 65, 69) were performed on pre-processed input data. We chose to keep only the best contig for each locus with a script developed by a Brown University team (available at <https://sites.google.com/a/brown.edu/bioinformatics-in-biomed/velvet-and-oases-transcriptome>). After that, contigs given by all k-mers were merged. Finally, anti-sense chimeras (accidentally produced by the assembly

step) were cut. Then, because similar collection of contigs were produced by close k-mers, a cd-hit-est clustering step (v.4.6; [192]) grouped similar contigs into clusters based on their sequence similarities (identity equal or greater than 0.95). TGICL (v.2.1; [234]), an OLC (overlap layout consensus) assembler, clusterized sequences sharing significant fragments. After this assembly process, all input reads were mapped back to rebuilt contigs using BWA (v.0.7.0-r313; [189]). For each contig, the longest ORF (region that is free of STOP codons) was extracted with the getorf EMBOSS tool (v.6.4.0.0; [264]) and contigs with longest ORF lower than 200 bp (66 aa) or very low coverage (less than 2 mapped reads for 1 million overall mapped reads) were discarded.

2.2.4 Meta-assembly

The second step of the assembly process was a meta-assembly resulting in the production of a single reference contig set. Contig fasta files coming from the four conditions were first concatenated. ORFs with sequence identity equal or greater than 0.9 were clusterized using cd-hit (v.4.6; [192]). Then, the contig with the longest ORF (or the longest contig if several contigs have an ORF of the maximal size) of each cd-hit cluster was extracted. A cd-hit-est job was performed to clusterize remaining contigs with sequence identity equal or greater than 0.95. This step will be able to clusterize close contigs inside which potential frameshifts abort ORF detection. Finally, input reads from all conditions were mapped to selected contigs using BWA (v.0.7.0-r313; [189]) and contigs with longest ORF lower than 200 bp (66 aa) or very low coverage (less than 2 mapped reads for 1 million overall mapped read) were discarded.

2.2.5 Comparative analysis

In order to compare the transcriptome of the red coral with those of other cnidarians, we performed a blastp [4] search in the following cnidarian ESTs libraries : *Acropora palmata*, *Acropora millepora*, *Pocillopora damicornis*, *Stylopora Pistillata*, *Porites Astreoides*, *Monastrea faveolata*, *Aiptasia Pallida*, *Anemonia Viridis*, *Metridium senile*, *Nematostella vectensis*, *Edwardsiella lineata*, *Paramuricea clavata*, *Eunicella verrucosa*, *Eunicella cavolinii*, *Gorgonia ventalina*, *Leptogorgia sarmen-*

tosa, *Hydractinia echinata*, *Hydra vulgaris*, *Clytia hemisphaerica* (Table 2.1). We applied an e value cut-off of 10^{-10} .

2.2.6 Transcriptome annotation

All contigs were annotated using a two steps method which allowed to maximize the reliability of our annotation. For all contigs, we found the longest open reading frame using the getorf EMBOSS tool [264]. First, we used orthoMCL [191] to search for potential ortholog reciprocal relationships between the translated contigs, sequences of seven very well annotated transcriptomes (*Adineta vaga*, *Arabidopsis thaliana*, *Caenorhabditis elegans*, *Drosophila melanogaster*, *Homo sapiens*, *Saccharomyces cerevisiae*, *Strongylocentrotus purpuratus*), and those of two transcriptomes of cnidarian species (*Hydra magnipapillata*, *Nematostella vectensis*). We applied an e value cut-off of 10^{-5} and we transferred the annotation from these orthologs, when available, to *C. rubrum* sequences. For all contigs which could not be annotated by this method, we performed blastp searches against the UniProt protein database [4, 313]. An e value cut-off of 10^{-5} was applied and we reported only the best hit of each query sequence. Among these UniProt hits, we listed the non metazoan hits to avoid the bias of contaminations in our analysis. In all cases, we used the UniProt identifier to assign each contig to Gene Ontology biological process categories [11] associated to the predicted protein. Annotation results were summarized using Generic GO-Slim, a cut-down version of Gene Ontology containing a small number of relevant functional categories (<http://www.geneontology.org/ontology/subsets/goslim-generic.obo>).

2.2.7 Differentially expressed genes

In order to identify differentially expressed genes between the two depths, we compared the results of three packages in R [309] : edgeR [266] and DESeq [6] using parametric statistical methods and NOISeq [308] using non parametric statistical method. In each case, we used parameters described as optimal by the authors. EdgeR was performed using the TMM (trimmed mean of M values) normalization

TABLE 2.1 – Number of ESTs used for comparative analysis of each species and number of BLASTP hits between the red coral transcriptome and the EST bank of each species (evalue = 10^{-10})

Classification	Species	Source	Number of ESTs	Number of contigs
Cnidaria	<i>Acropora palmata</i>	NCBI	43150	11671
Anthozoa	<i>Acropora millepora</i>	[215]	52 958	18 108
Hexacorallia	<i>Pocillopora damicornis</i>	[317]	70 786	14 716
Scleractinia	<i>Stylopora Pistillata</i>	[165]	21 810	16 664
	<i>Porites Astreoides</i>	NCBI	11 516	4 753
	<i>Monastrea faveolata</i>	NCBI	33 226	9 606
Cnidaria	<i>Aiptasia Pallida</i>	NCBI	10 295	6 391
Anthozoa	<i>Anemonia Viridis</i>	NCBI	39 939	10 535
Hexacorallia	<i>Metridium senile</i>	NCBI	29 412	9 815
Actinia	<i>Nematostella vectensis</i>	NCBI	163 314	14 906
	<i>Edwardsiella lineata</i>	[303]	90 440	17 166
Cnidaria	<i>Paramuricea clavata</i>	Mokthar-Jamaï <i>et al.</i>	6 529	4 341
Anthozoa	<i>Eunicella verrucosa</i>	[268]	15 461	15 130
Octocorallia	<i>Eunicella cavolinii</i>	[268]	20 731	17 982
	<i>Gorgonia ventalina</i>	[49]	90 230	20 172
	<i>Leptogorgia sarmentosa</i>	[268]	13 299	14 343
Cnidaria	<i>Hydractinia echinata</i>	NCBI	9 464	2 860
Medusozoa	<i>Hydra vulgaris</i>	NCBI	184 731	12 324
	<i>Clytia hemisphaerica</i>	NCBI	85 991	10 643
	<i>Aurelia aurita</i>	www.compagen.org	62 632	9 836

method [266] and a common negative binomial dispersion parameter for the variance estimation. In the case of NOISeq, we applied the RPKM (reads per kilobase per million) normalization method [214] and we used the noiseqbio function which is appropriate to treat biological replicates. In each case, we corrected p values for false discovery rate using the Benjamini-Hochberg procedure [23]. As our aim was to find the best candidates which could play a role in the depth differentiation, we chose to consider only genes detected as differentially expressed with an adjusted p value lower than 0.01 by these three statistical methods in order to be as stringent as possible. To identify functional categories enriched in our differentially expressed genes, we performed Fisher exact tests and applied the false discovery rate correction [23].

2.2.8 Detection and analysis of SNPs

The alignment files were first filtered to remove all PCR duplicates (samtools rmdup, [190]) and all reads were aligned in multiple locations. The reads were then realigned and recalibrated using GATK (v.2.4-9, standard parameters; [208]. Finally, all files were used to call high-quality SNPs and INDEL with GATK (same version, UnifiedGenotyper method; [95]) using a minimum Phred quality score of 30 (Q30, corresponding to an error rate in base calling lower than 0.1 %). From this SNP file, we searched for polymorphism patterns potentially linked with depth. Our experimental framework did not allow us to precisely estimate SNPs frequencies as individuals were pooled before sequencing and observed frequencies may also depend on expression levels. Therefore, we focused on the extreme case of differentially fixed SNPs (i.e. SNPs with the same allele for all individuals from the same depth and with an alternative allele for all individuals from the other depth). At this stage, to minimize errors in the detection of SNPs, we added a filter for the quality of each genotype. GATK variant calling generates the Genotype Quality (GQ) value that corresponds to the Phred-scaled confidence that the inferred genotype is true. We chose to apply the commonly used threshold of $GQ = 20$, and to discard low depth genotypes ($DP < 12$) in order to keep only high quality genotypes. We tested for enriched functional categories within contigs containing differentially fixed SNPs and non-synonymous SNPs using Fisher exact test, and we applied the false discovery rate correction [23].

2.3 Results

2.3.1 Transcriptome description

De novo assembly of reads resulted in 48 074 contigs with a mean length of 1 814 bp and N50 of 2 470 bp (Fig. 2.1). The Table 2.2 sums up the number of reads and alignments at different stages of the assembly process. The GC content of the *C. rubrum* transcriptome was 39.45 %. Our annotation method allowed us to successfully annotate 16 951 contigs (36 % of the red coral transcriptome). Among these contigs, 64 % could be annotated through orthoMCL. Among the 6 187 annotated contigs which failed to be annotated through orthoMCL, the majority (73 %) had a hit with species we did not use in the orthoMCL analysis. The last 27 % could be sequences not present in the transcriptomes used in orthoMCL analysis or eliminated by the orthoMCL filters, as the orthoMCL algorithm is more stringent than a simple blast. Among contigs which failed to be annotated through orthoMCL, 426 had non metazoan Uniprot hits, of which half (213) had similarities with bacteria. These non metazoan hits could be due to contaminations but also to horizontal transfers. In any cases, we chose not to include contigs with non metazoan hits in further analyses. Functional categories of annotated contigs were distributed into the 55 categories of biological processes of the Generic GO-Slim, but only eight categories contained more than 50 % of functions assigned to contigs (Fig. 2.2) : cellular nitrogen compound metabolic process (12 %), biosynthetic processes (10 %), signal transduction (6 %), cellular protein modification process (6 %), transport (6 %), small molecule metabolic process (5 %), catabolic process (4 %), anatomical structure development (4 %). Twenty-nine categories were represented by less than 1 % of functions attributed to contigs. The low number of data available on Uniprot for species phylogenetically close to the red coral could partly explain the absence of significant hit for 31 123 sequences. Moreover, the mean length of these non-annotated sequences was lower than for the annotated ones ($p < 2.2 \cdot 10^{-16}$) and could correspond to non coding RNA or technical artefacts (Fig. B.1).

TABLE 2.2 – Number of reads and alignments at different stages of the assembly process. The Q30 (Phred quality score of 30) corresponds to an error rate in base calling lower than 0.1 %.

Sample	Number of reads pairs	Number of alignments	Number of alignment after Q30 filtering	Number of alignments after PCR duplicates removing
FIG5A	57 446 789	114 893 578	47 979 281	15 155 126
FIG5B	51 612 319	103 224 638	43 854 431	14 846 834
MOR40A	69 032 155	138 064 310	58 804 098	17 071 186
MOR40B	62 839 889	125 679 778	55 166 310	16 897 159

2.3.2 Comparative analysis

14 135 of the red coral’s contigs had homologs within the four clades of cnidarians (Fig. 2.3, Table 2.1). The highest number of clade specific homologs was found in octocorals (2 680 contigs), followed by scleractinians (294 contigs), actinians (209 contigs) and medusozoans (87 contigs) (Fig. 2.3, Table 2.1).

2.3.3 Analysis of depth specific expression patterns

Expression patterns of our 48 074 contigs were highly similar between the two pools of the same depth, but also between sets of different depths : the Pearson’s correlation coefficient was ranged from 0.94 to 0.96 ($p < 0.0001$) for all comparisons (Table 2.3). Despite this correlation, 212 contigs were detected as differentially expressed with the three methods (FDR $\alpha = 0.01$, Table 2.4, Fig. B.2). Among these 212 contigs, 92 had higher expression at 5 m and 120 had higher expression at 40 m (Table 2.5). The level of over-expression was greater at 40 m than at 5 m ($p = 1.37 \times 10^{-6}$, Table 2.5). We found Uniprot hits for 77 of these 212 contigs (Table B.1), but only 41 could be assigned to 18 Gene Ontology functional categories. Four of these categories contained more than 50 % of functions assigned to contigs : cellular protein modification process (16 %), biosynthetic process (13 %), signal transduction (11 %), cellular nitrogen compound metabolic process (11 %). Among the functions assigned to the 41 differentially expressed contigs, two were detected as enriched by the functional enrichment analysis with a threshold of

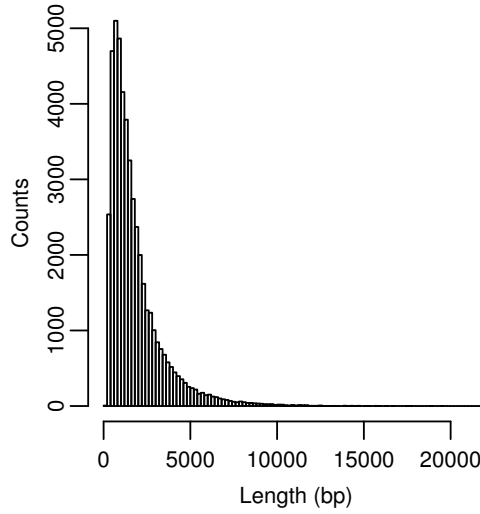


FIGURE 2.1 – Size distribution of contig lengths (bp).

0.05: cellular protein modification process ($p = 0.0054$) and cell death ($p = 0.048$). None of these functional categories remained significantly enriched after the false discovery rate correction (FDR $\alpha = 0.01$). The best informative Uniprot hit for the 77 differentially expressed contigs annotated are listed in Table B.1. Among them, we can highlight several noteworthy functions because they are classified into the two Gene Ontology functional categories “cellular protein modification process” and “cell death” mentioned before and potentially involved in corals adaptation (see discussion). Five contigs were homologous to *Tumor necrosis factor receptor-associated factors* (TRAFs), and all were over-expressed at 5 m. A homolog of *TNFAIP3 interacting protein* was over-expressed at 40 m. Two homologs of *E3 ubiquitin protein ligase* were over-expressed at 40 m and one at 5 m. Finally, we can cite a lectin homolog and a *F-type lectin* homolog over-expressed at 5 m. Among the 134 contigs differentially expressed without Uniprot hit, only 18 had homologs in the cnidarians datasets used for the comparative analysis (Table 2.1). Twelve of these contigs had hits only with octocorals.

2.3.4 Polymorphism analysis

The SNP detection allowed the identification of 691 940 SNPs within 43 474 contigs (5 000 contigs were not polymorphic). Two of these SNPs presented four alleles, 1 715 presented three alleles and all others SNPs had two alleles. The

TABLE 2.3 – Pearson correlation coefficient of population expression patterns.

	FIG5A	FIG5B	MOR40A	MOR40B
FIG5A	1			
FIG5B	0.96	1		
MOR40A	0.94	0.96	1	
MOR40B	0.94	0.95	0.96	1

TABLE 2.4 – Number of contigs detected as differentially expressed between depths by each R package. The total row indicates contigs detected as differentially expressed by the three methods.

	Differentially expressed contigs (FDR $\alpha = 0.01$)
DESeq	358
NOISeq	809
EdgeR	477
Total	212

TABLE 2.5 – Number of contigs overexpressed at each depth and their average fold change.

Overexpression	Average fold-change (x)				Total
	$x < 2$	$2 < x < 10$	$10 < x < 100$	$x > 100$	
Over-expressed at 5 m	4	58	24	6	92
Over-Expressed at 40 m	0	39	63	18	120

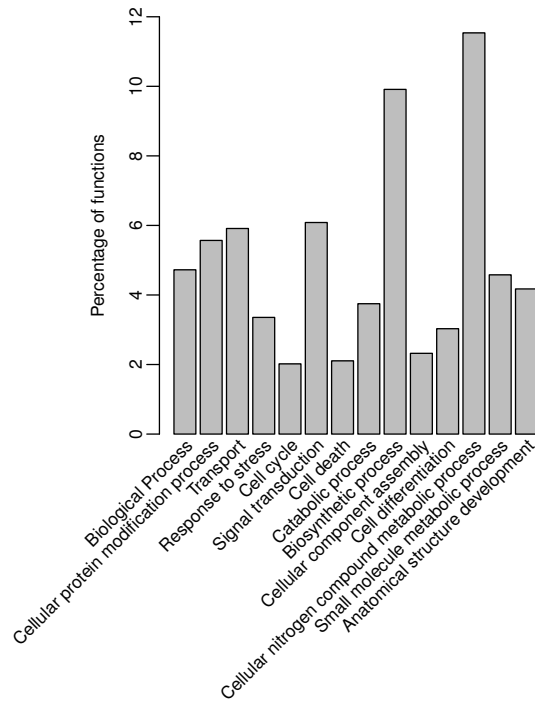


FIGURE 2.2 – Repartition of the functions attributed to contigs into the biological process GO functional categories. Functional categories represented by less than 2 % of functions attributed to contigs were not reported. The functional category ‘biological process’ represents the functions belonging to the biological process ontology, but which could not be classified in the categories presented here.

frequency of SNPs in the transcriptome was 1 per 126 bp. Although the identified SNPs had very good quality, the genotype quality was poor (Fig. B.3). Only 237 681 SNPs passed the genotype quality filter $GQ = 20$ and $DP = 12$. 73 % (173 825 SNPs) of these SNPs with high quality genotypes were in the longest open frame of the corresponding contig and 31 % were non synonymous mutation (72 954 SNPs). Among these SNPs, we found 56 which were differentially fixed SNPs and which corresponded to 46 contigs. Thirty six (63 %) of these differentially fixed SNPs were in the longest open reading frame of the corresponding contig. Among these 36 SNPs, 19 (distributed into 15 contigs) presented non-synonymous mutations. There was no enrichment of non-synonymous mutations among differentially fixed SNPs compared to the global set of SNPs with high quality genotype ($p = 1$). The best informative Uniprot hits (when it was available) for these contigs containing differentially fixed SNPs are listed in Table B.2. Only 19 of them could

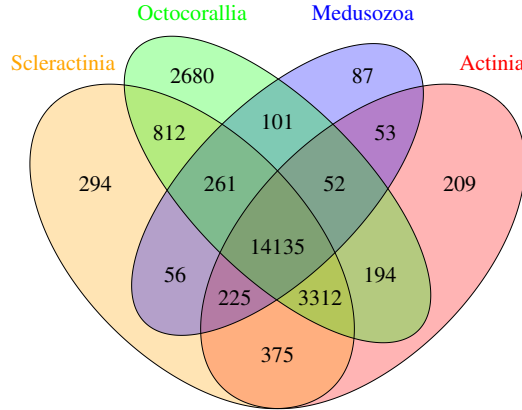


FIGURE 2.3 – Venn diagram of red coral’s homologues with the cnidarians (Scleractinia, Actinia, Octocorallia, Medusozoa).

be assigned to 21 Gene Ontology functional categories. The functional categories the more represented were cellular nitrogen compound metabolic process (14 %), cellular protein modification process (11 %), and transport (9 %). No enrichment for any functional category was detected by the functional enrichment analysis. Twelve differentially expressed contigs (including a lectin homolog and two of the five TRAFs homologs previously mentioned) contained at least one differentially fixed SNP before the correction for genotype quality, but none of them were detected by the two approaches after that correction.

2.4 Discussion

Our results enabled us to explore the molecular basis of adaptation to local environments for the red coral and to complement previous results on adaptive diversity in this species [143]. Indeed, the transcriptome variability according to depth was yet to be studied for this species, in terms of both expression levels and sequence polymorphism. More generally, our study is the first to compare transcriptomic patterns between individuals from contrasted environmental conditions for a temperate octocoral. Extending such approaches to an octocoral in a domain where hexacorals are mostly studied is an important step in the study of the evo-

lution of adaptive processes among cnidarians. The choice of a cnidarian that does not bear algal symbionts (*Symbiodinium*) also allowed us to overcome the effect of algal symbionts on stress response and on adaptation to local environment.

2.4.1 Transcriptome comparative analysis

Our comparative analysis allowed us to check for the congruence of the red coral transcriptome among others ESTs projects. As expected the highest number of clade specific homologs corresponded to octocorals. We found more clade specific homologs with hexacorals (scleractinians and actinians) than with medusozoans. These relationships correspond to the most usual phylogeny of cnidarians [68, 311] although two recent papers uncovered a closer relationship of octocorals with medusozoans rather than with hexacorals [229, 168]. This analysis would therefore support the usual phylogenetic relationships within Cnidaria, but we could not exclude biases associated with the size and content of the ESTs libraries compared here.

2.4.2 Biological functions involved in coral adaptation to thermal stress

We evidenced differentially expressed contigs between individuals from different depths in basal conditions (i.e. without thermal stress). These genes suggest an effect of environmental conditions at each depth acting on expression regulation either through genetic adaptation or acclimatization allowed by phenotypic plasticity at the transcriptomic level. Several functions identified here as differentially expressed between depths have been experimentally demonstrated as involved in other cnidarians' response to thermal stress, thereby supporting their potential involvement in adaptive response (broad sense) [267, 339, 179, 89, 93, 17, 326]. These genes could also constitute potential stress markers for experimental or in situ studies. Among these genes, we identified five homologs of *Tumor necrosis factor receptors associated-factors* (TRAFs), all over-expressed at 5 m. These genes play a role in several signalling pathways in cell death, survival and mechanisms of cellular response to stress such as innate immunity or apoptosis in other Metazoans [8, 41]. This gene family has been documented as a marker of acclimatization to

thermal stress in the hexacoral *Acropora hyacinthus* by Barshis *et al.* (2013) [17]. These authors observed a higher basal level (i.e. before experimental stress) of expression of TRAFs homologs in *A. hyacinthus* individuals from thermally variable environments compared with more stable ones. Moreover, Palumbi *et al.* (2014) [227], using reciprocal transplants, highlighted an overexpression of TRAF homologs in highly variable environments compared to moderately variable ones. Our results mirror these observations and extend them to temperate octocorals : the basal over-expression of these genes might be a conserved response of cnidarians to thermal stress. An homolog gene of the *TNFAIP3 interacting protein*, also called ABIN, was over-expressed at 40 m in red coral. Over-expression of ABIN has been shown to inhibit NF- κ B activation by tumor necrosis factor [324]. De Wit & Palumbi (2013) [93] found a gastropod homolog of *TNFAIP3 interacting protein* as outlier based on sequence polymorphism. Nevertheless we did not find any differentially fixed SNPs in the red coral homolog of TNFAIP3. This description of an TNFAIP3 as a potential candidate locus for the study of spatial adaptation, although interesting, requires further studies to be considered. The lectins have been described in all Metazoans and are involved in cell-cell interactions and pathogen recognition [337]. Concerning cnidarians, lectins have been mainly cited for their role in the interaction between host and *Symbiodinium* for symbiotic species [339, 89]. Additionally, lectin homologs from *Acropora millepora* and *Pocillopora damicornis* have been shown to be implicated in the bleaching process during thermal stress experiments [267, 326]. Nevertheless, Kvennefors *et al.* (2008) [179] highlighted the ability of pathogens recognition for the lectin of the coral *Acropora millepora* and suggested a conservation of the function of lectin in the innate immunity between cnidarians and the bilaterians. Moreover, Martin *et al.* (2002) [203] pointed that an increase in sea water temperature could favour the settlement of opportunistic pathogens or the increase of the sensibility of Mediterranean octocorals. Thus, the over-expression of the lectin red coral homolog at 5 m could be an adaptation to an environment where pathogens pressure would be higher : this remains to be investigated. We thus identified new potential expression markers for the study of adaptation, but other candidate genes were not recovered as differentially expressed. For example, Heat Shock Proteins homologs (three HSP60 and three HSP70) were expressed at the two depths, but without significant dif-

ferential expression. This contrasts with experimental stress studies in cnidarians where HSPs have been shown as involved in stress response [96, 17, 143, 170]. Here, HSPs would then correspond to an inducible response and the absence of expression in our samples could be considered as a negative control of the stress state of our colonies.

2.4.3 Phenotypic plasticity or genetic adaptation ?

TRAF proteins seem to be consistent markers of the native environmental origin of individuals but we still cannot discriminate the effects of phenotypic plasticity versus genetic adaptation. With the three-month period of common in situ temperature and a one-month period of acclimatization in common garden conditions, we expected to attenuate differences due to the environment of origin. Despite this, the effects of the original environment on individuals persisted. This has been much discussed recently and several non-genetic effects (plasticity, maternal effects, epigenetic processes) were probably maintained during our experiment and may cause the detection of some of the 212 differentially expressed genes [166, 235]. The effect of acclimatization on specific environmental conditions should decrease with the increase of the period of common acclimatization, but the time needed to reach the complete loss of the acclimatization effect is still unclear. Pespeni *et al.* (2013) [237] found a signal of this environment condition memory even after three years of acclimatization in common conditions in the purple sea urchin *Strongylocentrotus purpuratus*. Indeed, we can not characterize the proportion of genetic versus acclimatization effect at the transcriptome level, but with the combination of expression and annotation patterns, we highlight good candidates of the persistence of adaptation (broad sense) to thermally varying environments.

2.4.4 Polymorphism pattern

The frequency of SNPs in *C. rubrum* was 1 per 126 bp. This frequency is higher than those found for *Nematostella vectensis* (1 per 139 bp; [304]), *Acropora millepora* (1 per 207 bp; [210]) and *Acropora palmata* (1 per 272 bp; [245]). However, much higher frequencies have been described for other marine organisms like *Pinctata margaritifera* (1 per 23 bp; [310]), *Crassostrea gigas* (1 per 40 bp in

non-coding regions and 1 per 60 bp in coding regions ; [283]), and more particularly for the octocoral *Eunicella cavolinii* (1 per 63 bp ; [268]). We chose to analyse the sequence polymorphism of our transcriptome to search for potential signals of genetic adaptation in these populations. Indeed, sequence polymorphism has already been shown to be an indicator of local adaptation for cnidarians by Lundgren *et al.* (2013) [197] who highlighted correlations between allelic frequencies and environment for two hexacorals, *Acropora millepora* and *Pocillopora damicornis*. The conservative filter of genotype quality $GQ = 20$ and $DP = 12$ allows us to be confident about the quality of the remaining SNPs. The significant decrease caused by the application of this filter can be explained by the fact that individuals have been pooled and allelic frequencies drifted away from the expected values, although the red coral is probably diploid based on the observed microsatellite patterns [185, 184]. However, our method of using pools of three individuals and the sequencing of expressed genes, did not enable us to estimate precisely allele frequencies and we focused on genes with major effects. Apart from the consequences of pooling, the analysis of six individuals per population probably induced an ascertainment bias leading to the non-identification of SNPs with low-frequency alleles. We therefore focused on genes with strong differences in allelic frequencies between populations, the apparently differentially fixed loci. The 56 differentially fixed SNPs detected in our study are potential markers for the study of local adaptation in *C. rubrum* but we cannot completely exclude other genes as potential markers of local adaptation. Indeed one could also expect adaptation to be linked to various interacting loci with small effects. All candidate genes presented here still need to be individually tested, but they open the way to deeper investigations of the genetic adaptation to local conditions through the analysis of the polymorphism of these genes on more individuals with additional ecological replicates. This could be extended to other cnidarians as well to search for patterns of repeated evolution.

2.5 Conclusion

This study enabled the identification of the first candidates for the study of adaptation (broad sense ; i.e. including genetic adaptation and acclimatization) to

contrasted environmental conditions for the red coral. The consistence of some of these markers with other studies implicating hexacorals is a strong argument for the conservation of their adaptive function to thermal stress among cnidarian, whether symbiotic or not. Our results also support the key importance of studying the gene expression in basal conditions. Finally, we propose here SNPs data for population genetic studies, as well as new resources for studying adaptation to contrasted environments in terms not only of expression pattern, but also of polymorphism.

2.6 Acknowledgements

This work is a contribution to the Labex OT-Med (n ° ANR-11-LABX-0061) funded by the French Government “Investissements d’Avenir” program of the French National Research Agency (ANR) through the A*MIDEX project (n ° ANR-11-IDEX-000102). This project has been funded by the ADACNI program of the French National Research Agency (ANR) (project n ° ANR-12-ADAP-0016 ; <http://adacni.imbe.fr>). We thank ECCOREV Research Federation (FR 3098) for the financial support of this study. We are grateful to the genotoul bioinformatics platform Toulouse Midi-Pyrénées. A particular thanks to the molecular biology staff of the IMBE laboratory and the Pythéas Institute’s divers for their support and assistance.

2.7 Appendices

Additional Supporting Information may be found in the Annexe B.

Fig. B.1. Size distribution of contig lengths (bp) for annotated contigs (red) and non annotated ones (blue).

Fig. B.2. Expression values for the two depths (FIG 5: Figuiet – 5 m ; MOR40: Morgiou – 40 m. Differentially expressed genes between the depth detected by each method are indicated (red).

Fig. B.3. Effect of the increase of the Genotype Quality parameter on the number of SNPs. The red vertical line indicate the threshold applied.

Table B.1. Best informative UniProt Match and for annotated differentially ex-

pressed contigs (contigs with non informative UniProt hits are not indicated).

Table B.2. Best informative UniProt Match and for annotated contigs containing differentially fixed SNPs (contigs with non informative UniProt hits are not indicated).

Chapitre 3

Identification de locus d'adaptation locale dans un contexte de forte structure génétique

Chapitre 3 : Identification de locus d'adaptation locale dans un contexte de forte structure génétique

Ce chapitre correspond à un article prochainement soumis dans *Molecular Ecology*.

Objectifs

Des différences de thermotolérance entre individus issus de différentes profondeurs et expérimentant des conditions thermiques contrastées ont été mises en évidence chez le corail rouge dans la région de Marseille. Le chapitre précédent a permis de confirmer que des différences d'expression de gènes sont maintenues après une période en jardin commun d'individus issus de profondeurs différentes alors qu'ils sont dans un état physiologique basal. De plus, sur certains locus, des gènes présentant des fréquences alléliques corrélées à l'environnement d'origine des individus ont été mis en évidence. Ainsi, il semble que l'acclimatation individuelle ne soit pas le seul mécanisme entrant en jeu pour expliquer les différences de thermotolérance observées, et que l'adaptation locale doive jouer un rôle. L'objectif est ici de confirmer ces observations en recherchant des signaux d'adaptation locale à la profondeur chez le corail rouge dans une étude moléculaire dédiée. De plus, le corail rouge étant soumis à des régimes thermiques similaires dans différentes régions de Méditerranée, permettant de répliquer les paires de profondeurs d'études dans des clusters génétiques indépendants, il est intéressant de savoir si la réponse à une pression environnementale similaire est basée sur des mécanismes génétiques identiques ou non. La mise en place de l'adaptation locale est extrêmement dépendante des caractéristiques génétiques des populations. Il s'agit ainsi dans ce chapitre, de caractériser génétiquement le plus précisément possible des populations de corail rouge issues de deux profondeurs différentes afin de comprendre comment les processus évolutifs neutres et adaptatifs peuvent expliquer les relations et les différences qu'entretiennent ces populations.

Résultats

Trois clusters génétiques ont été observés chez le corail rouge dans notre zone d'échantillonnage, suivant les régions géographiques étudiées (Marseille, Corse, Ba-

nyuls). Les niveaux élevés de structure génétique observés entre les populations, même géographiquement proches, confirment les résultats des études antérieures quand aux capacités de dispersions extrêmement limitées chez les corail rouge. Enfin, la forte structure génétique observée entre les populations de surface suggère une connectivité réduite entre ces populations qui pourrait limiter leur recolonisation suite à des épisodes de mortalité ou à une extinction locale.

En milieu marin, des marqueurs d'adaptation locale ont été mis en évidence à travers des études de *genome scan* uniquement chez des espèces présentant une structure génétique très faible à grande distance. Chez ces espèces, l'importance de la différence entre le niveau de différenciation des marqueurs sous sélection et celui des marqueurs neutres est propice à la détection de ces marqueurs sous sélection via des méthodes de F_{ST} outliers. Chez le corail rouge, au contraire, nous avons été confrontés à plusieurs obstacles méthodologiques du fait du niveau moyen de différenciation observé entre les populations analysées. Chez cette espèce, les résultats indiquent que les processus neutres seuls ont très probablement généré une forte différenciation et une variance élevée de la valeur des F_{ST} , conduisant à la détection de nombreux faux-positifs.

En prenant en compte ces limites, des marqueurs potentiels d'adaptation à la profondeur dans les populations étudiées de la région de Marseille ont pu être mis en évidence. Un signal d'adaptation locale est aussi observé dans les populations de Corse, sans qu'il soit aussi marqué que dans la région de Marseille, où les différences environnementales entre les populations de profondeurs différentes sont les plus fortes. Enfin à Banyuls, où ces différences environnementales sont les plus faibles, la présence d'un signal d'adaptation locale n'a pas pu être confirmé. Aucun SNP et aucune fonction en commun aux trois régions étudiées n'ont pu être mis en évidence. Ces résultats suggèrent que l'adaptation locale joue probablement un rôle dans les différences de thermotolérance observées entre les individus dans la régions de Marseille.

Identifying adaptive loci in a context of high genetic structure, local adaptation to depth in the red coral (*Corallium rubrum*)

Pratlong M.^{1,2}, Brener K.³, Mitta G.³, Toulza E.³, Haguenauer A.¹, Pontarotti P.² & Aurelle D.¹

Pratlong M.^{1,2}, Haguenauer A.¹, Brener K.³, Mitta G.³, Toulza E.³, Bensoussan N.⁴, Pontarotti P.^{2*} & Aurelle D.^{1*}

1. Aix Marseille Univ, Univ Avignon, CNRS, IRD, IMBE, Marseille, France

2. Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Equipe Evolution Biologique et Modélisation, Marseille, France

3. Perpignan Via Domitia Univ., IHPE UMR 5244, CNRS, IFREMER, Montpellier Univ., Perpignan, France

4.

* These authors jointly supervised this work.

Abstract

Genome scans are promising approaches to identify local adaptation targets in natural populations. In marine environment it has been applied to several species characterized by very low levels of genetic structure, and despite theoretical predictions on the evolution of local adaptation in case of strong genetic structure, no empirical example has been developed yet. The red coral (*Corallium rubrum*, Cnidaria) is a highly genetically structured species and a promising model for the study of local adaptation in a context of strong genetic structure. Here, we used RAD-Sequencing in order to identify signals of local adaptation to thermal regime in the red coral. We designed a sampling scheme of six pairs of 'shallow vs deep' populations in three geographical regions at a scale at which local adaptation may occur according to previous studies of thermotolerance in this species. We discuss first of horizontal and vertical neutral genetic structures of the red coral, then we highlight the methodological obstacles and biases expected in the detection of selected loci for such a strongly structured species. Finally we discussed of the robustness of the candidates for local adaptation detected in each geographical region.

3.1 Introduction

Different responses may explain the life and evolution in contrasted environments of species with large ecological range. The balance between environmental conditions, biological traits and evolutionary constraints will shape the output of the interaction between genetic and plastic responses for each species facing different environmental conditions. In the absence of other evolutionary pressures (migration, genetic drift), and if the environmental gradient is persistent for an extended period of time, each local population submitted to local selection could become genetically adapted to the corresponding local environment [166, 124]. In the case of plasticity or acclimatization, a given genotype can develop during its lifetime morphologic or physiologic responses allowing an increase in fitness in different environments [97, 241]. Although particular situations are documented

which favor local adaptation or acclimatization, it is often difficult to disentangle the effects of these two mechanisms which may evolve together [227]. However understanding the adaptive mechanism of adaptation has a fundamental interest but is also useful for management purposes [217, 124]. Gene flow is generally considered to be a key factor opposing local adaptation through the input in a population of potentially maladapted individuals (migration load) [187]. This consideration, based on the hypothesis of random gene flow, does not take into account the possibility of habitat choice in which individual can choose to settle in habitats that maximize their fitness [104, 259, 103, 306]. Furthermore, several theoretical studies have shown that gene flow can counteract the effect of genetic drift and promote local adaptation [145, 113, 293, 218, 3]. These results were confirmed in several studies in both marine and terrestrial environments, where signals of local adaptation have been detected in presence of high gene flow [72, 219, 236, 216, 124, 306]. Contrary to expectations, no studies highlighted local adaptation in marine species with very strong genetic structure, based on genome scan methods. Indeed, several studies have been conducted at very large scale on highly dispersive teleost species [40, 193, 332, 211, 27, 141], others focused on benthic invertebrates with a highly dispersive, planctonic larvae stage [66, 20, 7, 22]. From a methodological point of view, the impact of demography and structure of populations on the neutral distribution of F_{ST} is well known today and can lead to a high number of false positives in outliers tests for the detection of selection simply by the increase of the variance of F_{ST} values [31, 149]. This is true in particular when the average level of differentiation is high, because the variance in F_{ST} values among loci increases with average F_{ST} [149]. Furthermore, in a context of low gene flow, neutral markers are likely to be highly differentiated, and it could be difficult to identify selected markers with a higher differentiation than expected under the neutral model. Finally, if genetic drift is strong, it can generate outlier markers with unexpected levels of differentiation among populations, or with apparent correlation with an environmental variable outside any selective effect [166, 152, 73]. Marine coastal environments offer particularly interesting conditions for studies of local adaptation, because of the gradual changes in environmental conditions along coastline at small scale, the gradual vertical changes from shallow to deep water and the patchy distribution of contrasted habitats at large scale [276, 197, 340].

Recently, following worries on global change and the increase in mortality events potentially linked with human pressures [330, 262, 295, 99, 296], studies of local adaptation became more frequent in coastal ecosystems [322, 102, 294, 288, 18, 223, 36, 205, 17, 197, 172, 143, 348, 227, 20, 183, 242, 250, 160]. By studying the basis of the adaptation to different current environmental conditions one might expect understanding the possibility of adaptation to expected future conditions. However, few examples confirmed the role of local adaptation in observed physiological differences between individuals from contrasted environments and the role of individual acclimatization can rarely be excluded [17, 197, 20, 183, 160]. Genome scans are powerful approaches to explore adaptive processes in natural populations [200]. However, both methodological and theoretical biases have been highlighted in case of highly structured populations and the study of local adaptation in this case remain unexplored and challenging. The red coral (*Corallium rubrum*) is an asymbiotic temperate octocoral distributed from 5 to 800 m depth in the Mediterranean sea and the near Atlantic [79]. It is a long-living species (more than 100 years), with low growth and recruitments rates [202, 280]. The red coral is a particularly low disperser, with populations separated by fewer than 10 m detected as genetically differentiated [185, 184] but see [204]. The shallowest populations, above the seasonal thermocline, are submitted to high maximum temperatures and to frequent and intense thermal fluctuations in summer [143]. The intensity and frequency of extreme thermal events decrease with depth, and deepest populations inhabit more stable thermal regions. Only individuals below the maximum depth of seasonal thermocline are expected to be completely naive concerning heat stress. Since the observation of mass mortality events affecting this species during summer 1999 and 2003 and linked with thermal anomalies, the relationship between the red coral and its thermal environment has been intensively studied in the region of Marseille (France) [133, 131]. Common garden experiments performed in aquarium highlighted differences in polyp activities, calcification rate, necrosis response and expression of HSP70 between shallow and deep individuals facing thermal stress (comparisons 11 / 40 m in [316] and 5 / 20 / 40 m in [143]). Ledoux *et al.* (2015) conducted reciprocal transplant between individuals from 20 m and 40 m and observed a local adaptation signal with a response validating the “local vs. foreign criterion” in Marseille [166], but this trend was not observed in Corsica.

In this study, the authors performed common garden experiment by transplanting individuals from 20 m and 40 m at 5 m in Marseille, and they demonstrated that 40 m individuals were significantly more affected by necrosis than 20 m individuals after 95 days at 5 m. Fecundity measurements were conducted in situ in Marseille, and no difference of fecundity was found between individuals from 15-22 m and 39-42 m [314]. Finally, transcriptomes of individuals from 5 and 40 m in Marseille were compared after a common garden acclimatization and several genes were detected as differentially expressed in a basal state (without the application of any stress) [250]. Together, these studies highlighted thermotolerance (i.e. phenotypic) differences between individuals from different depths in Marseille, but we still do not know if these differences are the result of local adaptation or of individual acclimatization, or both. The red coral is then a promising species for the study of local adaptation via genome scans approaches in a case of strong genetic structure. Furthermore, previous works on this species allowed us to have a precise idea of the geographic scale at which local adaptation may occur, and to optimize our sampling design accordingly. Because populations from different regions may have evolved similar responses to thermal stress, through similar or different genetic basis, it is interesting to investigate local adaptation in pairs of 'shallow vs deep' populations submitted to contrasted thermal environment in geographical distinct regions [162, 149]. Finally, the study of neutral genetic structure is essential for a good interpretation of outliers tests for the detection of selection, considering the biases highlighted above. Here we applied Restriction site Associated DNA sequencing (RAD-seq) to individuals from different depths in three geographical regions of the Mediterranean sea corresponding to distinct genetic clusters. The goal of this study was to characterize the neutral and adaptive genomic variation in this species. Our results allow us to discuss first the neutral genetic structure of the red coral. In a second time, we highlight the methodological obstacles and biases expected in the detection of local adaptation in the red coral considering its genetic characteristics, and in relation with our empirical results. Finally, we discuss the robustness of the candidates of local adaptation detected in each geographical region.

TABLE 3.1 – Characteristics of red coral sampling sites.

Population	Geographic region	Site	Depth (m)	GPS	GPS
FIG8	Marseille	Marseille South	8	43 ° 12.330'N	5 ° 26.790'E
MOR40	Marseille	Marseille South	40	43 ° 12.060'N	5 ° 27.100'E
ELV12	Marseille	Marseille North	12	43 ° 19.780'N	5 ° 14.210'E
MEJ40	Marseille	Marseille North	40	43 ° 19.700'N	5 ° 13.480'E
BANN20	Banyuls	Banyuls North	25	42 ° 26.890'N	3 ° 10.330'E
BANN40	Banyuls	Banyuls North	35	42 ° 26.890'N	3 ° 10.330'E
BANS20	Banyuls	Banyuls South	26	42 ° 26.390'N	3 ° 10.790'E
BANS40	Banyuls	Banyuls South	36	42 ° 26.390'N	3 ° 10.790'E
POR20	Corsica	Porto	21	42 ° 16.292'N	8 ° 41.255'E
POR40	Corsica	Porto	33	42 ° 16.292'N	8 ° 41.255'E
GAL20	Corsica	Galeria	26	42 ° 28.210'N	8 ° 38.950'E
GAL40	Corsica	Galeria	36	42 ° 28.210'N	8 ° 38.950'E

3.2 Material and methods

3.2.1 Sampling and DNA extraction

Corallium rubrum colonies were collected by scuba diving at two depths of two sites in three geographical regions (Marseille, Banyuls, Corsica) between February and August 2013 (Fig. C.1, Table 3.1). The two depths of each sites presented contrasted thermal characteristic (Table 3.2). The three geographical regions presented different strengths of the temperature differences between the two studied depths: a difference of 3.8 ° C between the maximum observed at the two depths in Marseille, 1.7 ° C in Corsica and 0.5 ° C in Banyuls (Table 3.2). Thirty individuals per site and depth were collected (360 individuals), preserved in 95 % ethanol and stored at -20 ° C until DNA extraction. Total genomic DNA was extracted according to the protocol of Sambrook *et al.* (1989) [275], followed by a purification using Qiagen DNeasy blood and tissue spin columns (Qiagen). Genomic DNA concentration was quantified using a Qubit 2.0 Fluorometer (Life Technologies).

TABLE 3.2 – Temperatures characteristics of the red coral sampling sites from March 2012 to October 2014.

	Minimum	Maximum	Mean	Standard Deviation
BANN20	12.22	24.29	17.20	2.63
BANN40	9.41	23.83	14.49	2.45
BANS20	12.22	24.29	17.20	2.63
BANS40	9.41	23.83	14.49	2.45
ELV12	11.81	26.70	16.60	3.24
MEJ40	11.86	22.87	15.29	2.18
FIG8	12.63	26.92	17.03	3.52
MOR40	12.73	23.06	15.40	2.11
GAL20	12.46	25.09	17.13	3.13
GAL40	12.56	23.83	16.26	2.45
POR20	12.51	25.91	17.51	3.41
POR40	12.56	23.83	16.26	2.45

3.2.2 RAD-Sequencing

Twelve RAD libraries were prepared according to the protocol described in Etter *et al.* (2011) [107], with small modifications. Briefly, 1 μ g of genomic DNA for each sample was digested using high-fidelity PstI during 60 min at 37 °C. P1 adapters, with 4-6 bp individual barcodes were then ligated to each sample using 0.5 μ L of T4 DNA ligase (NEB), 0.5 μ L of rATP 100 mM (Promega), 1 μ L of DTT 500 mM (Promega), 1 μ L of 10X T4 ligase buffer (NEB) and incubated during 60 min at 22 °C, 10 min at 65 °C and 1 min at 64 °C. Pooled DNA samples were sheared, size selected and P2-barcoded. Final PCR for RAD-tags enrichment were performed with 16 cycles and primers dimers were removed during a final AMPure Beads Purification (Agencourt). Libraries were sequenced on an Illumina HiSeq2000 using 100 bp single-end reads, at the Biology Institute of Lille (IBL, UMR 8199 CNRS) and at the MGX sequencing platform in Montpellier (France). The STACKS pipeline [57, 58] was used for the loci de novo assembly and genotyping. Quality filtering and demultiplexing were performed with the *process_radtags* module with default parameters. Exact-matching RAD loci (putative orthologous tags) were individually assembled using *ustacks* with a minimum depth of coverage of five reads per allele ($m = 5$) and a maximum of five nucleotide mismatches

between allele ($M = 5$). These parameters were optimized during preliminary runs. *Cstacks* was used to build a catalog of consensus loci from all individuals, with five mismatches allowed between individuals at the same locus ($n = 5$). Matches of individual RAD loci to the catalog of loci were searched using *sstacks*. Finally, the *population* module was used to obtain the loci that were successfully genotyped in at least 75 % of individuals from all populations.

Because of the increased of error rates toward the end of the reads, we observed an increase in the number of SNPs from position 86 bp to 91 bp and we removed these positions from the analysis. In order to filter for poor-quality SNPs and artifacts due paralogous sequences, we used VCFtools [87] to remove SNPs that were not at the Hardy-Weinberg equilibrium within at least one of the 12 populations with a p-value threshold of 0.01. SNPs with a minor allele frequencies below 0.01 were removed using VCFTools. Individuals with more than 30 % of missing genotypes were discarded. Finally, only the first SNP of each RAD-tag was kept for further analysis.

3.2.3 Diversity and neutral genetic structure

Global estimated F_{IS} over alleles and gene diversity were estimated using GENEPOP and ARLEQUIN v.3.5 [271, 110]. The *C. rubrum* genetic structure was first analyzed by principal component analysis (PCA) using the package adegenet in R [161, 258]. This analysis was performed on the total dataset (12 populations) and inside each of the three studied geographical regions (four populations in each). The dataset was centered and missing data were replaced by the mean allele frequency for each locus. In a second step, we performed a Bayesian population clustering implemented in the program STRUCTURE v.2.3.4 [254, 111, 112, 156]. We performed ten independent replicates for $K = 1$ to 10 with a burn-in of 50 000 and a number of MCMC iterations after burn-in of 100 000, with the model allowing for admixture and correlated allele frequencies between clusters. We calculated the ΔK statistic of [108] to help in the choice of the most appropriate number of genetic clusters. We used CLUMPAK to summarize the STRUCTURE results from the ten independent runs [178]. The global and pairwise populations F_{ST} and exact tests for populations differentiation were calculated with GENEPOP [271]. The correlation between the spatial distance between the two depths of the same site

and the corresponding populations pairwise F_{ST} was tested with the correlation test of Spearman implemented in R [258]. Finally, we conducted an analysis of molecular variance (AMOVA) in ARLEQUIN v.3.5 [110] with 10 000 permutations. The hierarchy for this analysis was chosen to follow the three geographical regions of our samples (Marseille, Corsica and Banyuls). This choice was justified by the PCA on all dataset. Finally, we performed the PCA and F_{ST} calculation using a dataset comprising only putative neutral SNPs (without the SNPs detected as outliers by ARLEQUIN, see below).

3.2.4 Detection of local adaptation

In order to search for loci potentially involved in local adaptation, we first tested BAYESCENV [92]. This method identifies F_{ST} outlier loci that show a relationship between genetic differentiation and environmental differentiation. Runs were performed using default parameters, except the number of pilot runs which was set at 40. The maximal temperature recorded in each site was used as environmental variable (Table 3.2). The convergence of runs was checked with the Gelman and Rubin’s diagnostic using the R package coda [244]. Second, we searched for F_{ST} outliers among red coral populations using ARLEQUIN v.3.5 [152, 110]. Because hierarchical genetic structures are known to lead to a high number of false positives [152], we performed this analysis independently in the three geographical regions in order to down a level in the structure. With this method, a distribution of F_{ST} across loci as a function of heterozygosity between populations is obtained by performing simulations under a hierarchical island model which is more appropriate than the island model in the case of the genetic structure of the red coral (two depths in one site and two sites in one geographical region). Outliers were identified as loci being in the tails of the generated distribution ($p < 0.01$). Outliers detected by ARLEQUIN could be false positives or the result of a selective pressure independent of depth. Therefore, we selected among these candidates loci, those linked with depth differentiation by searching, inside each geographical regions, loci with significant differences in genotypic frequencies between depths ($p < 0.01$). We corrected the obtained p-values using a false discovery rate of 0.05 [23].

TABLE 3.3 – Counts of SNP loci after each step of filtering.

Step	Number of SNPs	Software
After assembly raw data	138 810	Stacks [57, 58]
Excluding loci not in within population HWE	86 520	VCFtools [87]
MAF 1%	56 844	VCFtools [87]
One SNPs per RAD-tag	27 461	

3.2.5 Functional annotation and enrichment tests

The RAD tags were aligned on the red coral transcriptome [250] using the Burrows-Wheeler Alignment Tool (BWA) [189]. Blast2GO was used for the annotation of resulting contigs and functional enrichments analysis [71]. First, a blastp was first performed on the NCBI nr database with a e-value threshold of 10^{-10} [4]. Then, Blast2GO retrieved GO terms associated with the obtained BLAST hits. Finally, in order to identify function potentially over-represented in outliers, we performed an enrichment analysis using a Fisher’s exact test corrected using a false discovery rate of 0.05 [23].

3.3 Results

3.3.1 RAD-Sequencing and genotyping

An average of 191 ± 21 millions of reads by library was obtained after sequencing. After the demultiplexing and cleaning processes of the STACKS’s *process_radtags* module, an average of 180 ± 22 millions of reads by library was obtained. From these reads, we were able to assemble 138 810 unique consensus RAD-tags present in at least 75 % of our 360 sequenced individuals for all populations. After all quality filter steps (Table 3.3), 27 461 SNPs were available. Finally, we removed six individuals presenting more than 30 % of missing data (one individual from the MEJ40 population, two from the BANN40 population, two from the GAL20 population and one from the GAL40 population). Our final dataset used for further analysis consisted in 359 individuals genotyped on 27 461 SNPs.

TABLE 3.4 – Measures of genetic diversity of the red coral populations based on 27 461 SNPs.

Population	F_{IS}	Gene diversity
BANN20	0.018	0.15
BANN40	0.012	0.15
BANS20	0.019	0.15
BANS40	0.065	0.13
ELV12	0.005	0.17
MEJ40	0.053	0.18
FIG8	0.005	0.18
MOR40	0.036	0.18
GAL20	0.013	0.09
GAL40	0.019	0.13
POR20	0.023	0.13
POR40	0.009	0.15

3.3.2 Genetic diversity

Multilocus values of the F_{IS} ranged between 0.005 (ELV12) and 0.065 (BANS40) (Table 3.4). Expected heterozygosity varied from 0.12 (GAL20) to 0.18 (all populations of Marseille) (Table 3.4). Populations of Marseille had higher values of expected heterozygosity than populations from Corsica and Banyuls ($p = 0.02$, Wilcoxon–Mann–Whitney test).

3.3.3 Population structure analysis

The clustering of individuals with respect to the first two principal components (PCA1 and PCA2) reflected the geographical and depth origin of the individuals (Fig. 3.1A). Individuals from Marseille and from Banyuls formed two clear and homogeneous clusters while individuals from the two sites of Corsica formed two different clusters with an important distance between them on the second axis. The first PCA axis (PCA1) explained 7.28 % of the total genotypic variance and separated individuals from Marseille from individuals from Banyuls and Corsica. The second axis (PCA2) explained 4.4 % of the total genotypic variance and separated individuals from the Porto site in Corsica from other individuals. Concerning PCA inside geographical regions, individuals from the two sites of Corsica and Marseille

(north and south) were separated along the first axis (13.41 % and 6.74 % of the total genetic variance respectively ; Fig. 3.1B and 3.1C). The second axis (2.77 % of the total genetic variance) separated populations from the two depths of the south site of Marseille (FIG-MOR), and individuals from the two depths of the north site (ELV-MEJ). In Corsica and Banyuls, no PCA axis showed evident association with depth. Individuals from the two depths of the Galeria (GAL) site of Corsica were separated along the second axis (4.02 % of the total genetic variance) but this was not the case for individuals from the two depths of the Porto site (POR). Finally, individuals from Banyuls showed much less structure than individuals from Marseille and Corsica (Fig. 3.1D). The first axis (3.11 % of the total genetic variance) separated shallow populations from the two sites (north and south). Markers of sex have been removed in the Banyuls PCA (Pratlong *et al.*, in prep). The PCA on the overall dataset and inside each geographical region were identical to those results when only putative neutral SNPs were considered (Fig. C.2). The ΔK criterion [108] indicated that $K = 2$ was the best number of clusters for the STRUCTURE analysis. We present here the results for $K = 2$ to $K = 4$ which captured the main information of the results. In all cases, all clusters corresponded the main geographical boundaries and the two depths of each sites were always in the same cluster (Fig. 3.2). For $K = 2$, a clear separation between the Marseille region and the Corsica/Banyuls regions was observed, confirming the separation of populations along the first PCA axis (Fig. 3.1). $K = 3$ separated the three geographical regions in 7/10 replicates, and the remaining replicates grouped either one or the other Corsican site with Banyuls populations (Fig. C.5). Finally, $K = 4$ separated the two Corsican sites.

The overall F_{ST} of the total dataset was 0.13. Pairwise F_{ST} values ranged from 0.01 (BANS20 vs BANS40, BANS20 vs BANN40 and BANN40 vs BANS40) to 0.24 (ELV12 vs GAL20 and FIG8 vs GAL20) (Table 3.5). Genetic differentiation was highly significant for all pairwise comparisons ($p < 0.001$), even for populations separated by 10 m ($F_{ST} = 0.022$ for BANN20 vs BANN40, $F_{ST} = 0.012$ for BANS20 vs BANS40 and $F_{ST} = 0.10$ for GAL20 vs GAL40). Considering the between-depths F_{ST} , high F_{ST} values can be observed for different loci and different samples comparisons (Fig. C.9). The average F_{ST} between the two depths of the same site was 0.04 in Marseille, 0.08 in Corsica and 0.02 in Banyuls (0.04 for the total

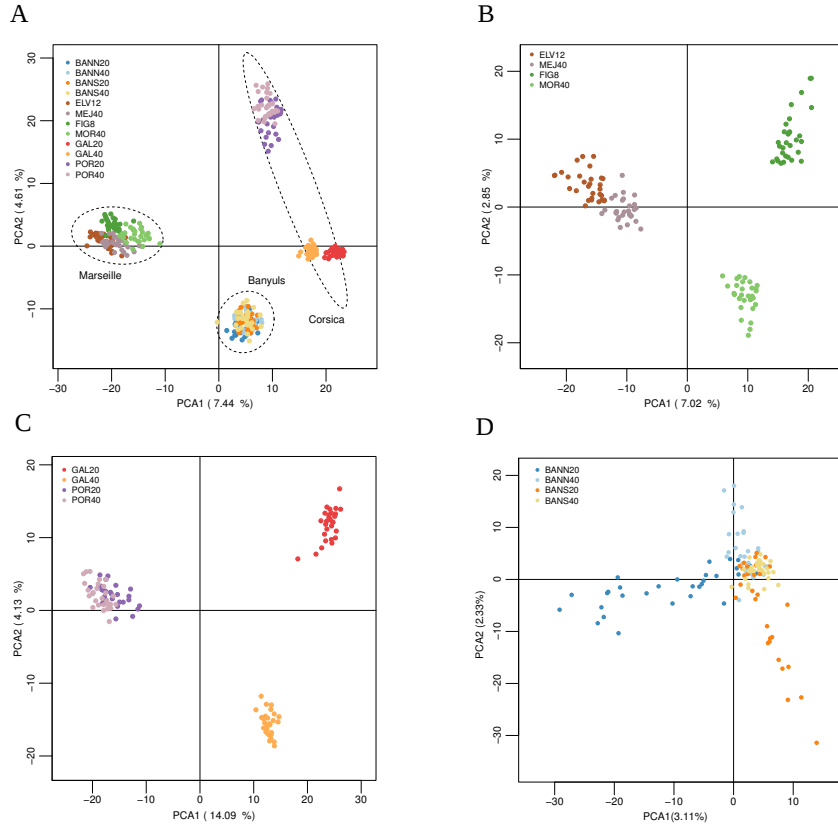


FIGURE 3.1 – Principal component analysis (Axes 1 and 2) of the **A.** 12 red coral populations ($n = 354$ individuals, 27 461 SNPs), **B.** four red coral populations from Marseille ($n = 119$ individuals, 27 461 SNPs), **C.** four red coral populations from Corsica ($n = 117$ individuals, 27 461 SNPs), **D.** four red coral populations from Banyuls ($n = 118$ individuals, 27 026 SNPs).

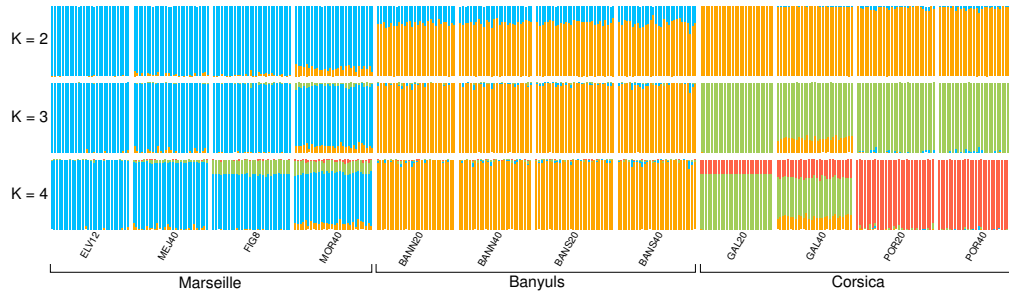


FIGURE 3.2 – Results from Bayesian individual clustering with STRUCTURE for $K = 2$ to $K = 4$. For $K = 2$ and $K = 4$, all ten replicates produced the same structure. For $K = 3$, the major mode presented here was the result of 7/10 replicates. Minor modes are presented in Figure C.5.

TABLE 3.5 – Pairwise F_{ST} estimates. All comparisons were highly significant. Intra-region comparisons are highlighted.

	BANN20	BANN40	BANS20	BANS40	ELV12	MEJ40	FIG8	MOR40	GAL20	GAL40	POR20
BANN20											
BANN40	0.02										
BANS20	0.03	0.01									
BANS40	0.03	0.01	0.01								
ELV12	0.13	0.13	0.13	0.13							
MEJ40	0.11	0.11	0.11	0.11	0.03						
FIG8	0.14	0.13	0.13	0.13	0.10	0.08					
MOR40	0.11	0.01	0.01	0.01	0.08	0.06	0.05				
GAL20	0.18	0.17	0.17	0.17	0.24	0.22	0.24	0.21			
GAL40	0.11	0.01	0.01	0.01	0.19	0.16	0.18	0.15	0.10		
POR20	0.14	0.13	0.13	0.13	0.17	0.16	0.17	0.14	0.20	0.13	
POR40	0.14	0.13	0.13	0.13	0.17	0.15	0.17	0.14	0.21	0.14	0.05

dataset). Considering only putative neutral loci (see below for outliers loci), the overall F_{ST} of the total dataset was 0.12 and pairwise F_{ST} values ranged from 0.01 (BANS20 vs BANS40, BANS20 vs BANN40 and BANN40 vs BANS40) to 0.23 (ELV12 vs GAL20) (Table C.1). There was no correlation between the distance between two depths of the same site and the corresponding population pairwise F_{ST} ($p = 0.75$). We obtained the same result if we removed the four populations of Marseille whose sampling sites for the two considered depths were not exactly the same (653 m between FIG8 and MOR40 and 995 m between ELV12 and MEJ40; $p = 1$). Finally, the F_{ST} between the two shallow sites inside a geographical region was in all three cases higher than those between the two deep sites of the same region (0.10 vs 0.058 in Marseille, 0.20 vs 0.14 in Corsica and 0.025 vs 0.014 in Banyuls, $p < 1.10^{-16}$ in all three comparisons). In a similar way, the F_{ST} between the two shallow sites of two geographical regions were in all cases higher than those between the two corresponding deep sites, except for the comparisons between the Porto sites and the Banyuls sites ($p = 0.38$ and $p = 0.02$ for the POR/BANN and POR/BANS comparisons respectively; $p < 1.10^{-16}$ for the other comparisons).

The AMOVA indicated a similar percentage of the molecular variance explicated by differences among group and within groups (7.8 and 7.07 %) and approximately 85 % of variance explicated by differences within populations (Table 3.6). There was significant genetic differentiation at the three studied levels ($F_{ST} = 0.15$,

TABLE 3.6 – Percent of the variation explained by grouping populations according to their geographical region on the analysis of molecular variance (performed with ARLEQUIN).

Source of variation	d.f.	Percentage of variation
Among groups	2	7.80
Among populations within groups	9	7.07
Within populations	696	85.13

$F_{SC} = 0.08$, $F_{CT} = 0.08$; $p < 0.001$ in the three cases).

3.3.4 SNPs outliers

We identified 82 outliers with BAYESCEENV. However, we noticed that all these outliers seemed to be driven by the divergence between populations as one allele was always fixed in one or several populations without logical association with depth. ARLEQUIN detected 563 locus potentially under selection in Marseille, 869 in Corsica and 397 in Banyuls. Among these SNPs, all were potentially under divergent selection in Marseille and Banyuls and 207 of the 869 were potentially under balanced selection in Corsica (the remaining 662 were potentially under divergent selection). Considering only these outliers, the overall F_{ST} of the total dataset was 0.25 and pairwise F_{ST} values ranged from 0.02 (BANS20 vs BANS40) to 0.42 (GAL20 vs POR40) (Table C.2). The 207 loci potentially under balanced selection in Corsica were linked with sex differentiation and were not further analyzed here (Pratlong *et al.* in prep). Eight outlier SNPs were detected both in Marseille and Banyuls, 12 both in Marseille and Corsica and 12 both in Corsica and Banyuls. No SNP was detected as potentially under divergent selection in the three regions. The complementary chi2 test of genotypic frequencies differences between depths inside each geographical region detected 162 candidates in Marseille, 1 371 in Corsica and 3 in Banyuls, of which 35, 248 and 2 where respectively common with the ARLEQUIN analysis. These numbers of candidates were correlated with the variance and the average of F_{ST} values inside each geographical regions (correlation coefficient of 0.97). Finally, the second axis of the PCA using the Marseille individuals showed apparent association with depth and appeared to be influenced by the variation of the candidates for local adaptation to depth:

51 % of these loci were in the top 1 % of the axis contributions, and 86 % were in the top 5 %.

3.3.5 Functional annotation

Among the 27 461 analyzed RAD-tags, 6 376 had hits on the red coral transcriptome (23.2 %). Concerning SNPs detected as outlier by ARLEQUIN and contributing to the depth divergence, 8 on the 35 detected in Marseille had hit on the transcriptome, 46 on the 248 detected in Corsica and 2 on the 2 detected in Banyuls (Table C.3). We observed no enrichment in coding regions among candidates SNPs. There was no functional enrichment in those outliers.

3.4 Discussion

3.4.1 Genetic diversity and structure

Our results confirm the high genetic structure of the red coral, which was already observed with microsatellite loci [77, 185]. These results could be the consequence of reduced mean larval dispersal distance (but see [204]). Our analysis of the genetic structure of the red coral revealed several clusters corresponding to the geographical distributions of individuals. This structure is close to the one evidenced by Ledoux *et al.* (2010) [185] by using microsatellites with similar sampling locations, and suggested that it may partly be the result of the absence of suitable habitats between these regions and of reduced dispersal (bathymetric barrier between Corsica and the continent and few shallow rocky substrates between Marseille and Banyuls). The relative proximity of the populations from Banyuls and the two populations of Porto (POR20 and POR40) according to the first component of our PCA was also observed in the STRUCTURE analysis of microsatellite data [185]. The high differentiation observed with PCA and STRUCTURE between the two sites of Corsica separated by around 20 km (but also for populations separated by around 15 km in the study of Ledoux *et al.* (2010) [185]) suggests the presence of a putative barrier to gene flow between Porto and the Punta Mucchilina in the south of the natural reserve of Scandola.

We reported here a significant vertical genetic structure between the two depths

of the same site (populations separated by less than 20 m). This fine-scale genetic structure has been repetitively reported for the red coral with microsatellites [185, 184, 79]. Contrary to Prada *et al.* (2008) [249] who showed the existence of two cryptic lineages at two different depths, the populations of red coral from the two studied depths clearly correspond here to the same genetic entity, as between depth differentiation was lower than differentiation between sites or regions. This structure may be the result of both inherent biological characteristics (negative-phototropism of larvae) and environmental factors. Indeed, the presence of a physical barrier or the stratification effect of seasonal thermocline during the larval emission (which could occur up to September in the Marseille region, Haguenaux A., pers. comm., [281]) could limit the connectivity between depths and conduct to a differentiation of both selected and neutral loci which may result in differences in vertical structure between regions [333, 270]. The strength of this genetic structure between two depths of the same site was indeed more or less pronounced according to the considered site, without correlation with the distance between the two depths. Thus, even if inherent biological characteristics (dispersal abilities, larval behavior) may have a significant impact on the genetic differentiation of such close populations, it seems that the differentiation between the two depths was influenced partly by local environmental constraints. The horizontal genetic differentiation between the two shallow sites being higher than those between the two deep sites, at the scale of a geographical region, but also for inter-region comparisons, suggested a higher connectivity or lower rate of genetic drift for deep populations compared to shallow ones. Low effective population size, repeated bottlenecks and founder events are known to lead to an impoverishment of gene diversity and higher genetic structure [48, 153, 10]. Although such effects could play a role in these red coral populations because of repeated colonization of shallow depths from deeper ones or following a preferential exploitation of shallow populations [270, 54], we did not observe here a reduction in gene diversity for shallow populations and connectivity differences seem to be more probable in these differences of genetic differentiation. This pattern was observed for comparable depths in the study of Costantini *et al.* (2011) [79] of the genetic structure and variability along a depth gradient in two western Mediterranean sites (Cap de Creus, Spain and Portofino, Italia) where the genetic differentiation between the two 30 m sites was higher than

those between the two 40 m sites. The lower connectivity at shallow depths could be related to the larvae photo-negative behavior which could prevent them to settle at the same depth than their origin habitat [333]. We reported here extremely low connectivity between shallow populations and deep populations of the same site as well as between shallow populations of distant sites. In case of shallow local mass mortality or over-exploitation, the ability of a population to escape local extinction will depend on horizontal and vertical recolonisation which seems to be, as we highlighted here both extremely limited at short term [185, 184, 79, 12].

3.4.2 Potential biases in the search of outlier loci

High F_{ST} values can be observed for different loci and different samples comparisons (Fig. C.9). As it is improbable that all these loci and comparisons are shaped by divergent selection (because adaptive polymorphism is expected to be quite rare at the genome scale), it indicates that neutral processes (i.e. mainly drift) alone generate high differentiation and high variance in F_{ST} in the red coral. Such patterns can then easily lead to false signals of divergent selection. Indeed whereas the red coral displays biological properties potentially favorable to the evolution of local adaptation (i.e. reduced dispersal limiting gene swamping; [187]), its genetic characteristics hinder the detection of adaptive loci: this species probably displays metapopulations with small, isolated, local populations. In each local population the reduced effective size could lead to important genetic drift [185]. Apart from leading to high neutral structure, genetic drift can also counteract the effects of local selection and could limit differences in allelic frequencies for low to moderately selected loci. Only alleles with strong effect might be selected here and could lead to an oligogenic architecture of adaptation. Moreover, the methods of detection of outlier loci used here are adapted to the detection of loci presenting high allele shifts between populations (adaptive traits controlling by few genes with large effects; [253, 124]). The high number of outliers detected in the three geographical regions by these methods raised thus concern of the biological pertinence of these candidates. Only a combination of different methods and a careful evaluation of the general structure and of the loci shaping differences between depths allowed to identify potentially selected loci.

Apart from the effects of genetic drift on the detection of local adaptation,

both the hierarchical genetic structure and the average level of differentiation of the red coral are known to lead to a high number of false positives and make the interpretation of results difficult [31, 149]. In this study, we set up a sampling design of pairs of geographically close 'shallow vs deep' populations in order to reduce signals from loci not under selection and to limit the effects of strong genetic structure on the detection of outliers (as recommended by Hoban *et al.* (2016) [149]). The correlation between the number of outliers detected and both the variance and average of F_{ST} values inside each geographical region raised however concerns concerning the potential presence of false positives remaining among our candidates. This is particularly true inside the Corsica region where the variance and the average of F_{ST} values were the highest.

Finally, the strength of differences in allele frequencies between populations from different depths as a result of divergent selection, should be correlated to the strength of divergent selective pressures between the two environments. In our case, temperature differences between the studied depths were higher in the Marseille sites than in the Corsican ones, and were very low in the Banyuls sites (Table 3.2). If, as hypothesized here, the local adaptation to depth in the red coral is driven by the thermal environment of individuals, we expected therefore to observe a stronger signal of selection in Marseille than in Corsica and Banyuls, as it is the case here. Thus, the number of candidates detected in the three geographical regions could be related to the rate of false positives which is correlated with the global genetic structure (see above) but also to the difference in selective pressure between the two studied depths. Unfortunately, in our study, the ecological and genetic distances were paired in such a way that we couldn't evaluate the respective impact of false positives and true positives in the number of outliers detected.

3.4.3 Local adaptation to depth in the red coral

By keeping in mind the aforementioned potential biases, we focused on candidates loci meeting the following criteria : i) detection with ARLEQUIN, ii) presence of a strong signal of differentiation between depth, iv) function congruent with the adaptation to thermal regime. These are the most promising loci for the study of genetic adaptation.

The absence of candidate SNPs or candidate function common to the three geo-

graphical regions, could indicate that the adaptation to comparable shallow environmental pressures in these independent regions may be based on different genetic pathways, or on no-genetic mechanisms [256]. However, the RAD-sequencing enable to sequence a reduced proportion of the genome and the probability to observe convergent positions is small, simply because they may not be sequenced. Furthermore, differences in the strength of selective pressure in the three regions could also explain the differences in the detected loci. In Marseille we evidenced a clear signal of differentiation between depths according to multivariate analysis, which was confirmed by outlier loci. This detection of a signal of local adaptation in the Marseille region is congruent with the observations of studies of thermotolerance differences in this region [316, 143, 183, 250]. Concerning Corsican populations, Ledoux *et al.* (2015) [183] reported no phenotypic signal of local adaptation after reciprocal transplant experiment. However, Ledoux *et al.* (2015) [183] studied a phenotypic trait, diameter growth, which may not be linked with adaptive abilities in all regions and genetic contexts. The most promising candidate for the adaptation to thermal regime detected in Corsica, is an homologous to an allene oxide synthase-lipoxygenase which is known to be involved in the response to thermal stress in octocorals [198]. Finally, the detection of only two candidates to local adaptation to depth in the Banyuls region was congruent with the selective pressure that should be weaker in this region (see above, Table 3.2). Moreover, the Banyuls populations were also the ones with the lowest genetic structure between populations from the two depths, a strongest effect of migration load can thus also limit the evolution of local adaptation and allelic differences on selected loci. The life in different environments can also rely on acclimatization. Until now, the role of local acclimatization has mostly been highlighted in thermotolerance differences studies in Cnidaria, probably because of the high plasticity abilities often observed in coral species together with the difficulty to demonstrate the impact of genetic adaptation [143]. Here, we showed that local adaptation seems to play a role in the observe thermotolerance differences between individuals separated by only several meters in the red coral. In this case, local adaptation could lead to different acclimatization abilities between populations as well. A similar pattern of correlation between gene sequences and thermal environment has been discovered by Bay & Palumbi (2014) [20] in *Acropora hyacinthus*, between non differentiated

sites of American Samoa but none of the functions highlighted were detected in our study. Pratlong *et al.* (2015) [250] conducted common garden experiment and transcriptome sequencing of individuals from two depths of a site in the Marseille region. We detected several genes differentially expressed and others presenting alleles differentially fixed between individuals from the two depths, but none of these genes were common with our analysis. However, differentially expressed genes may result from acclimatization differences between individuals linked with epigenetic differences, sequence polymorphism in a control region or related to a transcription factor as well as sequence polymorphism in the coding region. Without a sequenced genome, we couldn't infer the potential physical association between SNP and a gene, unless it is positioned in the coding region, which is rare. Indeed, and more generally, a sequenced genome, associated with the RAD-Seq data developed here, will enable to go further in the search of local adaptation in the red coral, especially with the identification of physical linkage between candidates loci [200].

3.5 Conclusion

To our knowledge, the red coral presented the highest level of differentiation among studies of local adaptation through genome scans approaches in marine environment [40, 193, 332, 66, 211, 20, 27, 7, 141, 22]. This study enabled us to empirically emphasize the limitations in the detection and the interpretation of signals of local adaptation using usual statistical methods in this strongly structured species. Both neutral and adaptive processes studied here highlighted the genetic singularity of shallow populations of the red coral, especially in the Marseille region where the known shallowest populations are found. Together, the strong genetic structure we observed between shallow populations, the low dispersal abilities of the red coral and the local adaptation of these individuals to the highly variable thermal conditions they experience raised strong concerns about the recolonization abilities of shallow populations and the possibility of loss of adaptive variations in case of local extinctions [134, 133, 135, 315].

3.6 Acknowledgements

This work is a contribution to the Labex OT-Med (n ° ANR-11-LABX-0061) funded by the French Government “Investissements d’Avenir” program of the French National Research Agency (ANR) through the A*MIDEX project (n ° ANR-11-IDEX-0001-02). This project has been funded by the ADACNI program of the French National Research Agency (ANR) (project n ° ANR-12-ADAP-0016 ; <http://adacni.imbe.fr>). We thank ECCOREV Research Federation (FR 3098) for the financial support of this study. We thank Nicolas Fernandez and Béatrice Loriod from the Marseille TGML platform for their precious help and advices with the RAD libraries preparation and Véronique Dhennin from the Biology Institute of Lille, the team of the MGX platform for the sequencing or the RAD libraries. We thank the molecular biology service of the IMBE. We want to thank the informatic service of the Pytheas Institute (especially Maurice Libes and Christophe Yohia) and Frédéric Zuberer of the Pytheas Institute for its support in sampling. We thank the Scandola Natural Reserve, especially Jean-Marie Dominici. We thank Manuela Carenzi for her help with statistical aspects of the detection of local adaptation. We thank Pierre-Alexandre Gagnaire, François Bonhomme and Nicolas Bierne for stimulating discussions.

Chapitre 4

Mise en évidence d'un déterminisme génétique du sexe chez le corail rouge

Chapitre 4 : Mise en évidence d'un déterminisme génétique du sexe chez le corail rouge

Ce chapitre correspond à un article prochainement soumis.

Objectifs

Chez les espèces gonochoriques, l'identité sexuelle des individus est définie par des mécanismes de détermination du sexe, allant d'une détermination purement génétique à une détermination purement environnementale. Alors que la détermination génétique du sexe est apparue indépendamment de façon répétée chez les organismes pluricellulaires, aucun exemple n'a encore été décrit chez les organismes non-bilatériens (Cnidaires, Eponges, Cténophores). L'objectif de ce chapitre est de rechercher l'existence d'une détermination génétique du sexe chez le corail rouge.

Résultats

Cette étude met en évidence deux types de marqueurs liés au sexe :

- des marqueurs pour lesquels les mâles sont très majoritairement hétérozygotes, et les femelles homozygotes, et qui suggèrent l'existence d'une région chromosomique présentant un blocage de la recombinaison entre mâles et femelles,
- des marqueurs présents chez les mâles et absents chez les femelles (présence et absence confirmées par PCR).

Les résultats obtenus par ces deux types de marqueurs sont cohérents et suggèrent un déterminisme génétique du sexe chez le corail rouge de type XX/XY. Trois individus présentent une différence entre leur sexe morphologique et leur sexe génétique et suggèrent la possibilité d'une réversion environnementale du sexe. Cette étude constitue la première description d'un déterminisme génétique du sexe chez un métazoaire non-bilatérien.

First evidence of genetic sex-determination in a non-bilaterian, the precious red coral

Pratlong M.^{1,2}, Haguenauer A.¹, Chenesseau S.¹, Brener K.³, Mitta G.³, Toulza E.³, Aurelle D.¹ & Pontarotti P.²

1. Aix Marseille Univ, Univ Avignon, CNRS, IRD, IMBE, Marseille, France

2. Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Equipe Evolution Biologique et Modélisation, Marseille, France

3. Perpignan Via Domitia Univ., IHPE UMR 5244, CNRS, IFREMER, Montpellier Univ., Perpignan, France

Abstract

Sexual reproduction is widespread among eucaryotes, but the sex-determining systems vary greatly among species [29]. In gonochoric species, the sexual identity of individuals is defined by sex-determination systems, going from purely genetic sex-determination (GSD), to purely environmental sex-determination (ESD) where the same genotype can produce both male and female phenotypes. While GSD has been intensively described in bilaterian species, no examples have so far been recorded among non-bilaterians [29, 14]. However, the quasi-ubiquitous repartition of GSD among multicellular species suggests that similar evolutionary forces repeatedly led to the evolution of GSD, and these forces could occur in non-bilaterians as well. Cnidarians display various sexual systems (from hermaphroditism to gonochorism) whose system determinations are poorly known: a few ESD examples have been reported but GSD has never been demonstrated in this group [292]. We tested the existence of sex-linked genes in the gonochoric precious red coral, *Corallium rubrum* (Cnidaria, Octocorallia) using Restriction site Associated DNA sequencing (RAD-Seq). We analyzed 27 461 Single Nucleotide Polymorphism (SNPs) in individuals that were morphologically sexed. We found strong association between the sex of individuals and the genotype of 472 SNPs which suggests a XX/XY sex-determination system, with a non-recombining XY-like region. This result was confirmed by the identification of 435 male-specific loci, markers of the putative Y chromosome. We developed a PCR-based protocol enabling the unambiguous identification of males. This is the first demonstration of GSD in a non-bilaterian species and a new example of its multiple convergences in multicellular organisms.

4.1 Results and Discussion

4.1.1 Morphological identification of sex

The red coral (*Corallium rubrum*) is an octocoral (Cnidaria) from the Mediterranean Sea and near Atlantic. Although hermaphrodite colonies have been

recorded in the first description of the reproduction of this species [90], it is nowadays considered to be gonochoric at colonies and polyps levels, as reported for the majority of octocorals [90, 281]. Red coral individuals from the two sexes are morphologically identical at the macroscopic scale, and the sex can be identified microscopically after dissection only during the period of gametogenesis (from May to September) [281]. We tested the existence of sex-linked genes in the red coral using Restriction site Associated DNA sequencing (RAD-Seq). We analyzed 27 461 SNPs in 354 individuals of 12 populations sampled from three geographical areas of the western Mediterranean Sea (Fig. D.1, Table D.1) between May and September. The sexing of 58 of these individuals from four populations in Corsica was undertaken on the basis of gonad analysis. Among them, 25 males and 28 females were formally identified (hereafter called morphological males and females). Five individuals were sexually undetermined because of the quality of the tissues or a sexually immature stage of development.

4.1.2 Identification of sex-linked loci

The presence of sex-linked loci (loci common to the two sex chromosomes but presenting differences in allele frequencies between sex; Fig. 4.1A) was explored by performing a principal component analysis (PCA). While the first four principal components highlighted neutral population genetic differences (Pratlong *et al.* in prep), the fifth axis (1.58 % of explained variance) separated individuals from all populations in two clear groups which matched the male/female repartition of morphologically sexed individuals (Fig. 4.2). Only three individuals, that were morphologically females, positioned among males in the PCA. From this PCA, we identified potential females and males among individuals for which sex-determination was not performed, on the basis of the visible separation in two groups. Considering this separation, we postulate that the dataset comprised 169 females and 185 males yielding to a balanced sex-ratio ($p = 0.40$, Table 4.1). The sex-ratio was balanced inside each population as well, except in ELV12 where it was significantly biased toward male individuals (70 % of males; $p = 0.03$, Table 4.1). There were 472 SNPs with a contribution to the fifth axis of the PCA higher than 1 % (Fig. D.2); these loci displayed the same allele fixed in almost all morphological females and were at the heterozygous state in almost all morphological males.

This observation was confirmed with PCA-identified males and females (Fig. 4.1B and 4.1C). The genotypes of these 472 SNPs allowed the sexing of 95 % of individuals (considering the three misclassified females). Among these SNPs, 379 were fixed in all morphological females, 55 were at the heterozygous state in all morphological males and 49 were in these two categories. These 49 markers were diagnostic of sex for the morphologically sexed individuals. Identification of such SNPs homozygous in females and heterozygous in males suggested a XX/XY sex-determination system in the red coral, with a non-recombining XY-like region. 347 of the SNPs leading the fifth PCA axis were fixed in all PCA-identified females. However, none of these SNPs were at the heterozygous state in 100 % of PCA-identified males. SNPs strictly homozygotes in females and heterozygotes in males were found inside each population and geographical region but none was common to the three geographical regions. The absence of markers diagnostic of sex in the overall dataset may indicate that these loci were submitted to a low but non-null recombination rate as it is the case when sexual chromosomes diverged recently or near the boundary of pseudoautosomal regions, where recombination is more frequent than in fully sex-linked regions [257, 26]. It could also be the result of polymorphism within the restriction sites on the Y chromosome (i.e. allele dropout) resulting in the impossibility of observing the Y allele with RAD-Seq [136]. Even if diagnostic markers of sex were not found at the scale of the overall dataset, the multilocus analysis allowed to clearly identify males and females. Sex-linked polymorphisms identified by RAD-Seq have been found in the pistachio (*Pistacia vera*) and the salmon louse (*Lepeophtheirus salmonis*), where authors identified markers being heterozygous in females and homozygous in males and suggested a ZW/ZZ system [163, 55], as well as in the Atlantic halibut (*Hippoglossus hippoglossus*), the Nile tilapia (*Oreochromis niloticus*) and the date palm (*Phoenix dactylifera*), where a XX/XY system was suggested [226, 65].

4.1.3 Identification of sex-specific loci

In order to identify potential sex-specific loci (Fig. 4.1A), we filtered the catalogs of RAD-tags to search among morphologically sexed individuals for loci present in all individuals from one sex (at least one sequence read by individual) and absent in all individuals from the other sex (no detected read). In order to

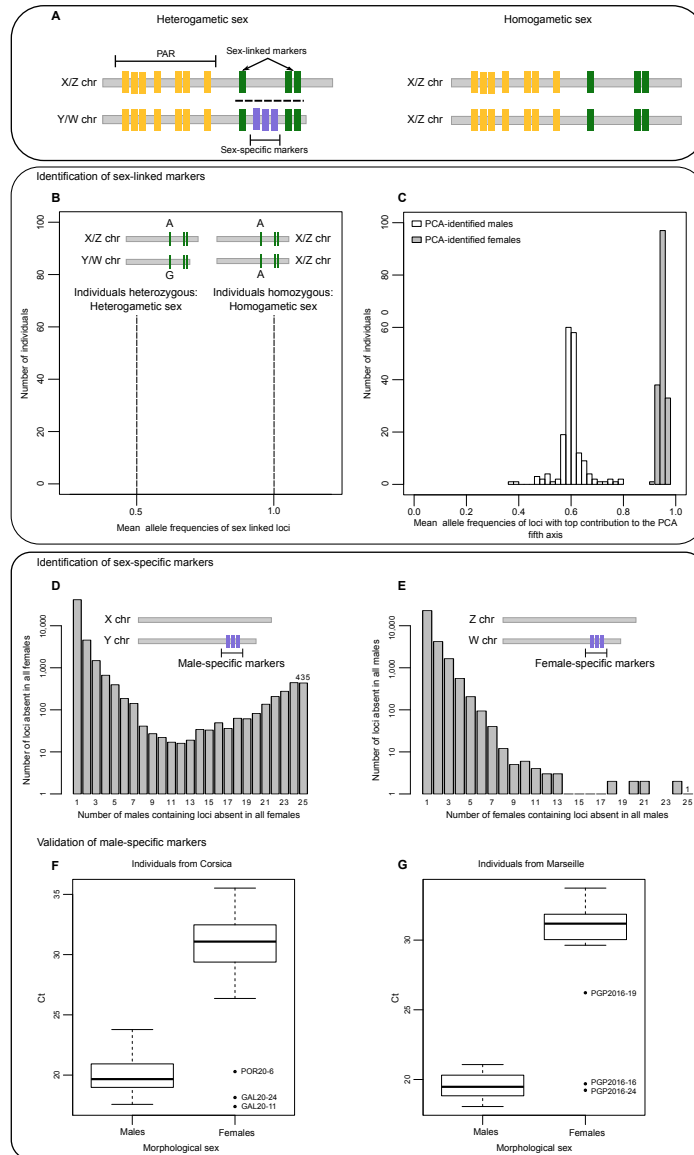


FIGURE 4.1 – **A**. Example of genetic sex-determination system. Sex-specific markers (blue) are present on the Y chromosomes in the XX/XY system and in the W chromosome in the ZW/ZZ system. Sex-linked markers (green) are common to the two sex chromosomes but should present differences in allele frequencies between sex because of the recombination arrest (adapted from [128]). **B**. Mean allele frequencies expected in case of sex-linked markers fixed on the X/Z and on the Y/W chromosomes. For those markers, individuals of the heterogametic sex (males in the XX/XY system, females in the ZW/ZZ system) should be heterozygous and individuals of the homogametic sex (females in the XX/XY system, males in the ZW/ZZ system) should be homozygous. **C**. Mean allele frequencies of loci with top contribution to the PCA fifth axis observed in PCA-identified males and females. **D-E**. Distribution of loci absent in all morphologically sexed individuals of one sex in morphologically sexed individuals of the opposite sex. **F**. Results of the real-time PCR amplification (threshold cycle Ct) of one male-specific locus (Locus_139082) on morphologically sexed males and females from the Corsica populations. Results for the six loci tested were similar (Fig. D.3). **G**. Result of the real-time PCR amplification (threshold, cycle Ct) of one male-specific locus (Locus_139082) on morphologically sexed males and females from the Marseille population. Results for the six loci tested were similar (Fig. D.4).

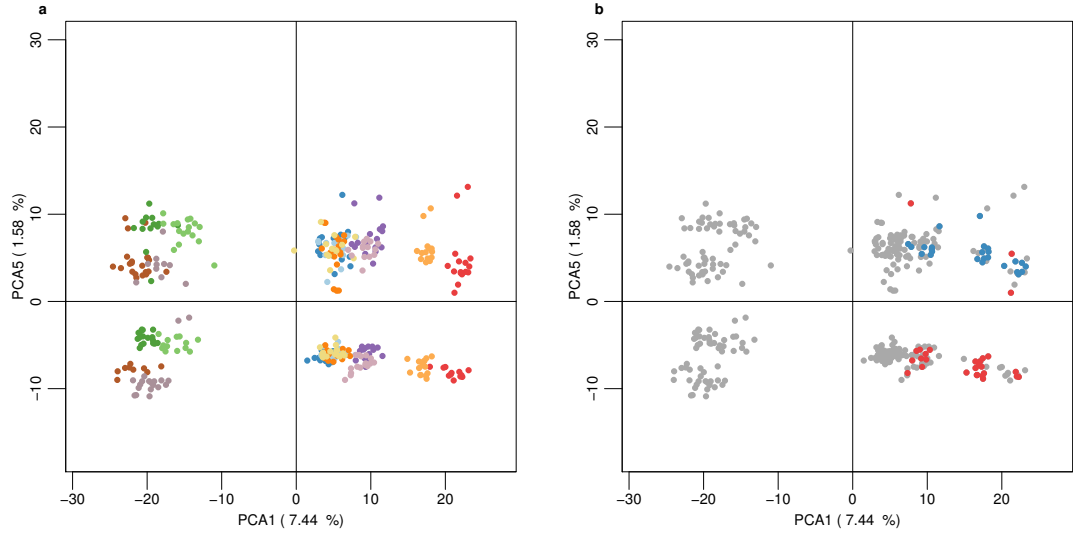


FIGURE 4.2 – Principal component analysis (Axes 1 and 5) of the 12 red coral populations ($n = 354$ individuals, 27 461 SNPs). Colors correspond to the **A.** population of individuals (light blue: BANN20, dark blue: BANN40, dark orange : BANS20, yellow: BANS40, brown: ELV12, grey: MEJ40, dark green: FIG8, light green: MOR40, red: GAL20, light orange: GAL40, dark purple: POR20, light purple: POR40; see Fig. D.1 and Table D.1), **B.** sex of individuals determined morphologically (red: females, blue: males).

avoid any bias caused by a potential misidentification of the sex of the three misclassified females, these individuals have not been taken into account for this analysis. 435 loci were present in all 25 morphological males and absent in the 28 females (Fig. 4.1D and 4.1E). In order to confirm the male specificity of the loci detected here, we targeted six of these 435 male-specific loci in real-time PCR of all morphologically sexed individuals. The six designed primer-pairs could be amplified in 100 % of morphologically sexed males. Almost no signal was observed in 100 % of morphologically sexed females (Fig. 4.1F, Fig. D.3; $p < 0.01$ in all tests, Table D.2). The presence of male-specific loci confirmed the previous results of sex-linked SNP and suggested a system with male heterogamety (i.e. XX/XY) [126, 127, 128]. The number of sex-specific markers should depend on the size of the non-recombining region and on the divergence between X and Y, from a minimal Y specific region for homomorphic sex chromosomes to several sex-specific markers in heteromorphic sex chromosomes [128]. The 435 male-specific loci detec-

TABLE 4.1 – Sex-ratio of red coral populations considering the PCA-identified males and females (see Fig. 4.2). Differences from balanced sex-ratio were tested by a chi2 test.

Population	Number of females	Number of males	p
BANN20	11	19	0.14
BANN40	15	13	0.71
BANS20	13	17	0.47
BANS40	17	13	0.47
ELV12	9	21	0.03
MEJ40	17	12	0.35
FIG8	17	13	0.47
MOR40	13	17	0.47
GAL20	12	16	0.45
GAL40	15	14	0.85
POR20	14	16	0.72
POR40	16	14	0.72
Total	169	185	0.40

ted here suggested at least 218 male-specific PstI sites which is far higher than the number of sex-specific loci in other studies based on RAD-Sequencing [127, 118]. However, considering the specificity of our analysis (size of the genome, frequency of the restriction site and restriction enzyme used), we could not directly compare the number of sex-linked markers (and the number of male-specific markers). Finally, we identified in the RAD-Sequencing catalog an homolog of a double-sex and mab3-related transcription factor (*Dmrt*). The *Dmrt* family of transcription factors is involved in sex-determination in numerous metazoans and the conservation of this function in cnidarians has recently been shown [63, 318]. In our case, this gene was neither a sex-specific marker nor a sex-linked marker, which stay compatible with a role in sexual differentiation. Finally, in order to confirm that the male-specific markers identified from only one geographical region were not the result of neutral genetic divergence between regions, we morphologically determined the sex of 40 additional individuals from a genetically different population, near Marseille. Among them, 15 males and 24 females were morphologically identified (one individual was sexually undetermined). The six male-specific primer pairs allowed to amplify 100 % of morphologically sexed males by real-time

PCR. Almost no signal was observed in 22 of the morphologically sexed females (Fig. 4.1G, Fig. D.4; $p < 0.01$ in all tests, Table D.2). Two females presented an amplification profile of the six markers similar to the males one, as it is the case for three of the females from the individuals from Corsica.

4.1.4 Cross-validation

Finally, as a validation test, we crossed the results obtained by PCA and by sex-specific loci by searching the presence of the 435 male-specific loci in all 354 individuals (Fig. 4.3). Twenty four percent of PCA-identified males possessed all 435 male-specific loci and 80 % possessed 90 % of these loci. The absence of some of these loci in some males may be due to their loss during library preparation and sequencing. Furthermore, the male-specific loci have been found from 25 individuals and it is likely that some of them have been detected erroneously if the pattern of technical missing data for a loci followed by chance the male/female distribution. All of the 435 male-specific loci were absent in 40 % of PCA-identified females, and 97 % of PCA-identified females contained less than 1 % of male-specific loci. Fig. 4.3 illustrates the correlation between the three methods of sex identification presented here (coordinate on the fifth axis of the PCA, number of male-specific loci and morphological identification). We observed an extremely good correlation between these three methods, the groups of males and females being well defined in each case. The three misclassified females identified from the PCA also possessed a high number of male-specific loci and seemed to be genetically males. Such incongruence between sex genotype and phenotype may result from an environmental sexual reversal during sexual differentiation [226], or from the existence of females with XY genotype as already described in several mammals [30, 282]. Finally, four individuals that have not been sexed morphologically presented female genotypes but between 15 and 65 male-specific loci. Further analysis is needed to determinate the sexual identity of these individuals.

4.2 Conclusion

In summary, both the exploration of sex-linked polymorphisms and sex-specific loci allowed us to identify a XX/XY genetic sex-determining system in the red

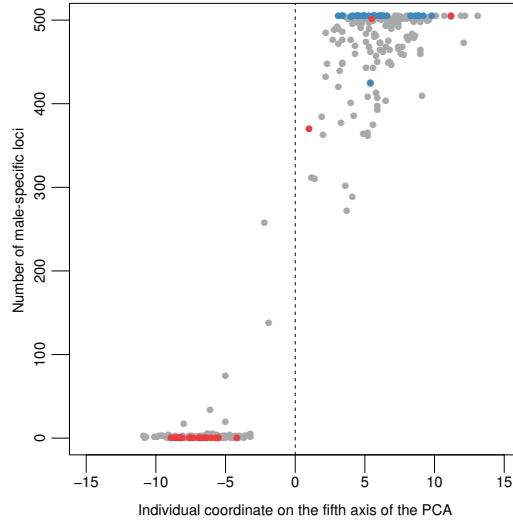


FIGURE 4.3 – Plot of the number of male-specific loci possessed by an individual as a function of the individual coordinate on the fifth axis of the PCA. Morphologically sexed individuals are indicated (red: females, blue: males).

coral. This is the first insight among non-bilaterian species and a new example of the convergence of GSD in multicellular organisms [25, 29]. As gonochorism predominates among octocorals [164], they constitute a promising group for further exploration of GSD among Cnidaria. Furthermore, these results pave the way towards a (re-)analysis of populations genomics datasets in the light of GSD in non-model organisms. Finally, we provided here real-time PCR primers that will allow to identify quickly red coral males and females, and facilitate the monitoring of the population dynamics for this emblematic species submitted to increased anthropic pressures, including harvesting and global change [42, 133, 252].

4.3 Experimental procedures

4.3.1 Sampling and DNA extraction

Corallium rubrum colonies were collected by scuba diving at two depths of two sites in three geographical regions (Marseille, Banyuls, Corsica) between February and August 2013 (Fig. D.1, Table D.1). Thirty individuals per site and depth were collected (360 individuals), preserved in 95 % ethanol and stored at -80 °C until DNA extraction. In order to validate our results, 40 additional individuals were

collected in one site near Marseille in June 2016 (43 ° 11.190'N, 5 ° 23.470'E) and conserved in both 95 % ethanol and formol. Total genomic DNA was extracted according to the protocol of Sambrook *et al.* (1989) [275], followed by a purification using Qiagen DNeasy blood and tissue spin columns (Qiagen). Genomic DNA concentration was quantified using a Qubit 2.0 Fluorometer (Life Technologies).

4.3.2 Morphological sex identification

Samples fixed in ethanol 95 % from the Corsica populations were decalcified in 10 % EDTA adjusted to pH 7.4 with NaOH for 48 h and dehydrated in graded alcohols, cleared in xylene substitute (Neo-clear VWR) and embedded in paraffin wax. Sections (7 μ M) were cut using a rotary microtome. Sections were stained using trichrome of Masson variant of Goldner protocols and examined using a Leica DMLB. The sex of individuals from the Marseille population collected in June 2016 and preserved in formol, was identified under a dissecting microscope after dissection.

4.3.3 RAD-Sequencing

Twelve RAD libraries were prepared according to the protocol described in Etter *et al.* (2011) [107], with small modifications. Briefly, 1 μ g of genomic DNA for each sample was digested using high-fidelity PstI during 60 min at 37 ° C. P1 adapters, with 4-6 bp individual barcodes were then ligated to each sample using 0.5 μ L of T4 DNA ligase (NEB), 0.5 μ L of rATP 100 mM (Promega), 1 μ L of DTT 500 mM (Promega), 1 μ L of 10X T4 ligase buffer (NEB) and incubated during 60 min at 22 ° C, 10 min at 65 ° C and 1 min at 64 ° C. Pooled DNA samples were sheared, size selected and P2-barcoded. Final PCR for RAD-tags enrichment were performed with 16 cycles and primers dimers were removed during a final AMPure Beads Purification (Agencourt). Libraries were sequenced on an Illumina HiSeq2000 using 100bp single-end reads, at the Biology Institute of Lille (IBL, UMR 8199 CNRS) and at the MGX sequencing platform in Montpellier (France). The STACKS pipeline [57, 58] was used for the loci de novo assembly and genotyping. Quality filtering and demultiplexing were performed with the *process_radtags* module with default parameters. Exact-matching RAD loci (putative orthologous

tags) were individually assembled using *ustacks* with a minimum depth of coverage of five reads per allele ($m=5$) and a maximum of five nucleotide mismatches between allele ($M=5$). These parameters were optimized during preliminary runs. *Cstacks* was used to build a catalog of consensus loci from all individuals, with five mismatches allowed between individuals at the same locus ($n=5$). Matches of individual RAD loci to the catalog of loci were searched using *sstacks*. Finally, the *population* module was used to obtain the loci that were successfully genotyped in at least 75 % of individuals from all populations. Because of the increased error rates toward the end of the reads, we observed an increase in the number of SNPs from position 86 bp to 91 bp and we removed these positions from the analysis. In order to filter for poor-quality SNPs and artifacts due to paralogous sequences, we used VCFtools [87] to remove SNPs that were not at the Hardy-Weinberg equilibrium within at least one of the 12 populations with a p-value threshold of 0.01 (Table D.3). SNPs with a minor allele frequencies below 0.01 were removed using VCFTools (Table D.3). Individuals with more than 30 % of missing genotypes were discarded. Finally, only the first SNP of each RAD-tag was kept for further analysis (Table D.3).

4.3.4 Identification of sex-linked loci

The presence of sex-linked polymorphisms was explored by performing a principal component analysis (PCA) using the package *ade4* in R [161, 258]. This analysis was performed on the total dataset (12 populations). The dataset was centered and missing data were replaced by the mean allele frequency for each locus.

4.3.5 Identification of sex-specific loci

We filtered the STACKS catalog in order to search among morphologically sexed individuals for loci present in all individuals from one sex (one read or more by individual) and absent in all individuals from the other sex (none read detected). In order to avoid any bias caused by an eventual misidentification of the sex of the three misclassified females, these individuals have not been taken into account for this analysis.

4.3.6 Real-time PCR

Primers were designed for 6 putative male-specific sequences using IDT online tool (<http://eu.idtdna.com/Primerquest/Home/Index>) (primers sequences in Table D.4). In order to avoid aspecific hybridization, we selected loci presenting no blast hit results other than themselves against the Stacks catalog. The presence or absence of target sequences was assessed by real-time PCR on 1 μ L of the same DNA extracts of the 58 morphologically sexed individuals that were used for RAD-sequencing experiments and those of the 39 morphologically sexed individuals from the Marseille population. Significance of amplification between males and females was calculated with Wilcoxon–Mann–Whitney test in R [258]. Real-time PCRs were performed on Eppendorf Master cycler, using Invitrogen platinumium SYBRGreen QPCR supermix (final volume 25 μ L, primers concentration 400 nM). PCR efficiency was checked with dilution range from 1 to 1 000, and specificity with melting curve analysis. PCR program consisted of initial enzyme activation (50 °C 2 min), denaturation (95 °C 2 min), 40 cycles (95 °C 15 s / 52 °C 30 s / 60 °C 30 s) and melting curve protocole. Ct values were measured for the same normalized fluorescence threshold between all experiments.

4.4 Authors contributions

D.A, P.P, A.H, G.M and E.T conceived the project. A.H, M.P performed library preparation. A.H performed the real-time PCR. S.C performed the sexual determination of individuals. M.P performed analysis. M.P wrote the paper. D.A and P.P supervised research. A.H, D.A, E.T, G.M, K.B, M.P, P.P, S.C discussed the results and contributed to the revision of the final manuscript.

4.5 Acknowledgments

This work is a contribution to the Labex OT-Med (n ° ANR-11-LABX-0061) funded by the French Government “Investissements d’Avenir” program of the French National Research Agency (ANR) through the A*MIDEX project (n ° ANR-11-IDEX-0001-02). This project has been funded by the ADACNI program of the French National Research Agency (ANR) (project n ° ANR-12-ADAP-0016 ;

<http://adacni.imbe.fr>). We thank ECCOREV Research Federation (FR 3098) for the financial support for this study. We thank Nicolas Fernandez and Béatrice Loriod from the Marseille TGML platform for their invaluable help and advice with the RAD libraries preparation, and Véronique Dhennin from the Biology Institute of Lille, the team of the MGX platform for the sequencing of the RAD libraries. We thank the molecular biology service of the IMBE and Frédéric Zuber of the Pytheas Institute for their support in sampling. We thank the Scandola Natural Reserve, especially Jean-Marie Dominici and thanks also go to Alexander Ereskovsky for his advice in red coral sexing.

Chapitre 5

Discussion générale

5.1 Résumé des résultats obtenus

L'objectif de cette thèse était de comprendre les origines des différences de thermotolérance observées chez le corail rouge. Le Chapitre 2 a mis en évidence que des différences d'expression de gènes étaient maintenues à moyen terme entre des individus issus de profondeurs différentes. Ce différentiel d'expression concernait certains gènes connus, et dont l'implication dans l'adaptation/acclimatation à la température chez d'autres cnidaires de l'ordre des scléactiniaires a déjà été décrite et suggère une possible conservation de certains mécanismes de réponse au stress thermique entre ces deux lignées très divergentes que sont les octocoralliaires et les hexacoralliaires (les scléactiniaires faisant partie des hexacoralliaires ; voir appendice A, voir Paragraphe 5.2). Cette surexpression “constitutive” hors d'une période de stress a été observée par Barshis *et al.* (2013) [17] chez le scléactinaire *Acropora hyacinthus* et confirmée par Palumbi *et al.* (2014) [227] qui parlent d'un effet fixé de la résistance au stress thermique (voir Paragraphe 5.3.1), qui pourrait être le résultat de l'adaptation locale ou de modifications épigénétiques (voir Paragraphe 5.3.2). L'identification de gènes présentant des SNPs différentiellement fixés entre profondeurs constitue un deuxième argument en faveur de l'adaptation locale. Ces fréquences alléliques calculées sur le transcriptome de 12 individus restent à être confirmées à l'échelle populationnelle (l'analyse des résultats d'un séquençage d'amplicons est en cours). Ce travail, bien que par certains aspects très préliminaire, a permis d'observer des différences physiologiques entre indivi-

dus issus de profondeurs différentes de manière plus large que les observations de nécrose ou d'expression de gènes candidats réalisées jusqu'à présent (voir Chapitre 1, [131, 133, 143, 183, 316]). Enfin, cette étude a confirmé que la distance écologique étudiée entre les populations de Marseille est suffisante pour observer des différences physiologiques entre individus et donc potentiellement favorable à la mise en place d'adaptation locale.

Le Chapitre 3 constitue le cœur de ce travail de thèse. Trois résultats principaux ont été obtenus et discutés dans cette étude. Le premier est une étude génomique de la structure génétique neutre verticale et horizontale du corail rouge. La structure génétique horizontale entre les trois régions étudiées ici a déjà été analysée via l'étude de marqueurs microsatellites. L'utilisation d'un nombre limité de marqueurs peut, dans certains cas, être insuffisante pour mettre en évidence une structure génétique complexe, et l'utilisation de marqueurs issus de séquençage haut-débit permet souvent d'analyser plus finement les patrons de structures génétiques [39]. Dans notre cas, nos résultats sont en accord avec les études microsatellites à l'exception de la relation entre les populations de Corse et de Banyuls. Notre échantillonnage de paires de populations 'surface vs profondeur' apporte en outre de nouvelles données sur la structure génétique verticale ainsi que sur la comparaison de la structure génétique horizontale aux deux profondeurs étudiées. Ainsi, nous avons pu identifier d'une part un flux de gènes réduit entre toutes les populations étudiées, mais surtout une différence de connectivité entre les populations de surface et les populations profondes de deux sites, pourtant séparés par la même distance géographique. Le second point de résultat porte sur la méthodologie de la recherche d'adaptation locale chez une espèce à forte structure génétique comme le corail rouge (voir Paragraphe 5.4). Par ailleurs, ce chapitre met en évidence des signaux d'adaptation locale chez le corail rouge qui sont confondus par des aspects fonctionnels ou par des patrons particuliers de structure génétique observés.

Enfin, le Chapitre 4 traite de l'identification d'un déterminisme génétique du sexe chez le corail rouge, du type XX/XY. Bien que ce sujet ne soit pas directement lié à la problématique générale de cette thèse et soit un résultat de l'ordre de la sérendipité, il nous permet de mieux comprendre la génétique du corail rouge et apporte de nouveaux éléments qui seront utiles pour la gestion des populations et leur conservation (voir Paragraphe 5.5). Il confirme également l'intérêt des oc-

tocoralliaires en général et du corail rouge en particulier en tant que modèle(s) en biologie évolutive et intégrative.

Notons ici que les analyses de recherche de marqueurs d'adaptation locale chez le corail rouge devaient être réalisées de façon similaire chez le scléractiniaire *Pocillopora damicornis*. Les premières analyses des données de RAD-Séquençage des échantillons de cette espèce ont mis en évidence la présence probable de plusieurs espèces cryptiques qui peuvent biaiser les études de connectivité [228]. Pour des raisons de cohérence, ces données n'ont pas été intégrées dans ce manuscrit, et feront l'objet d'une analyse dédiée.

5.2 Pertinence de l'interprétation fonctionnelle des gènes candidats pour l'adaptation locale

Cette thèse a mis en évidence, à travers différentes méthodes, plusieurs locus candidats, présentant des fonctions potentiellement pertinentes dans l'adaptation locale à la profondeur. Il faut séparer ici les locus détectés parce qu'il présentaient un différentiel d'expression entre les deux profondeurs en situation basale (Chapitre 2) et ceux détectés à travers l'analyse du polymorphisme de séquence (Chapitre 2 et 3). La surexpression constitutive des premiers types de marqueurs peut être le résultat de l'adaptation locale mais aussi d'un maintien à court ou moyen terme de l'acclimatation (voir Paragraphe 5.3.1). En ce qui concerne le deuxième type de candidats, le rôle de l'acclimatation est exclu, mais la présence de faux-positifs doit être envisagée. Dans tous les cas, l'identification de certaines fonctions très précises et déjà décrite chez des cnidaires nous permettent d'être plus confiants dans leur implication dans l'adaptation à la profondeur. C'est le cas des gènes de la famille des TRAFs (*Tumor Necrosis Factors Receptor Associated Factors*), et de la lectine dans le cas de la surexpression, ainsi que de l'allène oxide synthase-lipoxygenase dans le cas du polymorphisme de séquences [198].

Il faut cependant être prudent dans l'interprétation des fonctions des gènes candidats, notamment lorsqu'il n'y a pas d'enrichissement fonctionnel significatif. La stratégie de validation de vrais-positifs par leur interprétation fonctionnelle peut en effet conduire à des biais, étant donnée que la pléiotropie de nombreux gènes permet souvent de trouver une fonction pertinente dans le système étudié,

sans même invoquer un rôle de la sélection [232].

Finalement, de la même façon que les locus présents dans un gène présentant une fonction pertinente dans le système étudié ne sont pas forcément de vrais positifs, les locus n'étant pas dans un gène, ou pas dans un gène ayant une fonction connue, ne sont pas forcément de faux positifs :

- Ces locus, sans être directement sous sélection, peuvent être en déséquilibre de liaison avec des locus qui le sont. L'analyse de la fonction des locus détectés n'aurait donc, dans la majorité des cas, aucun sens biologique.
- Ces locus peuvent être non pas dans la région exprimée d'un gène mais dans sa région régulatrice.
- Ces locus peuvent être dans des gènes non annotés, ce qui peut être très fréquent chez le corail rouge. Les octocoralliaires sont en effet encore très peu étudiés au niveau génomique ou transcriptomique.

Notons ici que les deux premiers points pourront être explorés grâce l'apport d'un génome séquencé (voir Paragraphe 5.6.2). Finalement, les locus étudiés ici ont été *mappés* contre le transcriptome obtenu dans l'étude du Chapitre 2, en situation basale. Ainsi, la présence de marqueurs de stress dans nos candidats pour l'adaptation locale n'aurait probablement pas pu être identifiée, simplement parce que ces marqueurs ne sont pas dans ce transcriptome, car non ou trop faiblement exprimés. Les transcriptomes obtenus par Kelly Brener en situation de stress présenteront très probablement un contenu en gènes en partie différent, avec un enrichissement en marqueurs de stress. Il sera alors intéressant de *mapper* les locus candidats obtenus dans le Chapitre 3 contre ces nouveaux transcriptomes afin de potentiellement mieux caractériser certains de ces candidats.

5.3 Quelle(s) origine(s) aux différences d'expression de gènes observée(s) entre profondeurs ?

Plusieurs stratégies d'expression de résistance au stress thermique, via la mise en place d'un patron particulier d'expression de gène, peuvent expliquer les différences d'expression de gènes observées entre individus issus de profondeurs différentes.

5.3.1 Stratégies de résistance au stress thermique

Chez les coraux, deux formes de résistance à un stress thermique ont été identifiées : une surexpression constitutive d'un phénotype adaptatif ou la mise en place d'une plasticité plus élevée de l'expression de gènes. La mise en place de l'une ou l'autre de ces stratégies serait dépendante de la fréquence à laquelle les populations seraient soumises au stress : une surexpression constitutive sera favorisée à la plasticité si l'environnement fluctue plus rapidement que le temps de réponse moyen [169]. Notons ici que l'adaptation génétique et les modifications épigénétiques peuvent dans les deux cas être à l'origine de ces deux mécanismes. Une plasticité adaptative de l'expression de gènes a été décrite chez des populations de *Porites astreoides* présentant des variations de températures stressantes annuellement, de façon similaire aux populations de surface de corail rouge [169]. Au contraire, une surexpression constitutive de gènes de réponse au stress a été identifiée chez des populations de *Acropora hyacinthus* subissant des stress constants [17]. Ainsi, le corail rouge pourrait avoir évolué vers une plasticité phénotypique, comme cela a été envisagé par Haguenauer *et al.* (2013) [143]. Pourtant, l'observation de différentiel d'expression plusieurs mois après une période de stress est un argument en faveur d'une surexpression constitutive de certains gènes. On pourrait ainsi envisager que certains gènes, dont la surexpression ne présente pas de coût, soit surexprimés de façon constitutive alors que d'autres (dont l'HSP70 présenté par Haguenauer *et al.* (2013) [143]), de part le coût de leur expression, présenteraient plutôt une réponse de type plasticité adaptative. Certains de ces aspects pourraient être explorés à travers des séquençages de transcriptomes, avec un patron d'échantillonnage de paires de populations 'surface vs profondeur'. Ce type d'étude permettrait éventuellement d'observer des différences d'expression de gènes convergentes entre les deux environnements étudiés et apporter ainsi plus de puissance aux observations réalisées dans le Chapitre 2 de cette thèse.

5.3.2 Rôle des modifications épigénétiques

Chez le corail rouge, les différences de thermotolérance et d'expression de gènes entre individus issus de profondeurs différentes ont été décrites dans cette thèse comme étant le résultat de l'adaptation locale, ou du maintien à court ou moyen

terme de différences d’acclimatation. De nombreuses références traitent aujourd’hui de changements non génétiques, induits par l’environnement, stables au cours du temps et parfois héritables à travers des modifications épigénétiques. Ces modifications ne changent pas la séquence nucléotidique, mais la structure physique de l’ADN et sa disponibilité pour la machinerie transcriptionnelle [86]. Ainsi, ces modifications épigénétiques pourraient contribuer à la fois à l’acclimatation des individus au cours de leur vie, mais aussi à l’acclimatation des générations suivantes, et pourraient intervenir dans notre système d’étude. Des modifications épigénétiques associées à des différences de réponse à un stress ont été identifiées notamment chez des individus d’*Acropora millepora* dont la plasticité dans l’expression de gènes était corrélée avec le niveau de méthylation de l’ADN des lignées germinales [98], ou chez des individus de *Pocillopora damicornis* présentant une augmentation de la méthylation de l’ADN en réponse à une acidification [255]. Putman *et al.* (2015) [256] ont présenté pour la première fois un effet trans-générationnel d’une exposition à des températures élevées chez *Pocillopora damicornis*. Face à l’augmentation des pressions environnementales sur les coraux, et en relation avec ces résultats, certaines études proposent des protocoles d’évolution assistée à travers l’utilisation de mécanismes épigénétiques, par exemple en appliquant un stress à des individus, afin d’augmenter leur résistance à un stress environnemental futur [224, 255]. Chez le corail rouge, l’analyse d’acclimatation intra-générationnelle via la mise en place de marqueurs épigénétiques pourrait être mis en place, par exemple en appliquant des stress successifs à des individus en aquarium. Cependant, le cycle de vie du corail rouge ne pouvant pas actuellement être bouclé en aquarium, il constitue un mauvais modèle pour l’étude de la transmission de ces marqueurs entre générations. D’autres espèces modèles de cnidaires, comme *Clytia hemisphaerica* et *Nematostella vectensis* sont au contraire adaptées pour l’étude de la transmission de marqueurs épigénétiques et pourraient être utilisées pour l’étude du ‘souvenir’ d’un stress thermique au cours des générations [121].

Jusqu’à présent, il a souvent été suggéré que l’adaptation génétique des populations de coraux n’aurait pas le temps de se mettre en place face à la rapidité du changement climatique. Cette hypothèse ne prend cependant pas en compte l’acclimatation rapide, au cours de la vie des individus, et la transmission de marqueurs épigénétiques aux générations suivantes via l’hérédité non-génétique. Cette

dernière doit donc impérativement être prise en compte dans la prédiction de la réponse des coraux aux changements climatiques.

5.4 Obstacles méthodologiques à l'étude de l'adaptation locale chez le corail rouge

Les difficultés méthodologiques rencontrées lors de la recherche de locus impliqués dans l'adaptation locale ont été discutées dans le Chapitre 3. Cette discussion soulève une série d'arguments soutenant que les caractéristiques biologiques et démographiques du corail rouge sont particulièrement propices à la détection de faux-positifs dans la recherche de locus impliqués dans l'adaptation locale. Nous allons revenir ici uniquement sur le point concernant l'échelle d'échantillonnage, qui est un point central dans la discussion des problèmes rencontrés.

Comme cela a été indiqué dans le Chapitre 1, les études d'adaptation locale en milieu marin concernent essentiellement des espèces présentant une structure génétique très faible à grande distance géographique. Ces espèces sont en effet les meilleurs candidats pour l'étude de l'adaptation locale, notamment via des méthodes de recherche de locus outliers, du fait de la différence entre la différenciation génétique des locus neutres et des locus adaptatifs [219, 276, 124]. Il est ainsi primordial de construire le plan d'échantillonnage de façon à ce que l'échelle spatiale soit suffisamment réduite pour que la différenciation des locus neutres soit faible, mais assez importante pour observer des variations environnementales. Dans le cas de notre étude chez le corail rouge, les échelles de différenciations génétiques et environnementales sont quasiment identiques, ce qui augmente notamment la difficulté de différencier les marqueurs neutres des marqueurs sous sélection. Les niveaux de structure génétique du corail rouge ne permettent pas, à notre connaissance, de découpler différenciation génétique et environnementale, en ce qui concerne la variable étudiée ici (température). En effet la structure génétique verticale (mais aussi horizontale) neutre aboutit à des niveaux de différenciation déjà élevés susceptibles d'être confondus avec un signal d'adaptation locale. L'habitat pourrait être étudié plus finement afin d'identifier une variation environnementale pertinente à plus petite échelle, mais il n'existe aucune étude sur l'existence possible d'adaptation locale chez le corail rouge à de telles échelles.

5.5 Déterminisme génétique du sexe

Le déterminisme génétique du sexe est apparu indépendamment de nombreuses fois chez les eucaryotes. Les chromosomes sexuels évoluent à partir d'autosomes, par l'acquisition initiale d'un locus de déterminisme sexuel (Fig. 5.1A). L'apparition d'allèles sexuellement antagonistes (bénéfiques à un sexe mais diminuant la fitness de l'autre sexe) physiquement proches de ce locus conduit à une suppression de la recombinaison dans cette zone entre les chromosomes X et Y (cette première zone de suppression de la recombinaison est appelée Strate I) (Fig. 5.1B et 5.1C). Cette première zone de suppression de la recombinaison peut s'étendre par l'acquisition de nouveaux allèles sexuellement antagonistes (apparition de nouvelles states), conduisant à la différenciation des chromosomes sexuels (Fig. 5.1D à 5.1H).

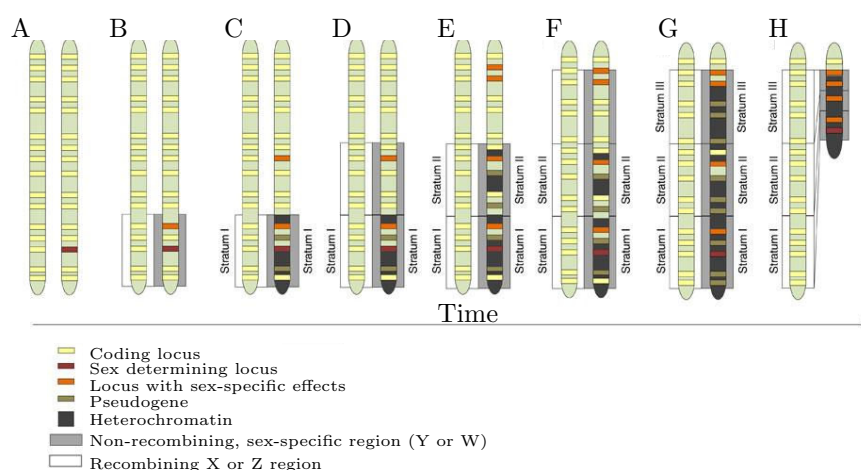


FIGURE 5.1 – Modèle théorique de différenciation des chromosomes sexuels. **A.** Acquisition d'un locus de déterminisme sexuel. **B.** Émergence de locus présentant des allèles sexuellement antagonistes. **C.** Apparition de la Strate I par suppression de la recombinaison. **D-G.** Expansion de la région de suppression de la recombinaison, conduisant à l'apparition de strates successives. **H.** Diminution de la taille du chromosome Y comme conséquence de la suppression de la recombinaison. Adapté de [341].

L'identification de marqueurs mâles-spécifiques et liés au sexe permet d'affirmer sans aucun doute qu'il existe un déterminisme génétique du sexe chez le corail rouge, mais la spécificité de notre analyse ne permet d'identifier plus précisément ni

le niveau de différenciation ou la longueur des chromosomes sexuels, ni les gènes de déterminisme sexuels impliqués. En effet, un nombre très faible de marqueurs mâle-spécifiques ont pu être *mappés* sur le transcriptome disponible, et leur fonction n'a pas pu être identifiée. Ce taux de locus rad mappant sur le transcriptome, inférieur à ce qu'on observe sur la totalité des locus rad, pourrait être le résultat d'une pseudogénisation des gènes du chromosome Y, mais pourrait aussi être le résultat d'un transcriptome développé uniquement sur des individus femelles. L'absence de polymorphisme dans le transcriptome au niveau des locus liés au sexe (i.e. présentant un allèle fixé chez les femelles) conforte cette hypothèse. L'expérience en aquarium et le séquençage des transcriptomes correspondants par Kelly Brener ont été réalisés sur des individus mâles et femelles, et l'apport de ces données permettront de préciser éventuellement la fonction des locus mâles-spécifiques, ainsi que d'analyser des différences éventuelles de réponse physiologique liées au sexe. Des analyses dédiées à la question du déterminisme génétique du sexe chez le corail rouge, notamment à travers un caryotype ou le séquençage de fragments plus longs de chromosomes sexuels, permettront de préciser le schéma observé ici.

Enfin, le gonochorisme a été décrit comme ancestral chez les octocoralliaires, mais aussi chez les anthozoaires en général [85, 171]. Il est majoritaire chez les octocoralliaires (89 % des octocoralliaires décrits sont gonochoriques [164]). Bien que déterminisme génétique du sexe et séparation des sexes soient deux mécanismes indépendants, la présence dominante de gonochorisme chez les octocoralliaires en fait un groupe d'intérêt pour l'étude de l'évolution du déterminisme génétique du sexe.

5.6 Perspectives

5.6.1 Liens avec le phénotype

Cette étude a permis de mettre en évidence des gènes candidats potentiellement impliqués dans l'adaptation locale à la température. Cependant, le lien entre ces locus et l'adaptation à la température reste ici théorique et leur relation avec le phénotype doit être confirmée. Cette relation pourrait être testée en analysant le différentiel de thermotolérance entre le génotype majoritairement observé à 20 m et celui majoritairement observé à 40 m pour ces locus d'intérêts. Ceci pourrait

être réalisé à travers des expériences de stress en aquarium en intra-population (afin de s'affranchir des effets de l'acclimatation à la profondeur d'origine) d'individus parallèlement génotypés. Des éléments préliminaires pourront être apportés par les expériences en aquarium et le séquençage de transcriptomes réalisés par Kelly Brener si les locus candidats pour l'adaptation locale peuvent être identifiés dans les transcriptomes obtenus, et que les génotypes d'intérêts sont présents par hasard dans l'expérience (avec cependant la limite du nombre d'individus analysés dans cette expérience). Une autre méthode de retour au phénotype consisterait à quantifier l'état physiologique d'individus *in situ* lors d'un événement de mortalité, et de le comparer au génotype de ces individus sur les locus candidats d'adaptation locale. Dans tous les cas, ces méthodes de retour au phénotype ne permettront pas de valider à 100 % le rôle direct des locus étudiés dans l'adaptation locale puisque le phénotype observé peut être le résultat de l'effet d'un ou plusieurs locus non identifiés en déséquilibre de liaison avec nos candidats.

Enfin, une estimation de l'héritabilité des traits qui pourraient être impliqués dans l'adaptation à la profondeur (par exemple la sensibilité au stress thermique) permettrait de préciser nos connaissances sur les bases génétiques de ces traits, et la part de plasticité phénotypique jouant un rôle dans leur variabilité en environnement naturel. La valeur de l'héritabilité est habituellement calculée à travers des études de pédigrée ou d'individus dont on connaît les relations de parenté qui peuvent difficilement être mises en place chez le corail rouge [301]. Des méthodes récentes permettent cependant d'estimer cette héritabilité grâce à la proportion de la variance phénotypique expliquée par les SNPs génotypés, dans des études de séquençage haut débit [301, 124]. Cette méthode, particulièrement pertinente dans notre cas d'étude, doit être associée à un phénotypage précis de traits pertinents, chez un grand nombre d'individus, qui semble difficile à mettre en place chez cette espèce.

5.6.2 Apport du génome du corail rouge

Comme cela a été mentionné dans la discussion du Chapitre 3, un génome de référence du corail rouge apporterait la possibilité d'une réanalyse et d'une réinterprétation complète des données de RAD-Séquençage [200]. L'association entre un génome et des données de SNPs issues de séquençage haut débit permet d'établir

la distribution des statistiques de génétique des populations le long du génome à la lumière de la relation physique qu'entretiennent les SNPs génotypés entre eux. Ainsi, des régions génomiques présentant des niveaux de différenciation entre populations significativement plus élevés ou plus faibles que le reste du génome, et potentiellement le résultat de sélection stabilisante ou diversifiante, peuvent être identifiées [154]. Ce type d'analyse des statistiques de génétique des populations, non pas en tant que données ponctuelles mais comme un continuum, apporte plus de puissance aux méthodes habituelles de détection de la sélection. Par exemple, une augmentation de la divergence a été observée autour du locus *EDA* (ectodysplasine), entre les populations d'eau douce et d'eau de mer chez l'épinoche [162] (Fig. 5.2). Dans ce cas, une substitution allélique sur le locus *EDA* a été associée à la perte de plaques latérales comme adaptation locale à la vie en eau douce [162]. Au contraire, une région du génome de l'épinoche présentant des niveaux de différenciation significativement plus faibles que le reste du génome, et associée à une hétérozygotie et une diversité nucléotidique plus élevées a été identifiée [154]. Les gènes de cette région, impliqués dans la réponse immunitaire, pourraient être la cible d'une sélection balancée [154] (Fig. 5.3). Un génome séquencé permet en outre d'identifier des zones présentant une diversité génétique faible pouvant être le résultat de balayage sélectif [200, 220].

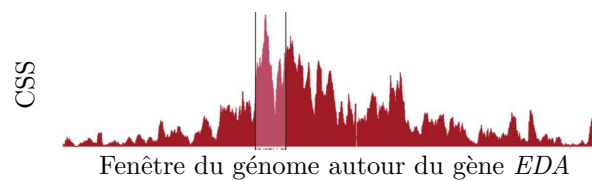


FIGURE 5.2 – CSS (statistique corrélée au F_{ST}) calculée entre les populations d'eau de mer et d'eau douce autour du gène *EDA* (ectodysplasine) (identifié par la zone éclaircie) [162].

Enfin, un génome, selon la qualité de la prédiction de gènes et de leur annotation, permet d'améliorer l'interprétation fonctionnelle des candidats détectés en identifiant les gènes codants qui leur sont le plus proche physiquement [200].

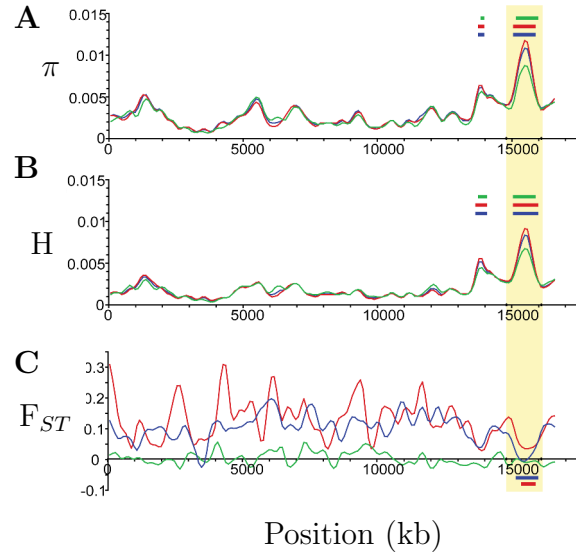


FIGURE 5.3 – Mesures de génétique des populations autour d’une région du génome de l’épinoche. **A.** Diversité nucléotidique (π), **B.** Hétérozygotie (H), **C.** Différenciation des populations (F_{ST}) entre les populations d’eau douce et d’eau de mer (bleu), d’eau douce (rouge) et d’eau de mer (vert). La région potentiellement sous sélection balancée est colorée en jaune [154].

5.6.3 Les autres gorgones méditerranéennes comme modèle d’étude ?

Nous avons discuté ici de l’adaptation locale chez le corail rouge en Méditerranée. Ce modèle a été choisi parce que les individus issus de profondeurs différentes présentent des différences de thermotolérance qui ont été très étudiées et bien caractérisées, mais aussi parce que le corail rouge est une espèce emblématique, exploitée et dont la conservation est une question d’actualité, notamment dans le cadre du changement climatique. Cependant, certaines caractéristiques génétiques du corail rouge ont soulevé des difficultés dans l’identification de marqueurs d’adaptation locale. D’autres espèces de gorgonaires, en particulier deux espèces du genre *Eunicella*, ayant une répartition similaire au corail rouge sur les gammes de profondeurs étudiées ici, présentent aussi des différences de thermotolérance, et des caractéristiques génétiques différentes du corail rouge. Ces espèces pourraient servir de modèles pour l’étude de la thermotolérance :

- *Eunicella cavolini* : les individus de surface sont plus thermotolérants que les individus plus profonds. Aucune structure génétique significative n’a été

observée entre les populations de profondeurs différentes à travers l'étude de neuf marqueurs microsatellites [242]. Des différences significatives ont été observées entre profondeurs pour l'intron de la Ferritine, mais pas pour celui du Facteur d'Induction de l'Apoptose (AIF) (Aurelle *et al.* submitted).

- *Eunicella singularis* : les individus profonds sont plus thermotolérants que les individus de surface. Cette espèce présente des symbiontes photosynthétiques du genre *Symbiodinium* qui pourraient expliquer cette inversion du patron de thermotolérance [114, 240]. Aucune structure génétique significative n'a été observée entre les populations issues de deux profondeurs différentes, que ce soit chez les gorgones ou le *Symbiodinium* [238]. Une étude récente a cependant mis en évidence une rupture au flux de gènes autour de 30 - 40 m entre deux écotypes liés à la profondeur [78].

Chez ces espèces, l'absence de structure génétique entre des populations pour lesquelles la distance écologique est suffisante pour la mise en place de différences de thermotolérance est particulièrement intéressante pour l'étude de l'adaptation locale (voir Paragraphe 5.4). La présence de *Symbiodinium* chez *Eunicella singularis* pourrait contribuer à complexifier un éventuel patron d'adaptation locale en superposant l'adaptation du symbionte avec l'adaptation de l'hôte, et *Eunicella cavonili* semble être le candidat idéal [155].

5.7 Conservation

Les études de génétique présentées ici chez une espèce aussi emblématique qu'est le corail rouge, en considérant en outre son importance économique et écologique, ne peuvent pas être dissociées d'une problématique de conservation. A travers la génétique des populations, il s'agit notamment d'identifier des populations isolées et génétiquement singulières, définir une protection adaptée au maintien de la connectivité entre populations, et envisager des scénarios possibles de l'évolution future des populations dans un contexte de changement global [228].

Dans l'hypothèse où les épisodes d'anomalies thermiques deviendraient plus fréquentes et plus intenses dans le futur, les populations de surface, bien que plus thermotolérantes que les populations profondes, seront les plus exposées au stress thermiques et aux épisodes de mortalité, comme cela a été le cas lors des épisodes

passés [133, 131]. Cette thèse a mis en évidence que ces populations de surface étaient peu connectées à la fois avec les populations profondes, mais aussi avec les autres populations de surface, et qu'elles étaient potentiellement génétiquement adaptées aux conditions thermiques de surface. Ce flux de gènes réduit pourrait limiter les possibilités de recolonisation à court terme d'une population suite à un épisode de mortalité partiel ou d'extinction locale, ainsi que l'apport d'un pool de variants potentiellement adaptatifs aux autres populations.

La Figure 5.4, réalisée par [157] (et reprise par [16]), présente les différents scénarios de prédiction des seuils de mortalité de coraux en relation avec le changement climatique. Du fait de la similarité de réponse au stress thermique entre le corail rouge et les scléractiniaires tropicaux (voir Chapitre 1), ces scénarios peuvent être transposés à notre cas d'étude. Selon le premier scénario, peu réaliste, tous les individus répondent de la même manière au stress thermique, et aucune adaptation ni acclimatation au changement climatique n'est observée (Fig. 5.4A). Les observations *in situ* des mortalités associées à des anomalies thermiques, ainsi que les expériences de stress thermiques en aquarium, nous indiquent au contraire que les seuils de thermotolérance sont différents selon les individus, en accord avec le deuxième scénario (Fig. 5.4B). Finalement, la mise en place d'acclimatation aux conditions environnementales locales, et/ou l'évolution de l'adaptation locale peut permettre d'augmenter ce seuil de thermotolérance au cours du temps (Fig. 5.4C). Les études réalisées chez le corail rouge, ainsi que les résultats obtenus dans cette thèse indiquent sans aucun doute que l'acclimatation et/ou l'adaptation locale interviennent dans certaines des différences de thermotolérance observées entre individus et pourraient donc jouer un rôle dans la persistance des populations de surface dans un contexte de changement climatique. Cependant, deux questions primordiales pour prédire le futur de ces populations restent ouvertes :

Une réponse évolutive peut-elle se mettre en place dans une échelle de temps cohérente avec la vitesse du changement climatique chez une espèce au temps de génération long tel que le corail rouge ?

Y-a-t-il une limite dans le seuil de thermotolérance qui peut être atteinte par l'acclimatation seule ?

Des éléments de réponse ont été donnés chez des scléractiniaires tropicaux dont certaines populations, présentes dans des régions thermiquement extrêmes, pour-

raient être les témoins de la réponse des coraux aux conditions environnementales futures. Par exemple, un rôle conjoint de l'acclimatation et de l'adaptation locale a été mis en évidence pour résister au stress thermique chez *Acropora hyacinthus* [227], l'association avec un symbionte photosynthétique thermorésistant a elle été décrite chez plusieurs espèces du golfe persique/arabique [158].

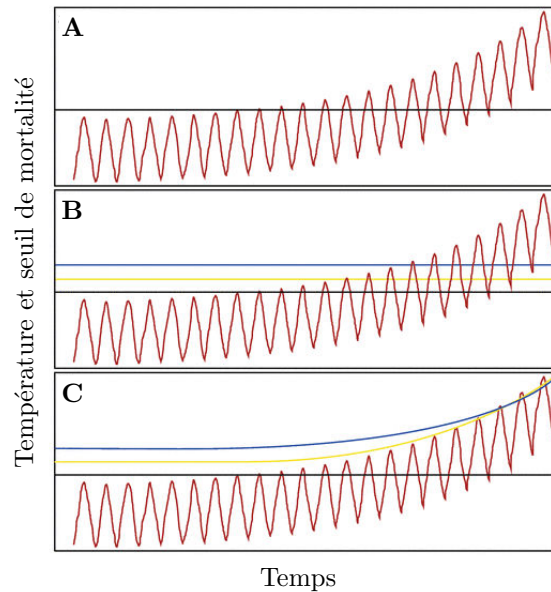


FIGURE 5.4 – Modèle schématisant différents scénarios du seuil de mortalité des coraux en fonction du réchauffement des océans. **A.** représente un seul seuil de mortalité pour tous les individus qui seraient alors excédés par l'augmentation de la température. **B.** représente la variation entre les populations et individus du seuil critique de thermotolérance. **C.** représente la capacité du seuil critique de thermotolérance de changer en réponse à l'environnement via l'acclimatation ou l'adaptation locale [157, 16].

Finalement, contrairement aux scléactiniaires tropicaux photosynthétiques présents uniquement en surface et dont la survie de l'espèce pourrait dépendre de ces capacités d'adaptation ou d'acclimatation, le corail rouge a été décrit jusqu'à des profondeurs pouvant atteindre 1 000 m, à l'abri des conséquences directes du changement climatique. Chez le corail rouge, les inquiétudes résident principalement dans la persistance à long terme des populations de surface, et des écosystèmes qui leurs sont associés.

5.8 Conclusion

Le corail rouge est une espèce emblématique qui est étudiée depuis des décennies à travers sa biologie, son écologie, et plus récemment sa génétique. Ce travail de thèse a permis d'apporter une pierre à cet édifice en précisant nos connaissances sur la génétique du corail rouge et en apportant des informations nouvelles, inattendues pour certaines, et confirmant les attendus théoriques pour d'autres. Trois points ont été explorés et contribuent ainsi à construire une vision intégrative de la génétique du corail rouge :

- la structure génétique du corail rouge dans l'ouest méditerranéen et notamment les relations verticales entre les populations,
- l'adaptation locale du corail rouge. Même si les analyses réalisées ne permettent pas d'identifier un patron clair d'adaptation locale, cette hypothèse restait à tester. Cette étude présente en outre les premiers résultats empiriques de l'application des méthodes usuelles de recherche d'adaptation locale chez une espèce à forte structure génétique,
- le déterminisme génétique du sexe chez le corail rouge.

Finalement, une grande partie de cette thèse a été dédiée à l'adaptation du protocole de RAD-Séquençage chez le corail rouge et *Pocillopora damicornis*. Les données de RAD-Séquençage développées peuvent encore être analysées sur certains aspects qui ne faisaient pas partie des problématiques étudiées ici. Une étude de l'histoire démographique des populations de corail rouge sera notamment réalisée dans le futur.

Annexe A

Phylogénie des anthozoaires (Cnidaires) : comment expliquer les différences entre phylogénies nucléaires et mitochondriales

Ce chapitre correspond à un article publié dans *Zoologica Scripta* (2016).

Monophyly of Anthozoa (Cnidaria): why do nuclear and mitochondrial phylogenies disagree?

Pratlong M.^{1,2}, Rancurel C.³, Pontarotti P.² & Aurelle D.¹

1. Aix Marseille Univ, Univ Avignon, CNRS, IRD, IMBE, Marseille, France

2. Aix Marseille Univ, CNRS, Centrale Marseille, I2M, Équipe Evolution Biologique et Modélisation, Marseille, France

3. INRA, Univ. Nice Sophia Antipolis, CNRS, UMR 1355-7254 Institut Sophia Agrobiotech

Abstract

The phylum Cnidaria is usually divided into five classes: Anthozoa, Cubozoa, Hydrozoa, Scyphozoa and Staurozoa. The class Anthozoa is subdivided into two sub-classes: Hexacorallia and Octocorallia. Morphological and molecular studies based on nuclear rDNA and recent phylogenomic studies support the monophyly of Anthozoa. On the other hand, molecular studies based on mitochondrial markers, including two recent studies based on mitogenomic data, supported the paraphyly of Anthozoa, and positioned Octocorallia as sister group to Medusozoa (the monophyletic group of Cubozoa, Hydrozoa and Scyphozoa). On the basis of 51 nuclear orthologs from four hexacorallians, four octocorallians, two hydrozoans and one scyphozoan (with poriferans and Homo sapiens as outgroups), we built a multilocus alignment of 9 873 amino acids, which aimed at minimizing missing data and hidden paralogy, in order to understand the discrepancy between nuclear and mitochondrial phylogenies. Our phylogenetic analyses strongly supported the monophyly of Anthozoa. We compared the level of substitution saturation between our dataset, the datasets of two recent phylogenomic studies and one of a mitogenomic study. We found that mitochondrial DNA is more saturated than nuclear DNA at all the phylogenetic levels studied. Our results emphasize the need for a good evaluation of phylogenetic signal.

A.1 Introduction

The phylum Cnidaria is usually divided into five classes: Anthozoa, Cubozoa, Hydrozoa, Scyphozoa and Staurozoa [69, 84]. Historically, Octocorallia (soft corals, sea fans, sea pens and gorgonians) and Hexacorallia (hard corals or Scleractinia and sea anemones) were considered as sister groups, forming the clade of Anthozoa (Fig. A.1A ; [28, 43, 59, 67, 84, 173, 209, 222, 274, 338]. The free-swimming medusa stage is absent in the Anthozoa clade, with no known exceptions. The life cycle of anthozoans is comprised of a sessile, reproductive, polyp stage with a larval stage. The remaining four classes belong to the clade of Medusozoa where the alternance between a sessile polyp stage and a free-swimming medusa stage is generally considered as an apomorphy lost several times [70, 84]. Recently, a phylogenomic study placed the clade of Myxozoa, a group of obligate endoparasites, as sister group to Medusozoa with good support [61]. Understanding the phylogeny of Cnidaria is therefore of great interest to the study of the evolution of life cycles and life history traits in this group. The Anthozoa and Medusozoa clades are supported by nuclear phylogenies [28, 59, 61, 67, 173, 209, 222, 338, 347], by anatomical and life cycle features [43, 70, 85, 84, 201, 260, 274, 338]. However, several studies based on mitochondrial genes have shown the paraphyly of Anthozoa, and the monophyly of a group constituted by Medusozoa and Octocorallia [168, 167, 225, 229, 286] (Fig. A.1B), without however being able to reject the hypothesis of anthozoan monophyly in Kaya *et al.* (2013) [168]. This clade (Octocorallia + Medusozoa), is not supported by morphological phylogenies [338], and does not present evident synapomorphies. The results of Park *et al.* (2012) [229] and Kayal *et al.* (2013) [168] have been recently reanalyzed by Figueroa & Baco (2014) [115]. These authors observed that the level of saturation of mitochondrial sequences was high, with a significant loss of information in the case of the deepest phylogenetic relationships. They suggested using nuclear gene phylogenies to resolve the high level phylogenetic relationships inside Cnidaria. Recently, phylogenomic data clarified the question of relationships among Hexacorallia, Octocorallia and Anthozoa by showing the monophyly of Anthozoa [61, 347]. Nevertheless, we still lack studies dedicated to understanding the discrepancy between mitogenomics and phylogenomics studies. This discrepancy has been observed in other phylogenetic studies, including those of octocorals [5]. It is well known today that the phylogeny of

a specific gene or gene family does not necessary represent the phylogeny of the corresponding species because of incomplete lineage sorting [150], horizontal gene transfer [138], introgression [265], hidden paralogy or lack of phylogenetic information (because of long branch attraction or substitution saturation, for example ; [115]) [44, 125]. Today, Next Generation Sequencing methods allow to fill the gap between mitochondrial and nuclear gene sampling in cnidarian class-level phylogeny. It is now pertinent to use phylogenomics in order to better understand the origins of different phylogenetic signals between mitochondrial and nuclear markers (Fig. A.1). We used a method comparable to the one of Zapata *et al.* (2015) [347] to build a new phylogeny of the Cnidaria phylum with 51 nuclear genes for 11 cnidarian species in which we attempted to maximize our control on hidden paralogy. First, we confirmed the congruence of our phylogeny using a concatenated gene alignment with previous phylogenomics results. In a second step, we used the individual alignments to check if the alternative scenario (i.e. anthozoan paraphyly) was well supported at the individual gene level. Finally, we tried to understand the differences between nuclear and mitochondrial phylogenies by analyzing our dataset and the datasets of Kayal *et al.* (2013) [168], Chang *et al.* (2015) [61] and Zapata *et al.* (2015) [347] in light of substitution saturation which is a mechanism known to lead to incongruent phylogenies.

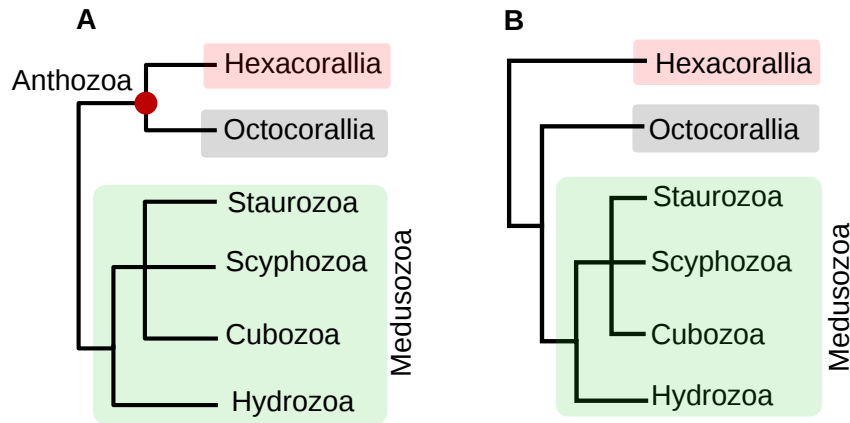


FIGURE A.1 – Hypotheses of the cnidarian phylogeny tested in this study. **A.** The monophyly of Anthozoa. **B.** The paraphyly of Anthozoa.

A.2 Material and Methods

A.2.1 Markers selection

Our strategy aimed at identifying a set of markers with the lowest levels of missing data and sequence saturation, as well as minimizing putative paralogs. Therefore we used stringent criteria in the choice of the markers. We used OrthoMCL [191] to search for potential ortholog reciprocal relationship between the available ESTs banks or transcriptomes of four hexacorallians (Cnidaria, Anthozoa) : *Acropora palmata* (order Scleractinia), *Pocillopora damicornis* (order Scleractinia), *Anemonia viridis* (order Actiniaria) and *Nematostella vectensis* (order Actiniaria); four octocorallians (Cnidaria, Anthozoa) : *Corallium rubrum* (order Alcyonacea), *Eunicella verrucosa* (order Alcyonacea), *Leptogorgia sarmentosa* (order Alcyonacea), *Gorgonia ventalina* (order Alcyonacea); two hydroidolins (Cnidaria, Hydrozoa) : *Clytia hemisphaerica* (order Leptothecata) and *Hydra vulgaris* (order Anthoathecata); one Discomedusa (Cnidaria, Scyphozoa) : *Aurelia aurita* (order Semaestomeae). We downloaded directly amino acid sequences from Ensembl and Compagen [83]. Concerning all remaining species, we downloaded nucleic acid sequences and found the longest open reading frame using the getorf EMBOSS tool [264]. The characteristics and references of each dataset are indicated in Table A.1. First, we manually chose orthologous groups containing at least one sequence of each species. Then, we aligned all the sequences for each orthologous group using MUSCLE with default parameters [105] and we built individual maximum likelihood trees using FASTTREE [251]. Concerning inparalogs (sequences of the same species forming a monophyletic group), we kept the sequence with the shortest branch length in order to select genes with the lowest evolution rate (and therefore minimize substitution saturation). In order to avoid the problem of unrecognized paralogy, we used PhyloTreePruner [177] that kept only orthologous groups forming a monophyletic group containing one sequence per species. At this step, we lacked an outgroup, which we did not include in the orthoMCL analysis for reasons of memory limitation. We performed a blastp [4] search of the sequences of each orthologous groups against the peptides banks of the sponges *Amphimedon queenslandica* (Porifera, Demospongiae, Heteroscleromorpha, order Haplosclerida), *Ephydatia muelleri* (Porifera, Demospongiae, Heteroscleromorpha, order Spongillida),

Sycon ciliatum (Porifera, Calcarea, Calcaronea, order Leucosolemida), *Oscarella carmela* (Porifera, Homoscleromorpha, order Homosclerophorida), and of *Homo sapiens* (Chordata, Mammalia, order Primates). These outgroups were chosen because they represent two main Metazoan lineages : sponges and bilaterians. Inside bilaterians we chose *H.sapiens*, for which a wealth of genetic data is available, in order to maximize the probability of identifying genes orthologous to those analyzed in Cnidarians. We used an e.value of 10^{-20} and kept all the hits. Finally, we performed the same procedure as before for the control of paralogy.

TABLE A.1 – Characteristics and references of each data set used in the phylogeny.

Phylum	Classification			Species	Source	Number of ESTs
	Class	Sub Class	Order			
Cnidaria	Anthozoa	Hexacorallia	Scleractinia	<i>Acropora palmata</i> <i>Pocillopora damicornis</i>	NCBI Vidal-Dupiol <i>et al.</i> (2013) [326]	43 150 72 890
			Actiniaria	<i>Anemonia viridis</i> <i>Nematostella vectensis</i>	NCBI – EST Bank NCBI – EST Bank	39 939 163 314
		Octocorallia	Alcyonacea	<i>Eunicella verrucosa</i> <i>Gorgonia ventalina</i> <i>Leptogorgia sarmentosa</i> <i>Corallium rubrum</i>	Romiguier <i>et al.</i> (2014) [268] Burge <i>et al.</i> (2013) [49] Romiguier <i>et al.</i> (2014) [268] Pratlong <i>et al.</i> (2015) [250]	15 461 90 230 13 299 48 074
				<i>Hydra vulgaris</i> <i>Clytia hemisphaerica</i>	NCBI – EST Bank NCBI – EST Bank	184 731 85 991
				<i>Aurelia aurita</i>	www.compagen.org	62 632
			Primates	<i>Homo sapiens</i>	Ensembl	99 436
Chordata	Mamalia	Discomedusae	Spongillidae	<i>Ephidatia muelleri</i> <i>Amphimedon queenslandica</i>	www.compagen.org NCBI	29 154 29 883
			Haplosclerida			
			Leucosolemida	<i>Sycon ciliatum</i>	www.compagen.org	50 731
Porifera	Calcarea	Calcarenea	Homosclerophorida	<i>Oscarella carmela</i>	www.compagen.org	29 220

A.2.2 Phylogenetic analyses

For each individual alignment, we selected conserved regions using Gblock [56] with the option allowing gaps for all positions. Next, we concatenated all individual alignments and obtained a final alignment of 9 873 positions (with 8.9 % of missing data; Appendix 1). ProtTest [1] was used to choose the model for protein evolution that best fitted our dataset. We performed Maximum-likelihood phylogenetic analyses of our final alignment using RaxML [299] with 1 000 fast bootstraps iterations and a PROTGAMMAILG model according to the ProtTest output. A Bayesian phylogenetic analysis was performed for the final concatenated alignment and for each individual alignment / gene using MrBayes [269] (number of generations: 200 000; sample frequencies: 100; burnin = 25 %; number of chains: 4; model of protein evolution: GAMMA-I-LG). Using the resulting tree files of the individual alignments analyses, we performed a Bayesian concordance analysis with Bucky [182] which estimates a concordance tree from clades supported by the largest proportions of genes. We then applied three commonly used tests to test phylogenetic hypotheses for output trees from RaxML: the Approximately Unbiased (AU) test [289], the Hasegawa (KH) test [176], and the Shimodaira-Hasegawa (SH) test [290]. We specifically tested the following hypotheses: 1: monophyly of Anthozoa, 2: monophyly of Octocorallia and Medusozoa, 3: monophyly of Hexacorallia and Medusozoa. All these tests were performed with the software CONSEL [291].

A.2.3 Substitution saturation analysis

Because we aimed to compare the levels of substitution saturation between datasets from different studies with different number of species, we first analyzed the effect of the number of species on two commonly used analyses of substitution saturation (Appendix 2). We showed that the statistical test of Steel *et al.* (1993) [302] implemented in DAMBE [344] was sensible to the number of species in the dataset and thus was not appropriate in the case of our comparative analysis. This was not the case for the slope of the regression of the plot of phylogenetic distances against observed distances. Therefore, we chose to keep only this last analysis of substitution saturation for the following analysis. We drew the plot of

TABLE A.2 – Number of contigs overexpressed at each depth and their average fold change

Characteristics				This study	Zapata <i>et al.</i> (2015)	Chang <i>et al.</i> (2015)	Kayal <i>et al.</i> (2013)
Missing data on complete dataset (%)				8.9	43	12	5
Number of amino acids				9 873	365 159	51 940	3 485
Number of sequences	Porifera			4	1	13	21
	Cnidaria	Anthozoa	Hexacorallia	4	9	10	30
			Octocorallia	4	4	2	19
		Meduzosoa		3	17	10	29

phylogenetic distances against observed distances for the individual alignments. Phylogenetic and observed distances were calculated with PATRISTIC [117] using the RaxML trees, and PAUP software [305] respectively. For each plot, we calculated the linear regression using R [309]. Additionally, we performed this analysis for the mitochondrial dataset of Kayal *et al.* (2013) [168]. For this second dataset, we built individual ML trees using RaxML and the PROTCATGTR model used in the original paper. In order to compare the nuclear and mitochondrial concatenated alignments, we performed a hierarchical substitution saturation analysis in a similar way than Figueroa & Baco (2014) [115] using our dataset and the datasets of Kayal *et al.* (2013) [168], Zapata *et al.* (2015) [347] and Chang *et al.* (2015) [61] (characteristics of these datasets are detailed in Table A.2). We plotted phylogenetic distances against observed distances for the concatenated alignments of these four studies using i) Cnidaria + Porifera datasets, ii) Cnidaria datasets, iii) Anthozoa datasets, iv) Hexacorallia datasets, v) Octocorallia datasets, vi) Medusozoa datasets. In each case, we calculated the linear regression using R [309]. Because divergent clades could mislead our comparison if they are not present in all considered studies, we removed Myxozoa from the dataset of Chang *et al.* (2015) [61] and Cerianths from the dataset of Kayal *et al.* (2013) [168] for this analysis. We used here original trees and alignments provided by the authors.

A.3 Results and discussion

A.3.1 Nuclear phylogeny of Cnidaria

We found 335 orthologous groups between the eleven Cnidaria species included in our orthoMCL analysis. From these 335 groups, 123 were kept after the PhyloTreePruner filtration, which reduced the inaccurate orthology assignment bias. Finally, 51 orthologous groups were retained after the addition of the out-group sequences. The Maximum Likelihood (ML) and Bayesian trees based on the concatenation of the sequences of these 51 genes were identical with the exception of the relationships between poriferans (Fig. A.2), Appendix 3). In the two cases, we found the monophyly of Cnidaria, Octocorallia, Hexacorallia and Medusozoa which were supported by most of the previously published phylogenies [28, 61, 62, 69, 84, 120, 168, 167, 229, 272, 297, 338, 347]. Regarding the relationships between the main cnidarian groups, the ML and Bayesian trees gave high support to the monophyly of Anthozoa (Fig. A.2, Appendix 3; ML bootstrap support of 99 and Bayesian posterior probability of 1), which confirmed the results obtained in other studies based on nuclear markers [59, 120, 222, 347]. The two alternative hypotheses (i.e. grouping Octocorallia – Medusozoa and grouping Medusozoa – Hexacorallia) were strongly rejected by all AU, KH and SH tests (Table A.3). However, our study did not take into account the possible paraphyly of Hexacorallia due to the uncertain position of Ceriantharia, which is placed either as sister taxon to Hexacorallia, to Octocorallia or to Anthozoa [300, 347]. Furthermore, these taxonomic results should be taken cautiously due to the limited taxa sampling and lack of any basal taxa at the level of the phylum, classes and subclasses. More than strictly resolving cnidarian phylogeny, we highlighted here the congruence of our phylogeny in order to explore the differences between individual gene based phylogenies as well as differences between nuclear and mitochondrial phylogenies.

The topology of the Bucky’s primary concordance tree (using individual – gene alignment – Bayesian phylogenies) was identical to the one resolved with the ML and Bayesian phylogenies based on concatenated sequences (Fig. A.2A). The monophyly of Anthozoa was supported here by 32 % [24 ; 41] of individual trees while

TABLE A.3 – Results of the tests of alternative topologies for the ML analyses (only P values are indicated).

Topology	AU	KH	WKH	SH	WSH
(Med, (Hex, Oct))	1	1	1	1	1
(Hex, (Med, Oct))	8.10^{-12}	0	0	0	0
(Oct, (Med, Hex))	7.10^{-7}	0	0	0	0

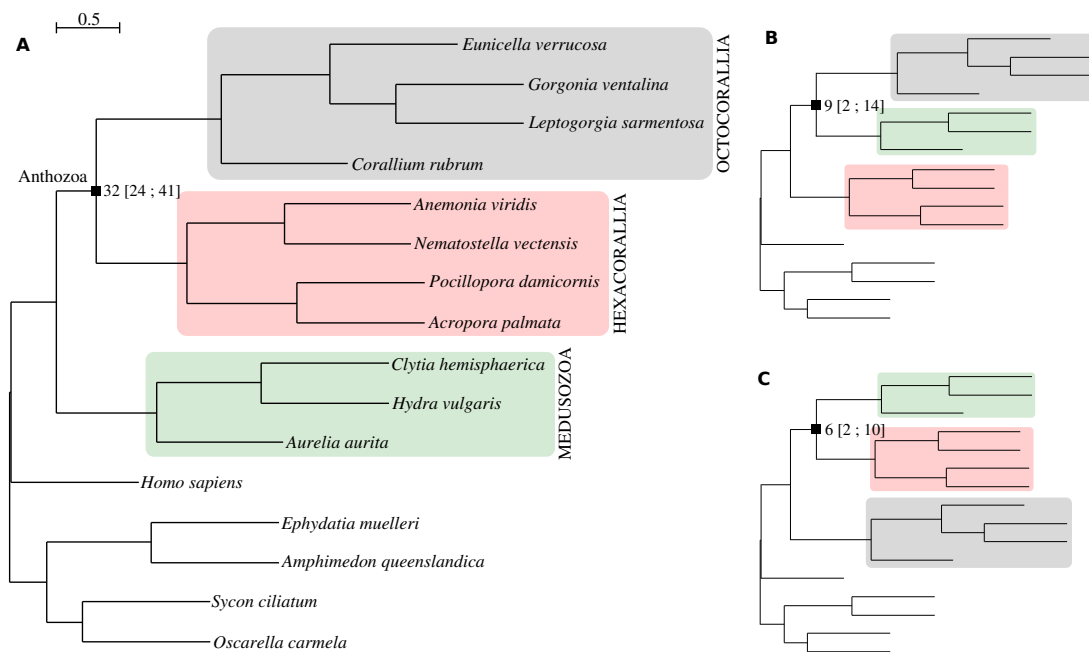


FIGURE A.2 – **A**. Bayesian primary concordance tree resulting from the analysis of MrBayes results (bayesian analysis) using the program Bucky [182]. Numbers on branches represent the average concordance factor. The 95 % confidence intervals are included. **B-C**. Results of the average concordance factors for the two alternative trees for the nodes of interest.

the two alternative hypotheses had lower support (9 % [2;14] for the monophyly of (Octocorallia – Medusozoa) and 6 % [2;10] for the monophyly of (Hexacorallia – Medusozoa)) (Fig. A.2**B-C**). Our results therefore confirmed other phylogenies obtained using nuclear markers and underlined the incongruence between mitochondrial and nuclear phylogenies. Moreover, this result raised limitations concerning the use of a limited number of genes in phylogenetic analyses. In our case, although

the main signal was congruent with the accepted cnidarian phylogeny, 15 % of our genes showed a different signal. Concerning nuclear markers, all concatenated and individual plots of patristic distance versus observed distance demonstrated substitution saturation signal with different strengths (Fig. A.3) : for small distances, the phylogenetic distances followed the observed distances but for larger distances, phylogenetic distances increased faster than observed distances. As a complementary test, we also built a new phylogeny using a concatenated alignment of the 20 % less saturated genes (i.e. those with the highest slopes). This phylogeny again supported the monophyly of Anthozoa with a bootstrap value of 84. Therefore our results based on nuclear genes do not seem to be strongly biased by saturation.

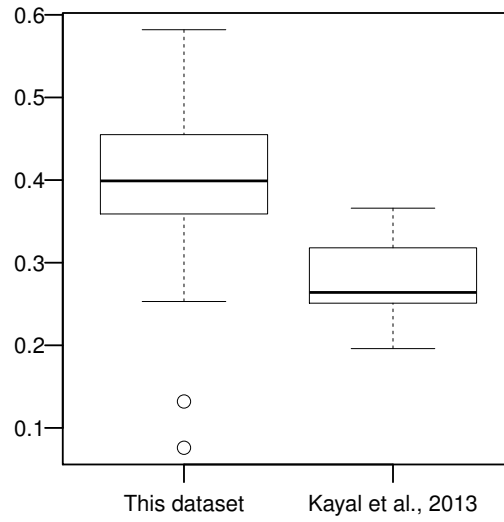


FIGURE A.3 – Boxplot of the slopes' distributions of the saturation plots for this dataset and the dataset of Kayal *et al.* (2013) [168].

A.3.2 Possible origin of the discrepancy between nuclear and mitochondrial phylogenies

Substitution saturation could be one of the mechanisms explaining the discrepancy between our results and previous results based on mitochondrial data [115]. To test this hypothesis, we compared the slope of the linear regression of the phylogenetic vs observed distances plots performed on hierarchical datasets (Fig. A.4, Appendix 4).

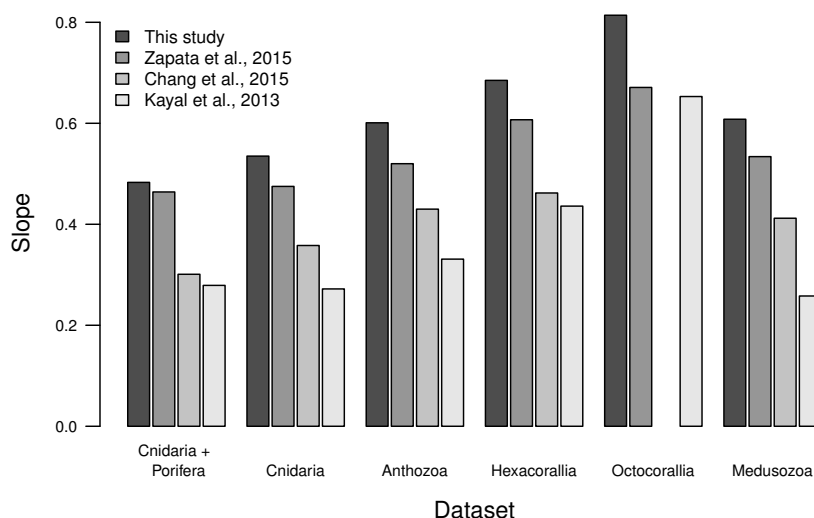


FIGURE A.4 – Slopes of the linear regression of the saturation plots from the hierarchical substitution saturation analysis.

A low slope is the signal of saturated sequences. Our hierarchical substitution saturation analysis showed that in all studies, the slope of the linear regression was low in the dataset Cnidaria + Porifera, increased in Cnidaria, Anthozoa, Hexacorallia, and Octocorallia where it reached its maximum. In Medusozoa, the slope was weaker than in Octocorallia and Hexacorallia. This result is congruent with divergent times proposed for Cnidarian clades: around 600 Ma for Medusozoa, 300 Ma for Octocorallia and Hexacorallia [229]. The difference in the saturation levels observed between Hexacorallia and Octocorallia, despite similar divergence times, could be due to a slower evolution rate in Octocorallia, for nuclear and mitochondrial markers. In all datasets, the slope was much higher in our study and in the study of Zapata *et al.* (2015) [347] than in the mitochondrial study of Kayal *et al.* (2013) [168]. This confirms that the mitochondrial sequences are globally more saturated than nuclear sequences in Cnidaria, with an increase of the level of saturation with taxonomic level. The slope of the regression in our study was always higher than the one in the study of Zapata *et al.* (2015) [347]. This may be the result of our conservative control of paralogy resulting in the selection of slowly evolving genes. The difference between the nuclear and mitochondrial slope was very small in Octocorallia. This clade seems to be the only one where mitochondrial sequences could be used with the same effectiveness as nuclear sequences. In

the case of the study of Chang *et al.* (2015) [61], the slope of the regression was also higher than for the mitochondrial study, but with much less difference than the two other nuclear studies. This could be the consequence of the presence in the concatenated alignment of Chang *et al.* (2015) [61] of 27 % of ribosomal and mitochondrial genes. As ribosomal genes have been reported to be more saturated than other nuclear genes and may contain a different phylogenetic signal [221], Chang *et al.* (2015) [61] state that their results was not influenced by these genes, but they still could induce a decrease of the saturation level of their dataset. Finally, we compared the slope of the linear regression of the phylogenetic vs observed distances plots between the individual genes, in our study and in the study of Kayal *et al.* (2013) [168], in order to check if some of the mitochondrial genes are less saturated than others. The 13 mitochondrial genes used in Kayal *et al.* (2013) [168] showed signal of substitution saturation, because of their relatively low slopes. The slopes of the linear regression of nuclear genes were significantly higher than those of mitochondrial genes ($p = 2.4 \cdot 10^{-08}$; Fig. A.3). This confirms that mitochondrial DNA is more saturated than nuclear DNA at the gene level in Cnidaria. The level of saturation of all mitochondrial genes is similar, with less variation than in nuclear genes. The gene COB (cytochrome b) was the least saturated with a slope of 0.366. When substitution saturation is high, similarity between sequences does not necessarily reflect phylogenetic relationships [302, 345, 343]. The results observed here suggest that mitochondrial sequences may not be appropriate to resolve deep cnidarian phylogeny, as suggested by Figueroa & Baco (2014) [115], Osigus *et al.* (2013) [225] and Van Iten *et al.* (2014) [323]. The paraphyly of Anthozoa, observed in several studies based on mitochondrial markers [168, 167, 225, 229, 286], could indeed be the result of substitution saturation. Here, on the basis on an extended gene dataset, we confirm the analysis of Figueroa & Baco (2014) [115], namely that nuclear markers are less affected by substitution saturation and more appropriate to resolve deep phylogeny of Cnidaria. Mitochondrial DNA is usually considered as poorly informative for anthozoan phylogeny at the species and genus level because of its slow evolution rate [64, 119, 207, 287]. Stampar *et al.* (2014) [300] argued that the most parsimonious scenario for this slow mtDNA evolution in Anthozoa is that the rate of evolution was high in the common ancestor of Cnidaria and Bilateria, and decreased in anthozoans. Because we highlighted here that mitochondrial

DNA seems to be much more saturated than nuclear DNA in Cnidaria, including Anthozoa, it seems indeed that the slow mitochondrial evolution rate appeared in the branch leading to Anthozoa.

A.4 Conclusion

The validation of this Cnidaria tree of life has major impact for the study of morphological and life history traits among Cnidaria. We argue here that nuclear markers should be favored for resolving cnidarian class-level phylogeny. More generally, our results confirm the caution concerning the use of a reduced number of markers for phylogenetic analyzes. Indeed, even within nuclear markers, different phylogenies were inferred depending on the gene. The high diversity of phylogenetic histories and levels of informations between genes therefore requires the use of enough markers and adequate statistical tests before validating a phylogenetic result.

A.5 Acknowledgements

This work is a contribution to the Labex OT-Med (n ° ANR-11-LABX-0061) funded by the French Government “Investissements d’Avenir” program of the French National Research Agency (ANR) through the A*MIDEX project (n ° ANR-11-IDEX-0001-02). This project has been funded by the ADACNI program of the French National Research Agency (ANR) (project n ° ANR-12-ADAP-0016 ; <http://adacni.imbe.fr>). We thank ECCOREV Research Federation (FR 3098) for the financial support of this study. The authors thank Xegen Company (www.xegen.fr) for their advice. We thank Anne Haguenaier for the development of *Corallium rubrum* transcriptome, Etienne Danchin for his collaboration for the OrthoMCL analysis and Abigail E. Cahill for her contribution to the correction of the manuscript.

A.6 Appendices

The following supplementary material is available :
Appendix 1. Amino acid alignment used in this study.

Appendix 2. Analysis of the sensitivity to the number of species of two markers of substitution saturation.

We tested here the sensitivity of our analysis of saturation to the number of sequences included in the dataset. For this purpose, we used the large dataset of Kayal *et al.* (2013) [168] in order to compare datasets with very different number of sequences. From this dataset, we randomly sampled five, ten and 19 species from the clades Porifera, Hexacorallia, Octocorallia and Meduzosoa, generating new datasets of 15, 50 and 95 species. We generated three replicates for each of the three datasets of different sizes. For each dataset, we first applied the statistical test of Steel *et al.* (1993) [302] implemented in DAMBE [344] to determine the number of phylogenetically informative sequences. The test of Steel *et al.* (1993) [302] calculate for each sequence a score representative of how the distribution of phylogenetically informative sites differs from what can be expected for random sequences (i.e. without phylogenetic information). Sequences with a ϕ score below 0.04 are considered as lacking phylogenetic information [343]. Then, we drew for each dataset the plot of phylogenetic distances against observed distances. Phylogenetic and observed distances were calculated with PATRISTIC [117] using the tree available in Kayal *et al.* (2013) [168], and PAUP software [305] respectively. For each plot, we calculated the linear regression using R [309]. For the two tests, we calculate the mean and the 95 % confidence interval for each dataset of different sizes using R. We used the Student's test implemented in R to test the equality of the results of the two analysis of substitution saturation (ϕ and slope of the regression) between datasets of different sizes when the hypothesis of normality was verified, and the test of Wilcoxon otherwise.

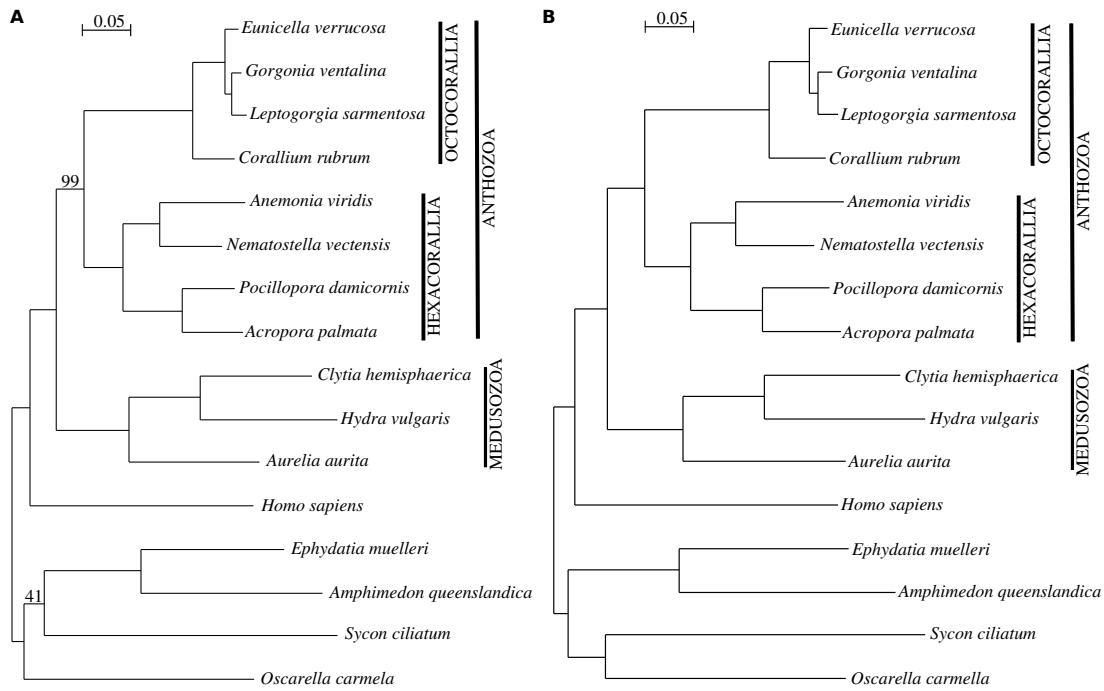
For datasets of 15 sequences, results indicated that all sequences were phylogenetically informative with a mean ϕ of 0.13 [0.12 – 0.13]. In the case of datasets of 50 sequences, all sequences were phylogenetically informative with a mean ϕ of 0.062 [0.061 – 0.063]. Finally, for the datasets of 95 sequences, no sequence was phylogenetically informative with a mean ϕ of 0.031 [0.030 – 0.031]. The mean ϕ was significantly different in the datasets with different number of sequences (Student's test, $p < 2.2 \cdot 10^{-16}$).

In the case of the slope of the regression of the plot of phylogenetic distances

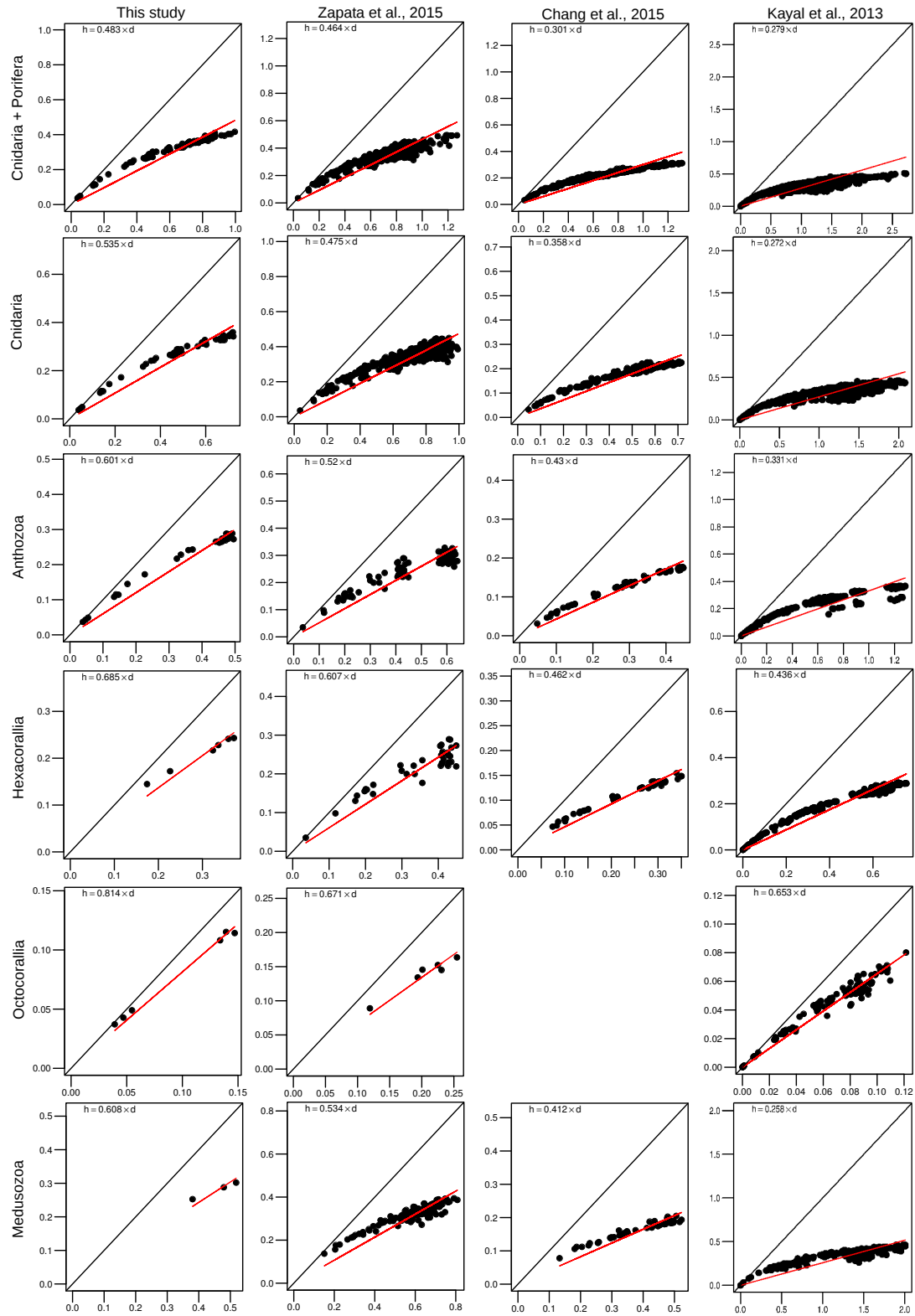
against observed distances, datasets of 15 sequences presented a mean slope of 0.30 [0.28 – 0.30], datasets of 50 sequences presented a mean slope of 0.29 [0.26 – 0.31] and datasets of 95 sequences presented a mean slope of 0.28 [0.28 – 0.29]. The mean slope was not significantly different in the datasets with different number of sequences (Student’s test, $p = 0.26$ for the comparison 15 sequences vs 50 sequences; Wilcoxon’s test, $p = 0.077$ for the comparison 15 sequences vs 95 sequences and $p = 0.66$ for the comparison 50 sequences vs 95 sequences).

Our analysis of sensitivity showed that the test of Steel *et al.* (1993) [302] was sensitive to the number of sequences and was not appropriate in our comparative study including datasets with very different number of sequences (Table A.2). The slope of the regression of the plot of phylogenetic distances against observed distances was less affected by the number of sequences. Therefore, we chose to keep only this last analysis of substitution saturation for our analysis.

Appendix 3. Phylogenetic relationships among cnidarians based on **A**. Maximum-likelihood analysis; **B**. Bayesian analysis. Bootstrap supports in **A** were all 100 % apart from the two indicated values. In **B**, all posterior probabilities were equal to 1.



Appendix 4. Saturation plots of the hierarchical substitution saturation analysis.



Annexe B

Annexes du Chapitre 2

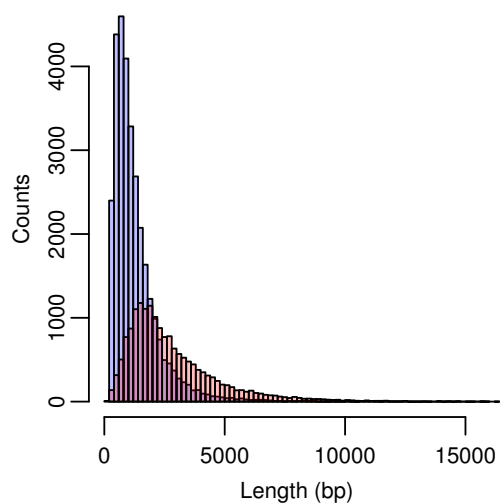


FIGURE B.1 – Size distribution of contig lengths (bp) for annotated contigs (red) and non annotated ones (blue).

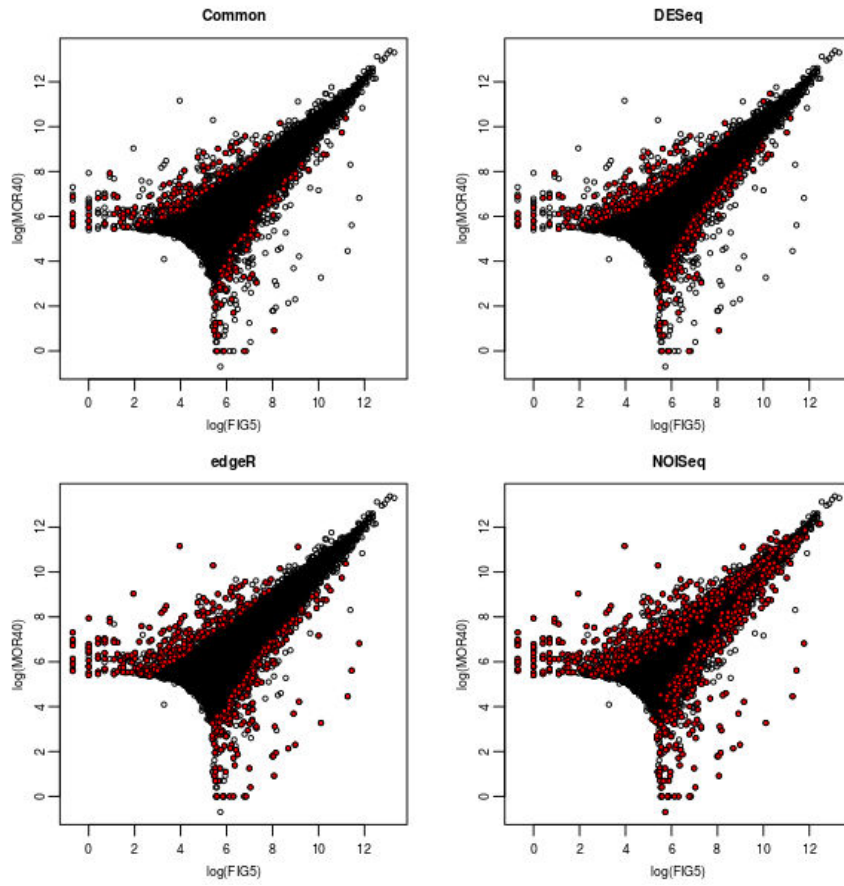


FIGURE B.2 – Expression values for the two depths (FIG 5 : Figuiet – 5 m ; MOR40 : Morgiou – 40 m. Differentially expressed genes between the depth detected by each method are indicated (red).

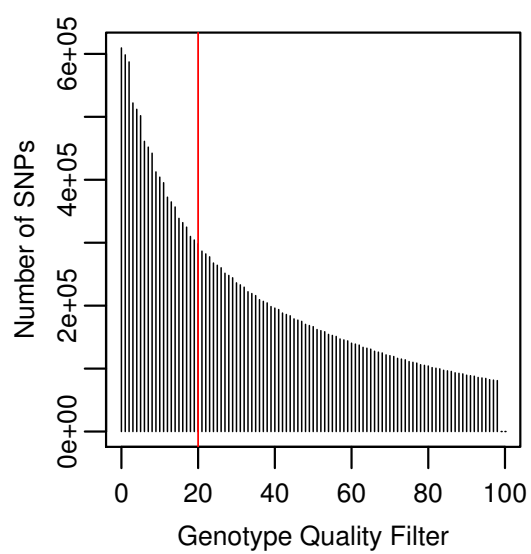


FIGURE B.3 – Effect of the increase of the Genotype Quality parameter on the number of SNPs. The red vertical line indicate the threshold applied.

TABLE B.1 – Best informative UniProt Match and for annotated differentially expressed contigs (contigs with non informative UniProt hits are not indicated).

Contig name	Best Uniprot Match	Uniprot ID	Overexpression depth (m)
Contig_13537	TNF-receptor-associated factor	I3WEV1	5
Contig_13607	Lectin	D4QD85	5
Contig_14781	Polycystic kidney disease protein	K1RA29	5
Contig_14885	Guanine nucleotide-binding protein subunit gamma-e	Q9NFZ3	5
Contig_15308	B-box zinc finger family protein	L8GEK5	5
Contig_16117	E3 ubiquitin-protein ligase	G9L0H4	40
Contig_16238	Protein KEL-3, isoform a	O18147	5
Contig_16253	Puromycin-sensitive aminopeptidase	H9ZFL2	5
Contig_17918	Up-regulator of cell proliferation	Q8TCY9	5
Contig_1856	TNF-receptor-associated factor 1	Q706I7	5
Contig_18560	Putative flagellar-associated protein	M1T236	5
Contig_1895	TLDc domain-containing protein	J9IEY4	5
Contig_19627	Cell adhesion molecule	B7QEU6	5
Contig_20033	Metallo-beta-lactamase superfamily protein	A2E3J2	5
Contig_23477	PugilistDominant	B0WQG9	40
Contig_24277	Phospholipase A2-like protein	K1QJTJ0	5
Contig_24664	Golgi-associated plant pathogenesis related protein	R0JF26	40
Contig_24677	Low-density lipoprotein receptor	B7P761	5
Contig_26186	Protein sidekick-1	R0J TZ8	40
Contig_26239	Superfamily II DNA/RNA helicase	S2RL03	40
Contig_28601	TNF-receptor-associated factor 1	Q706I7	5
Contig_29949	TLDc domain-containing protein	J9J3R9	40
Contig_30208	Zinc finger protein	B7QGF2	40
Contig_3040	TLDc domain-containing protein	J9IEY4	5
Contig_31003	Receptor-type tyrosine-protein phosphatase	S7Q5K6	40
Contig_31055	Transposase domain-containing protein	A0JPR6	40
Contig_3328	TNF-receptor-associated factor 1	Q706I7	5
Contig_33864	40S ribosomal protein	B6RB17	40
Contig_34160	Protein sidekick-2	K1QWM8	40
Contig_34357	Nucleolar and spindle-associated protein	F1NYB3	40
Contig_35953	60S ribosomal protein	D6PVS7	40
Contig_36032	Alpha-protein kinase	K1QAN8	40

Contig_36262	Vasohibin-1	T2MBD5	40
Contig_36724	Putative tick transposon	L7MMP3	40
Contig_37569	E3 ubiquitin-protein ligase	V8NUV2	40
Contig_37744	Polyprotein	Q6IMA0	40
Contig_37887	TLDc domain-containing protein	J9J3R9	40
Contig_3920	TLDc domain-containing protein	J9IE60	5
Contig_40045	B-box zinc finger family protein	L8GEK5	40
Contig_40372	Kinetochore protein Nuf2	Q9BZD4	40
Contig_40374	Putative DUF962 domain protein	K1XE20	40
Contig_40896	Apextrin-like protein	G0YL87	40
Contig_41232	Putative tick transposon	L7M0W9	40
Contig_41234	F-type lectin	F7J048	5
Contig_41274	Serine/threonine protein kinase	D9SFB2	40
Contig_4304	Tumor necrosis factor receptor-associated factor 4	A2TK69	5
Contig_44779	Elongation factor 1-delta	K1QWN2	5
Contig_44801	Collagen triple helix repeat-containing protein	M7BN13	40
Contig_44808	Glutathione S-transferase	D8LHF0	40
Contig_46035	Exportin-7	V5G3C9	40
Contig_4656	PDK domain-containing protein	E1VVF4	5
Contig_46704	Membrane metallo-endopeptidase-like	K1PRU0	40
Contig_46885	TNFAIP3 interacting protein 1	S5DIK3	40
Contig_47741	Kinetochore protein Nuf2	M2Y9U9	40
Contig_5325	Metallo-beta-lactamase superfamily protein	A2E3J2	5
Contig_5842	Exonuclease 3-5 domain-containing protein	L5KHD4	5
Contig_5983	PT repeat superfamily protein	E6KPT8	5
Contig_5994	Elongation factor 2 kinase	N9UTX9	5
Contig_6966	Lysosomal protein NCU-G1	Q8WWB7	40
Contig_7730	E3 ubiquitin-protein ligase	M7BQ33	5
Contig_779	Histone H2A	A7RJL4	5
Contig_8253	Potassium channel	A7SI05	5
Contig_9470	Kelch-like protein	S9YPT0	5
Contig_9497	Cysteine and glycine-rich protein	Q28IA0	5
Contig_9811	Mannuronan C-5-epimerase C-terminal	D7G651	5
Contig_9890	Putative tick transposon	L7MLC2	5
Contig_9957	Phospholipase A2	G5B779	5

TABLE B.2 – Best informative UniProt Match and for annotated contigs containing differentially fixed SNPs (contigs with non informative UniProt hits are not indicated).

Contig name	Best Uniprot Match	Uniprot ID	Number of differentially fixed SNPs
Contig_1033	Receptor-type tyrosine-protein phosphatase F	A4IFW2	1
Contig_5010	Peptidoglycan-binding domain 1 protein	K9QCD2	2
Contig_9277	Pescadillo homolog	A7SWH1	1
Contig_9432	Solute carrier family 43, member 3	K9IVI0	1
Contig_17273	Exocyst complex component 6	T2M341	2
Contig_17293	Cyclin B	E3UP94	1
Contig_18167	WD repeat-containing protein mio	R0KCH8	2
Contig_22756	Chymotrypsin-like protease CTRL-1	G3HKW8	1
Contig_22987	Histone H2B	C1BLJ0	1
Contig_31385	Beta-1,4-N-acetyl-galactosaminyl transferase 3	E9CI98	1
Contig_31567	Fas-binding factor 1	T2M956	1
Contig_33142	Translocon-associated protein subunit alpha	T2MG75	1
Contig_34221	CMP-Neu5Ac—lipooligosaccharide alpha 2-3 sialyltransferase	E1X9Z0	1
Contig_34479	SCO-spondin	K1PYZ9	3
Contig_36584	Lysine-specific histone demethylase 1B	V9K7V6	1
Contig_37373	RING finger protein, putative	B7P1R6	1
Contig_40193	Putative helicase	K9IPI0	1
Contig_41655	Selenoprotein J	V9KMI2	2
Contig_41682	Collagen triple helix repeat-containing protein 1	M7BN13	2
Contig_44321	Uracil-DNA glycosylase	A7SYX6	1
Contig_46226	Genome polyprotein	G0Y2E1	1
Contig_47087	CKLF-like MARVEL transmembrane domain containing protein 4	A0A034WF78	1

Annexe C

Annexes du Chapitre 3

C.1 Figures et Tables Supplémentaires

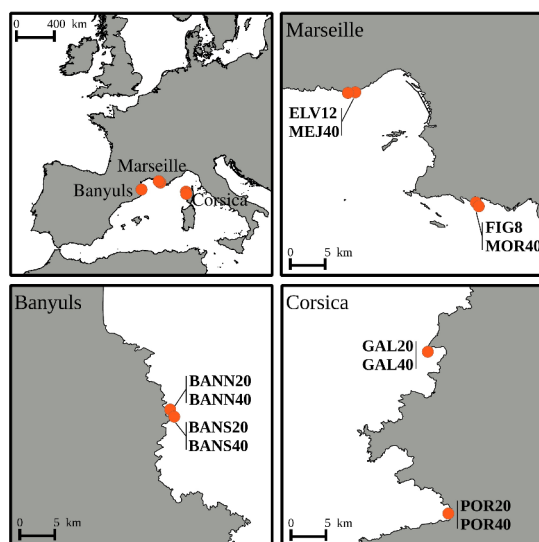


FIGURE C.1 – Locations of the sampling sites of the red coral individuals among the studied three geographical regions.

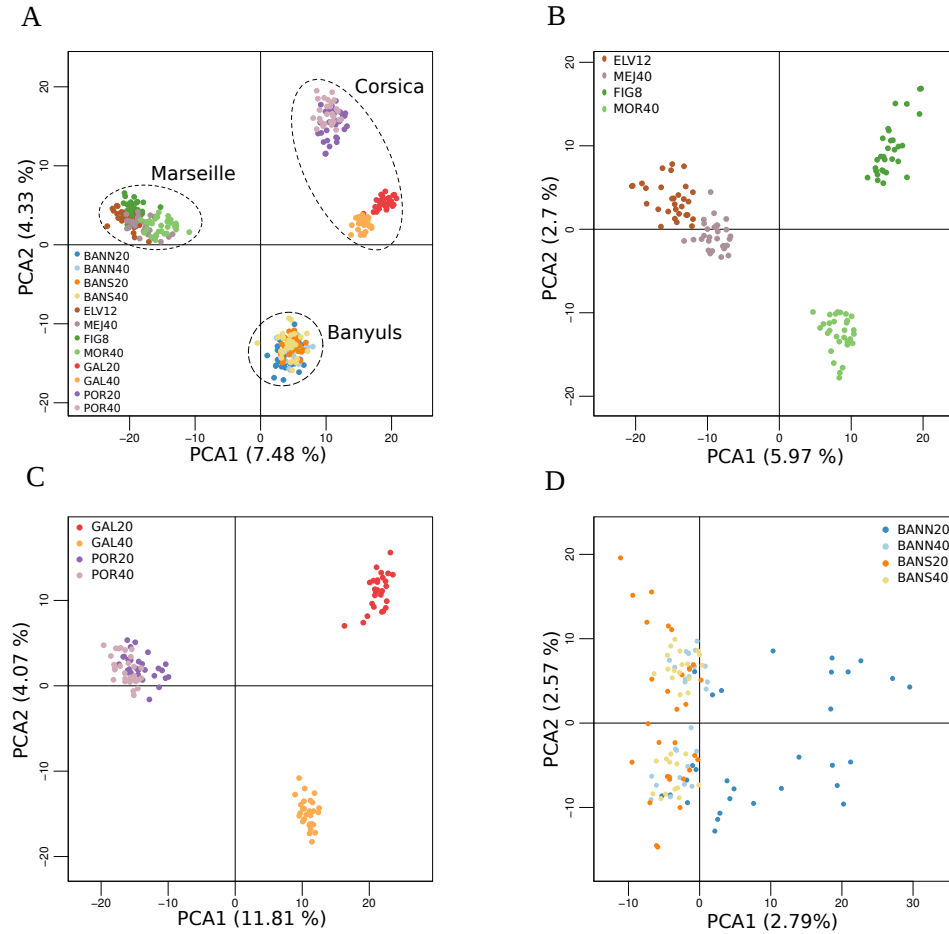


FIGURE C.2 – Principal component analysis (Axes 1 and 2), using only putatively neutral SNPs, of the **A.** 12 red coral populations ($n = 354$ individuals, 25 669 SNPs), **B.** four red coral populations from Marseille ($n = 119$ individuals, 26 898 SNPs), **C.** four red coral populations from Corsica ($n = 117$ individuals, 26 592 SNPs), **D.** four red coral populations from Banyuls ($n = 118$ individuals, 27 069 SNPs).

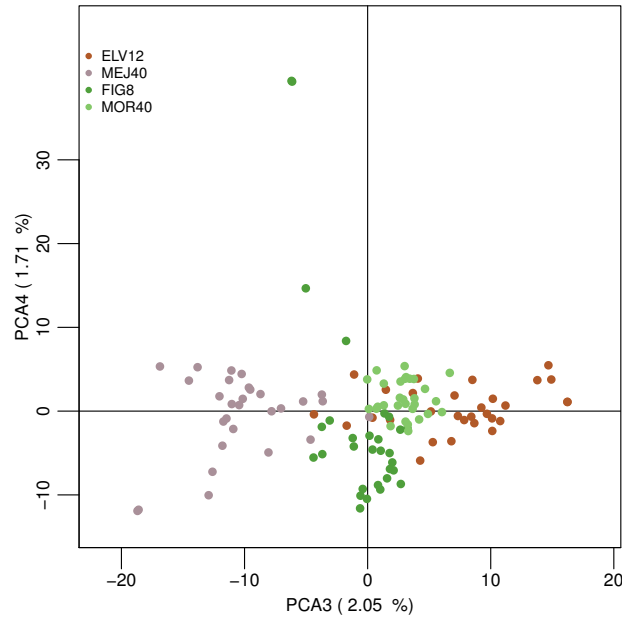


FIGURE C.3 – Analyse en composante principale (axes 2 et 3) réalisée sur les quatre populations de Marseille ($n = 119$ individuals, 27 026 SNPs) (figure non présentée dans l'article).

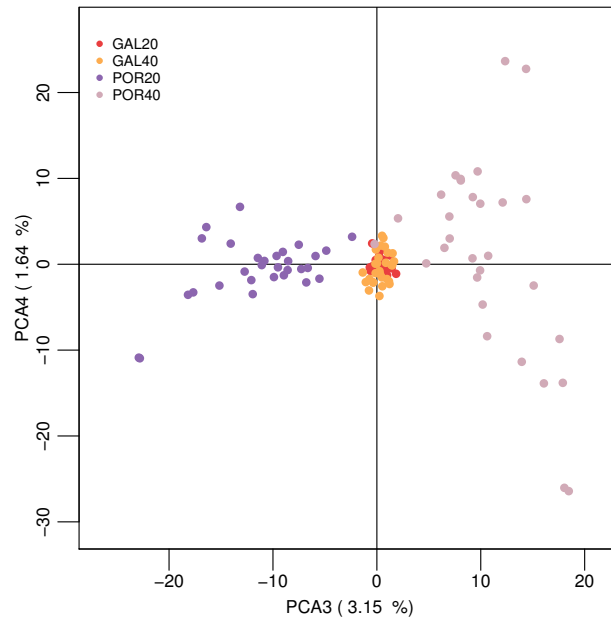


FIGURE C.4 – Analyse en composante principale (axes 2 et 3) réalisée sur les quatre populations de Corse ($n = 117$ individuals, 27 026 SNPs) (figure non présentée dans l'article).

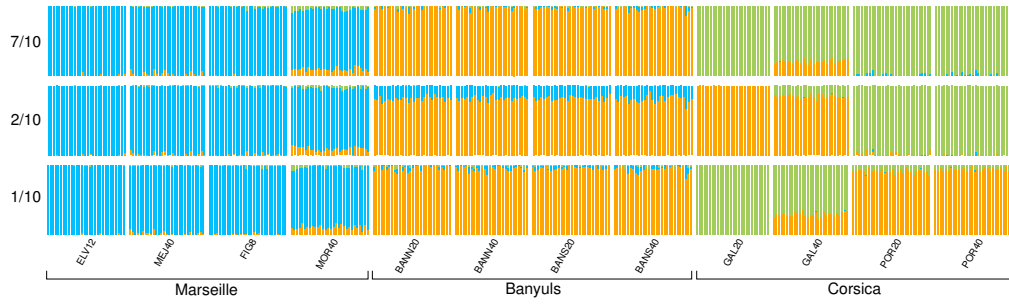


FIGURE C.5 – Results from Bayesian individual clustering with STRUCTURE for $K = 3$. The three figures correspond to major and minor modes detected.

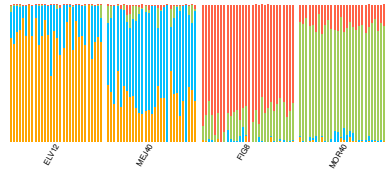


FIGURE C.6 – Résultats de STRUCTURE pour $K = 4$ pour les populations de Marseille (figure non présentée dans l'article).

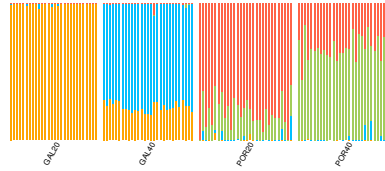


FIGURE C.7 – Résultats de STRUCTURE pour $K = 4$ pour les populations de Corse (figure non présentée dans l'article).

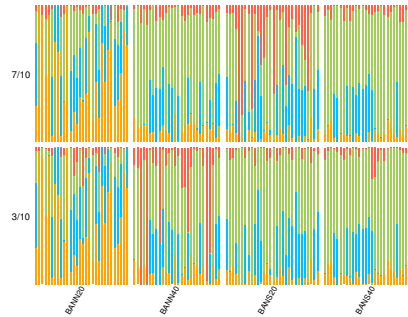


FIGURE C.8 – Résultats de STRUCTURE pour $K = 4$ pour les populations de Banyuls. Les deux figures correspondent aux modes majeur et mineur détectés (Figure non présentée dans l'article).

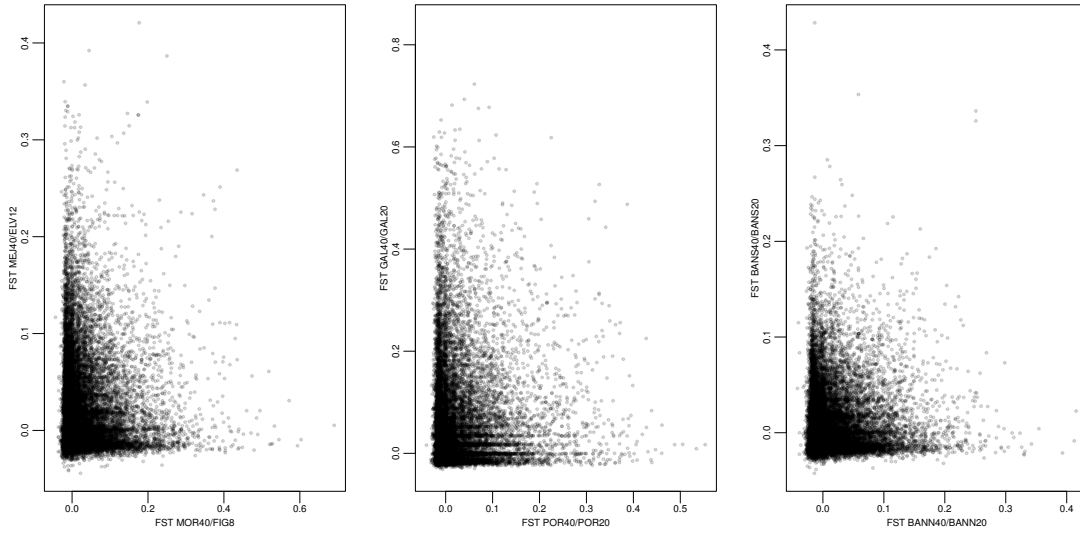


FIGURE C.9 – Joint distribution of between-depths F_{ST} in the three geographical regions.

TABLE C.1 – Pairwise F_{ST} estimates using only putatively neutral SNPs. All comparisons were highly significant. Intra-region comparisons are highlighted.

	BANN20	BANN40	BANS20	BANS40	ELV12	MEJ40	FIG8	MOR40	GAL20	GAL40	POR20
BANN40	0.02										
BANS20	0.02	0.01									
BANS40	0.02	0.01	0.01								
ELV12	0.13	0.12	0.12	0.12							
MEJ40	0.11	0.10	0.10	0.10	0.03						
FIG8	0.13	0.12	0.12	0.12	0.09	0.07					
MOR40	0.10	0.10	0.09	0.09	0.07	0.05	0.04				
GAL20	0.17	0.16	0.16	0.16	0.23	0.21	0.22	0.20			
GAL40	0.11	0.09	0.09	0.09	0.18	0.16	0.18	0.15	0.10		
POR20	0.13	0.12	0.12	0.12	0.17	0.15	0.16	0.14	0.17	0.11	
POR40	0.13	0.12	0.12	0.12	0.16	0.15	0.16	0.14	0.18	0.12	0.04

TABLE C.2 – Pairwise F_{ST} estimates using only outlier SNPs from the ARLEQUIN analysis. All comparisons were highly significant. Intra-region comparisons are highlighted.

	BANN20	BANN40	BANS20	BANS40	ELV12	MEJ40	FIG8	MOR40	GAL20	GAL40	POR20
BANN40	0.04										
BANS20	0.06	0.03									
BANS40	0.06	0.03	0.02								
ELV12	0.19	0.18	0.19	0.18							
MEJ40	0.16	0.14	0.15	0.15	0.04						
FIG8	0.19	0.17	0.18	0.18	0.23	0.19					
MOR40	0.14	0.12	0.13	0.13	0.19	0.14	0.08				
GAL20	0.27	0.27	0.28	0.27	0.37	0.35	0.36	0.32			
GAL40	0.15	0.13	0.15	0.13	0.26	0.24	0.26	0.22	0.17		
POR20	0.23	0.21	0.21	0.21	0.26	0.22	0.22	0.18	0.41	0.30	
POR40	0.24	0.23	0.24	0.23	0.26	0.23	0.23	0.19	0.42	0.31	0.07

TABLE C.3 – Results of the annotation analysis of candidates for local adaptation in the three geographical regions.

Contig Name	evalue	Description	GO	Region
Contig_16793	1.1.10 ⁻¹¹⁷	LON peptidaseN-terminal domain and RING finger partial	ubiquitin-protein transferase activity zinc ion binding metal ion binding proteolysis ATP-dependent peptidase activity protein ubiquitination	Marseille
Contig_20016	0	Chromodomain-helicase-DNA-binding 1-like		
Contig_23068	2.8.10 ⁻¹¹⁷	RNA-directed DNA polymerase from mobile element jockey-like	nucleic acid phosphodiester bond hydrolysis RNA-directed DNA polymerase activity endonuclease activity RNA-dependent DNA biosynthetic process	
Contig_38936	4.1.10 ⁻²³	E3 ubiquitin-ligase DZIP	nucleic acid binding zinc ion binding	
Contig_44372	3.1.10 ⁻⁶¹	PREDICTED : uncharacterized protein LOC107346707	binding	
Contig_47492	7.8.10 ⁻⁹⁸	No description	metal ion binding oxidoreductase activity metabolic process oxidation-reduction process	Corsica
Contig_7346	2.3.10 ⁻¹⁵⁰	FAM46C-like		
Contig_10570	4.8.10 ⁻³⁷	PREDICTED : uncharacterized protein		
Contig_10731	2.8.10 ⁻⁶⁸	allene oxide synthase-lipoxygenase	oxidoreductase activity, acting on single donors with incorporation of molecular oxygen, incorporation of two atoms of oxygen	

				metal ion binding fatty acid metabolic process	
Contig_11268	2.3.10 ⁻⁹⁰	neuronal acetylcholine receptor subunit alpha-9	integral component of membrane plasma membrane part extracellular ligand-gated ion channel activity ion transport synaptic transmission response to stimulus biological regulation		
Contig_11731	0	E3 ubiquitin-ligase RNF213	binding		
Contig_12920	5.5.10 ⁻¹⁴⁹	Stonin-2 isoform X2	cytoplasmic part vesicle-mediated transport single-organism process intracellular transport		
Contig_13771	4.0.10 ⁻¹⁵¹	A-kinase anchor mitochondrial	hemopoiesis cell differentiation		
Contig_16202	6.5.10 ⁻⁶⁵	PIN2 TERF1-interacting telomerase inhibitor 1	Nucleus chromosome intracellular organelle part nucleic acid binding regulation of cellular process		
			aldehyde dehydrogenase (NAD) activity succinate-semialdehyde dehydrogenase (NAD+) activity succinate-semialdehyde dehydrogenase [NAD(P)+] activity		

Contig_16843 0 Succinate-semialdehyde mitochondrial

				glycosylceramide metabolic process multicellular organism development glutamine family amino acid metabolic process gamma-aminobutyric acid catabolic process dicarboxylic acid metabolic process oxidation-reduction process
Contig_16868	4.3.10 ⁻²⁴	glioma tumor suppressor candidate region gene 1-like	Membrane integral component of membrane	
Contig_17255	7.0.10 ⁻⁵³	Nanos 1	zinc ion binding RNA binding	
Contig_18282	1.3.10 ⁻⁴⁴	centromere K	nucleus	
Contig_19099	0	epithelial growth factor receptor substrate 15-like partial	Calcium ion binding	
Contig_19611	2.0.10 ⁻¹³⁴	DDB1- and CUL4-associated factor 5	mitochondrion Cul4-RING E3 ubiquitin ligase complex	
Contig_24102	1.6.10 ⁻⁶²	nuclease HARBII		
Contig_24221	0	tRNA (guanine(26)-N(2))-diethyltransferase	nucleus mitochondrion tRNA (guanine-N2-)-methyltransferase activity binding tRNA N2-guanine methylation	
Contig_31623	0	No description	binding	
			nucleus cytoplasm	

Contig_32690 1.4.10⁻¹⁶³ ubiquitin carboxyl-terminal hydrolase
isozyme L5

				thiol-dependent ubiquitin-specific protease activity binding ubiquitin-dependent protein catabolic process protein deubiquitination regulation of proteolysis forebrain development regulation of cellular protein metabolic process
Contig_33144	6.6.10 ⁻⁵⁶	Dok-7		protein kinase binding positive regulation of protein tyrosine kinase activity insulin receptor binding
Contig_35407	3.8.10 ⁻¹¹	hypothetical protein AC249_AIPGENE14243		
Contig_36059	5.9.10 ⁻⁴⁷	cytosolic non-specific dipeptidase		exopeptidase activity metabolic process
Contig_36102	2.8.10 ⁻¹³⁵	nucleolar complex 4 homolog		Nuclear part membrane cell part
Contig_37478	8.9.10 ⁻²⁸	fibroblast growth factor receptor 3		protein kinase activity phosphorylation positive regulation of phosphorylation regulation of primary metabolic process
Contig_38721	2.3.10 ⁻⁴¹	L-seryl-tRNA(Sec) kinase		phosphorylation kinase activity
				Ada2/Gcn5/Ada3 transcription activator complex

Contig_38739 2.0.10⁻⁵⁶

Dr1

				DNA binding transcription corepressor activity TBP-class protein binding protein heterodimerization activity negative regulation of transcription from RNA polymerase II promoter transcription, DNA-templated histone H3 acetylation
Contig_39033	7.1.10 ⁻¹⁵⁵	homeodomain transcription factor 1 isoform X1		intracellular membrane-bounded organelle
Contig_40626	1.0.10 ⁻⁵¹	RNA-directed DNA polymerase from mobile element jockey-like		nucleic acid phosphodiester bond hydrolysis RNA-directed DNA polymerase activity RNA binding endonuclease activity RNA-dependent DNA biosynthetic process
Contig_41259	9.9.10 ⁻⁴⁵	serine protease 23-like		proteolysis serine-type endopeptidase activity serine-type peptidase activity peptidase activity
Contig_41360	8.9.10 ⁻¹³⁴	tonsoku		nuclear lumen protein complex DNA metabolic process cellular response to DNA damage stimulus single-organism metabolic process

			single-organism cellular process
Contig_41417	3.3.10 ⁻²⁷	eukaryotic translation initiator factor 4E-binding 1	cytosol eukaryotic initiation factor 4E binding translational initiation signal transduction negative regulation of translational initiation
Contig_41500	1.8.10 ⁻¹²⁸	Cytoplasmic 1-like	cytoskeleton plasma membrane focal adhesion dense body
Contig_42135	0	No description	membrane exopeptidase activity
Contig_42435	3.8.10 ⁻⁸⁴	Tensin-partial 3	focal adhesion actin binding cell-substrate junction assembly fibroblast migration
Contig_43083	2.0.10 ⁻⁴⁷	No description	
			early endosome cytosol kinesin complex phosphatidylinositol-3,4,5-trisphosphate binding ATP-dependent microtubule motor activity, plus-end-directed phosphatidylinositol-3-phosphate binding phosphatidylinositol-3,4-bisphosphate binding

Contig_45623 0 kinesin KIF16B isoform X2

			phosphatidylinositol-3,5-bisphosphate binding formation of primary germ layer regulation of receptor recycling Golgi to endosome transport microtubule-based movement epidermal growth factor receptor signaling pathway endoderm development fibroblast growth factor receptor signaling pathway cytoskeleton-dependent intracellular transport receptor catabolic process early endosome to late endosome transport		
Contig_45771	8.3.10 ⁻⁹⁴	No description	dopamine neurotransmitter receptor activity G-protein coupled receptor signaling pathway		
Contig_47280	2.0.10 ⁻¹⁶⁸	No description	regulation of Rho protein signal transduction metabolic process Rho guanyl-nucleotide exchange factor activity transferase activity positive regulation of GTPase activity		
Contig_47344	0	allene oxide synthase-lipoxygenase	oxidoreductase activity, acting on single donors with incorporation of molecular oxygen, incorporation of two atoms of oxygen		

				metal ion binding fatty acid metabolic process	
Contig_5993	1.5.10 ⁻⁹⁵	GA-binding subunit beta-partial		nucleus protein homodimerization activity transcription regulatory region DNA binding protein heterodimerization activity transcription, DNA-templated positive regulation of transcription from RNA polymerase II promoter	
Contig_7742	0	Indole-3-acetaldehyde oxidase-like		oxidoreductase activity ion binding single-organism metabolic process	
Contig_8936	5.2.10 ⁻¹⁰⁵	drebrin isoform X1		actin filament binding synapse assembly ruffle assembly actin binding neuron projection morphogenesis receptor-mediated endocytosis intracellular	
Contig_8963	3.4.10 ⁻¹⁶⁹	tubuline delta chain		Intracellular part	
Contig_26377	6.7.10 ⁻¹³⁸	ribokinase isoform X1 2		ribokinase activity D-ribose metabolic process carbohydrate phosphorylation	Banyuls

C.2 Aspect phylogéographique - corail rouge d'Atlantique

Grâce à une collaboration avec Joana Boavida, Sophie Arnaud-Haond et Ester Serrao, des individus issus d'un échantillonnage en Atlantique ont pu être intégrés au RAD-Séquençage réalisé pendant cette thèse [34]. L'aspect phylogéographie des échantillons d'Atlantique et de Méditerranée sera développé ultérieurement. Une première analyse de la structure génétique de ces échantillons indique que les différences génétiques entre les populations de Méditerranée étudiées pendant cette thèse sont inférieures aux différences entre les populations de Méditerranée et les individus d'Atlantique (Fig. C.10).

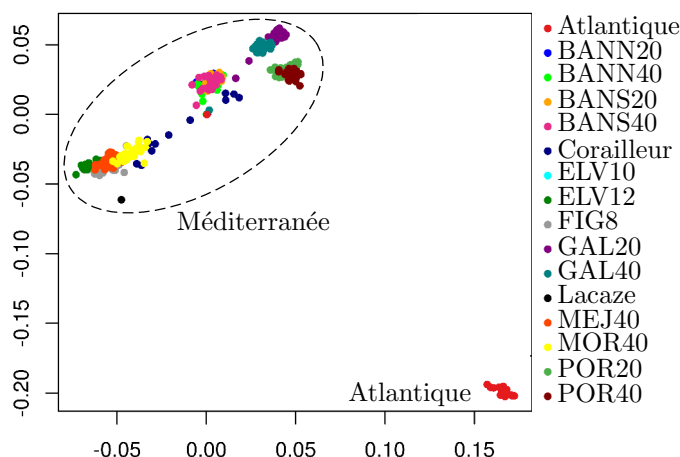


FIGURE C.10 – Analyse en Composantes Principales (Axes 1 et 2) des individus échantillonnés en Atlantique et en Méditerranée ($n = 416$, 32 687 SNPs). L'échantillon "Lacaze" correspond à un individu prélevé à 350 m dans le canyon de Lacaze-Duthier. Les échantillons "Corailleurs" correspondent à des individus profonds prélevés en Corse.

C.3 Analyses exploratoires réalisées non présentées dans l'article

C.3.1 Recherche statistique d'un effet “profondeur” dans les données de polymorphisme issues de RAD-Séquençage

Nous avons ici adapté la méthode mise en place par Soria-Carrasco *et al.* (2014) [298] pour étudier l'adaptation de populations de phasmes à leur plante hôte. Dans notre cas) implique fréquemment le même SNP. Dans notre cas, il s'agit de tester si la divergence entre les paires de populations issues de profondeurs différentes implique fréquemment les mêmes locus et donc d'identifier un patron de convergence dans l'adaptation à la profondeur.

Schéma expérimental : Considérons ici trois lieux : Marseille, Corse et Banyuls. Chacun de ces lieux contient deux sites (Site 1 et Site 2), et chacun de ces sites deux profondeurs (Fig. C.11). On appelle “population” l'ensemble des individus échantillonnés à une profondeur donnée, pour un site et un lieu donné.

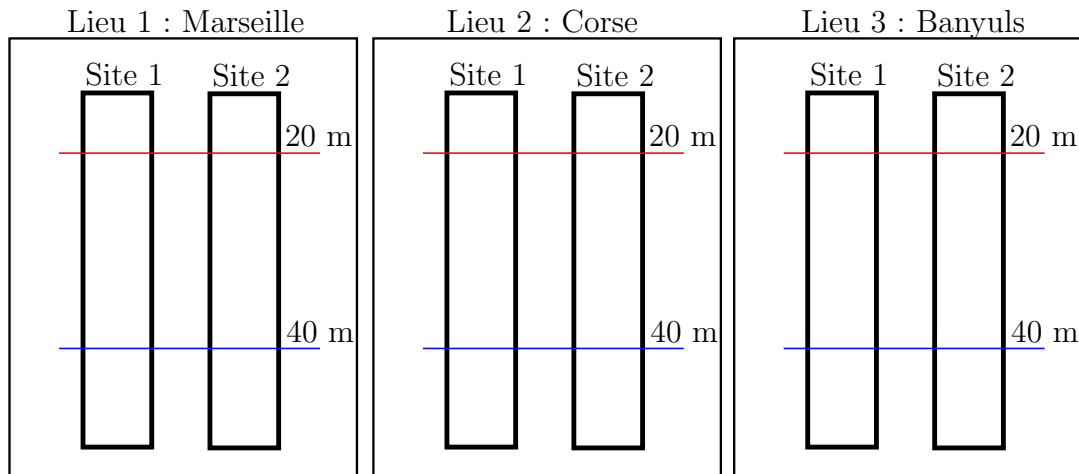


FIGURE C.11 – Schéma expérimental.

Problématique : Peut-on identifier des SNPs présentant des F_{ST} élevés dans plus de comparaisons entre populations de profondeurs différentes que ce que l'on attendrait sous l'effet du hasard ?

Méthode : La structure génétique de nos populations est telle que des F_{ST} entre des populations de lieux différents auraient tendance à être d'un ordre de grandeur supérieur par rapport aux F_{ST} au sein d'un même lieu, ce qui pourrait masquer un éventuel effet profondeur. Afin de s'affranchir de cet effet du lieu, deux analyses indépendantes ont été réalisées :

- Analyse inter-lieux : regroupement des deux sites du même lieu et comparaison des F_{ST} entre populations de lieux différents (six comparaisons entre même profondeur et six comparaisons entre profondeurs différentes, Fig. C.12A et C.12B).

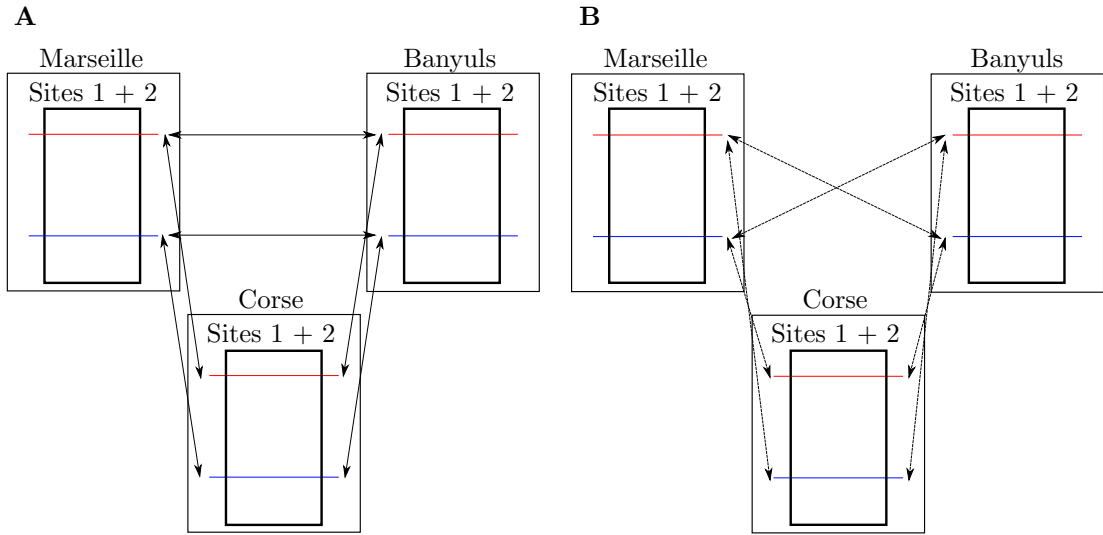


FIGURE C.12 – Schéma expérimental inter-lieux.

- Analyses intra-lieux : comparaisons de trois analyses indépendantes réalisées au sein d'un même lieu. A l'intérieur de chaque lieu, comparaison des F_{ST} entre populations de sites différents (deux comparaisons entre même profondeur et deux comparaisons entre profondeurs différentes, Fig. C.13A et C.13B).

Dans tous les cas, les SNPs pourront être classés théoriquement dans plusieurs catégories (Table C.4) :

- Pas d'effet profondeur ou lieu.
- Effet lieu.
- Effet profondeur plus fort que effet lieu.
- Effet biologiquement difficilement interprétable.

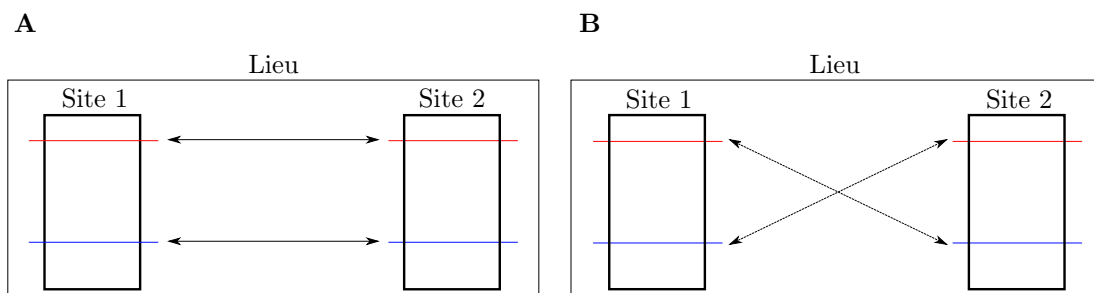


FIGURE C.13 – Schéma expérimental intra-lieu.

TABLE C.4 – Interprétation en fonction des résultats théoriques des tests réalisés.

Nombre de comparaisons entre populations de mêmes profondeurs significativement plus élevé que l'attendu théorique	Non	Oui	Non	Oui
Nombre de comparaisons entre populations de profondeurs différentes significativement plus élevé que l'attendu théorique	Non	Oui	Oui	Non
Interprétation	Pas d'effet profondeur ou lieu	Effet lieu	Effet profondeur plus fort que effet lieu	?

Mise en place du test :

1) Génération de la distribution observée des F_{ST} :

- Pour chaque comparaison de paires de populations, calcul du F_{ST} de Weir et Cockerham pour chaque SNP [334].
- Pour chaque comparaison de paires de populations, extraction des 10 % plus grands F_{ST} .
- Calcul, pour chaque SNP, du nombre de comparaisons entre populations de **même profondeurs** pour lesquelles il présente un F_{ST} élevé, et du nombre de comparaisons entre populations de **profondeurs différentes** pour lesquelles il présente un F_{ST} élevé.

TABLE C.5 – Interprétation en fonction des résultats des tests réalisés.

Nombre de comparaisons entre populations de mêmes profondeurs significativement plus élevé que l'attendu théorique	1783 SNPs significatifs	0	590
Nombre de comparaisons entre populations de profondeurs différentes significativement plus élevé que l'attendu théorique		538	0
Interprétation	Effet lieu	Effet profondeur plus fort que effet lieu	?

2) Génération de la distribution attendue des F_{ST} :

- Pour chaque comparaison de paires de populations, permutation des noms des SNPs afin de briser la relation entre le SNP et son F_{ST} .
- Calcul, pour chaque SNPs, du nombre de comparaisons entre populations de **même profondeurs** pour lesquelles il présente un F_{ST} élevé, et du nombre de comparaisons entre populations de **profondeurs différentes** pour lesquelles il présente un F_{ST} élevé.

3) Comparaison de la distribution observée et attendue :

Pour un SNP donné, on considère le test significatif si le nombre de comparaisons pour lesquelles il présente un F_{ST} élevé est supérieure à 99 % des données théoriques (Table C.4).

Résultats :

1) Analyse inter-lieux : Plusieurs SNPs présentant un effet profondeur ont été mis en évidence, mais autant de SNPs ayant un effet biologiquement difficilement interprétable (Table C.5).

2) Analyse intra-lieux : mêmes observations que pour l'analyse inter-lieux. Plusieurs SNPs présentant un effet profondeur ont été mis en évidence, mais autant de SNPs ayant un effet biologiquement difficilement interprétable (Table C.6 et C.7).

TABLE C.6 – SNPs significatifs dans les comparaisons entre même profondeurs dans l’analyse intra-région. La diagonale correspond au nombre de SNPs détectés à l’intérieur de chaque région. Les SNPs détectés dans deux régions sont indiqués dans la moitié supérieure du tableau. Aucun SNP n’a été détecté dans les trois régions.

	Marseille	Corse	Banyuls
Marseille	471	9	9
Corse		548	19
Banyuls			526

TABLE C.7 – SNPs significatifs dans les comparaisons entre deux profondeurs différentes dans l’analyse intra-région. La diagonale correspond au nombre de SNPs détectés à l’intérieur de chaque région. Les SNPs détectés dans deux régions sont indiqués dans la moitié supérieure du tableau. Aucun SNP n’a été détecté dans les trois régions.

	Marseille	Corse	Banyuls
Marseille	532	4	12
Corse		669	6
Banyuls			503

Discussion :

Cette analyse a permis de détecter plusieurs SNPs présentant des F_{ST} élevés entre populations de profondeurs différentes mais pas entre populations de même profondeur pouvant être le résultat d’adaptation locale. Cependant, autant de SNPs présentant le patron inverse ont aussi été détectés. La présence de ces derniers SNPs soulève des doutes sur le fait que les SNPs d’intérêts détectés soient bien le résultat d’adaptation locale : si un certain nombre de SNPs peuvent être élevés dans plusieurs comparaisons de même profondeurs et pas entre populations de profondeurs différentes sans explication biologique évidente, alors il peut en être de même dans le cas inverse.

C.3.2 Autres tests de recherche de marqueurs d'adaptation locale

Analyse de Redondance

Nous avons mis en place une analyse de redondance (RDA) pour analyser l'effet de plusieurs variables sur la distribution des géotypes étudiés. Afin de mettre en évidence le rôle de la température ou de la profondeur indépendamment de celui de la géographie, nous avons conditionné la RDA par les coordonnées GPS de chaque population dans un premier temps, puis par les coordonnées de chaque population sur l'axe 1 de l'analyse en composante principale. La significativité de la RDA a été testée par un test de permutations (1 000 permutations).

Les premières analyses ont montré un effet significatif de la température et de la profondeur sur la distribution des géotypes, que la RDA soit conditionnée par la géographie ou pas. Afin de vérifier la pertinence de ces résultats, nous avons testé la significativité d'une RDA estimant les effets d'un vecteur de nombres généré aléatoirement. Quel que soit le vecteur de nombres généré, et que la RDA soit conditionnée ou non par la géographie, le résultat était significatif.

Cet effet significatif de nombres aléatoires pourrait être le résultat d'un taux de dérive génétique très élevé, qui induirait, quelle que soit la variable étudiée, l'existence de gènes dont la fréquence allélique lui serait corrélée.

Bayenv2

Bayenv2 [142] teste si un modèle qui inclut un facteur environnemental s'adapte mieux aux données qu'un modèle nul incluant uniquement la structure génétique neutre. Bayenv2 propose une méthode paramétrique en cas de relation linéaire entre fréquence allélique et variable environnementale, et une méthode non paramétrique quand cette relation n'est pas linéaire. Dans notre cas, nous avons utilisé l'ensemble des SNPs pour estimer la structure génétique neutre. Ce programme étant basé sur une méthode MCMC, plusieurs runs indépendants doivent être réalisés.

Résultats : Malgré différents paramètres testés, il n'a pas été possible d'atteindre une corrélation suffisante entre les différents runs réalisés (aucune corrélation entre les Bayes factor de deux runs quel que soit le nombre d'itérations réalisées).

Cette absence de corrélation peut être due à mauvaise convergence des runs ou au fait que la relation entre notre variable environnementale et les fréquences allélique ne soit pas linéaire. La corrélation est au contraire quasi-parfaite sur le test non paramétrique. Ce test non-paramétrique n'a pas été utilisé par la suite du fait qu'il n'est pas recommandé d'utiliser ce test seul dans les cas où on s'attend à la présence de faux-positifs.

BayeScan

Contrairement à Bayenv2 et BayeScEnv, BayeScan [116], ne peut pas être paramétré par un set de locus neutres. Dans le programme, on définit la probabilité que l'on a d'observer des locus sous sélection dans notre jeu de données (ou autrement dit à quel point le modèle neutre est probable par rapport au modèle avec sélection). Il est recommandé d'utiliser $PO = 100$ (le modèle neutre est 100 fois plus probable que le modèle avec sélection) quand le nombre de locus utilisés est compris entre 1 000 et 10 000. Une valeur de PO trop faible par rapport au nombre de locus utilisé peut conduire à la détection d'un grand nombre de faux positifs. Bien que l'augmentation de la valeur de PO permette de limiter ce nombre de faux positifs, certains d'entre eux peuvent être détectés de façon stable dans le cas de certains modèles démographiques ne respectant pas les hypothèses du modèle.

BayeScan repose sur une méthode bayésienne basée sur un modèle de régression logistique qui sépare les effets locus spécifiques de la sélection des effets de démographie population-spécifique. L'éloignement à la neutralité a lieu quand la composante locus-spécifique est nécessaire pour expliquer le patron de diversité.

Résultats : aucun SNP outlier n'est détecté quelque soit la région considérée. Il semble que le nombre de populations considéré ici (4) soit très insuffisant pour ce programme. Cent cinquante trois SNPs outliers sont détectés sur le jeu de données total, mais ne seront pas exploités par la suite du fait du risque trop important de faux positifs (structure hiérarchique à deux niveaux sur le jeu de données total).

Annexe D

Annexes du Chapitre 4

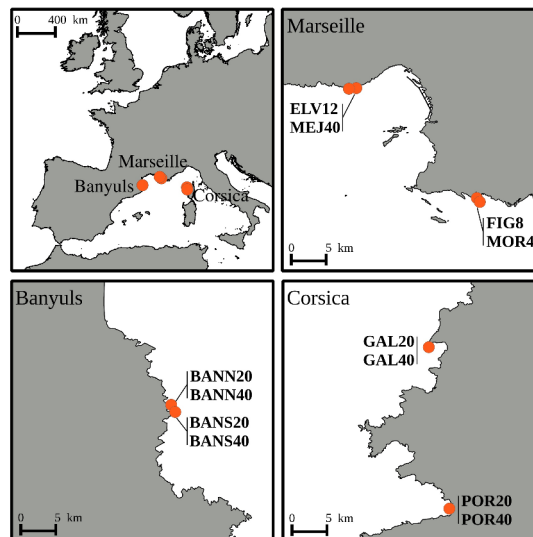


FIGURE D.1 – Locations of the sampling sites of the red coral individuals among the studied three geographical regions.

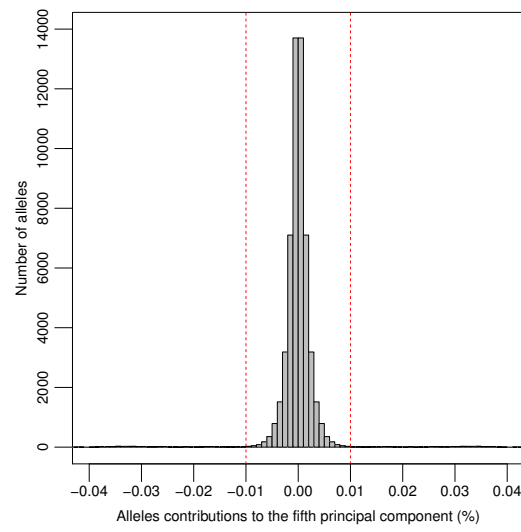


FIGURE D.2 – Loci contribution to the fifth axis of the PCA (see Fig. 4.2). The dash line indicate the threshold of selected loci.

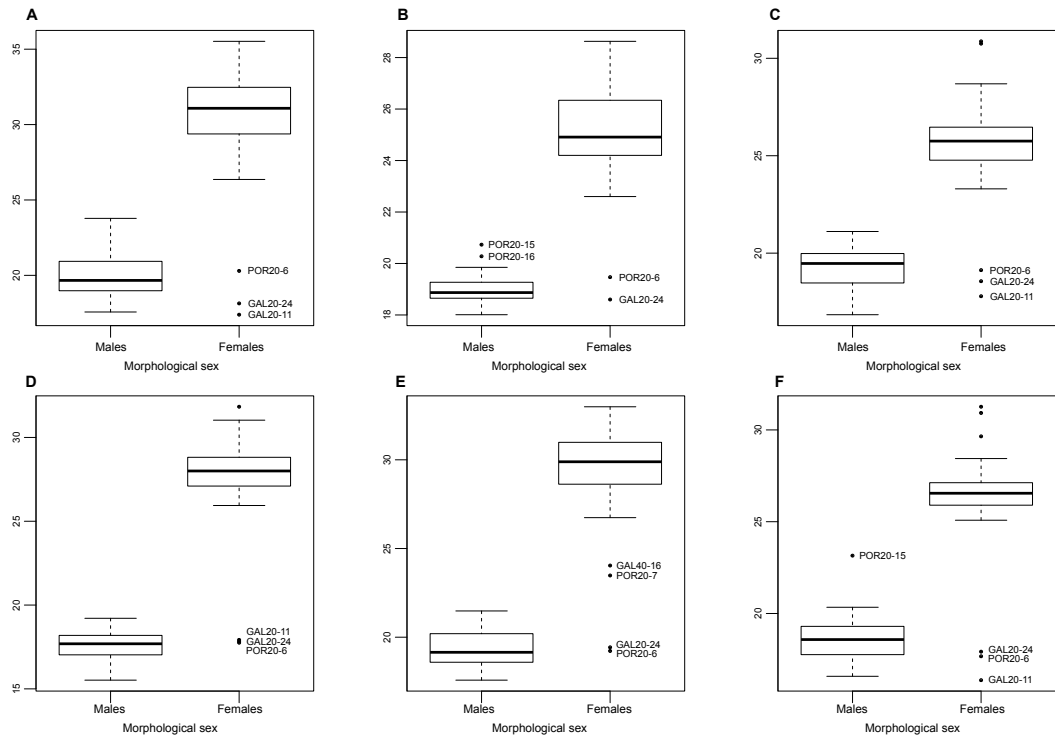


FIGURE D.3 – Result of the real-time PCR (threshold cycle, Ct) of six male-specific loci on morphologically sexed males and females. **A.** Locus_139082. **B.** Locus_144530. **C.** Locus_147437. **D.** Locus_139282. **E.** Locus_158229. **F.** Locus_154834.

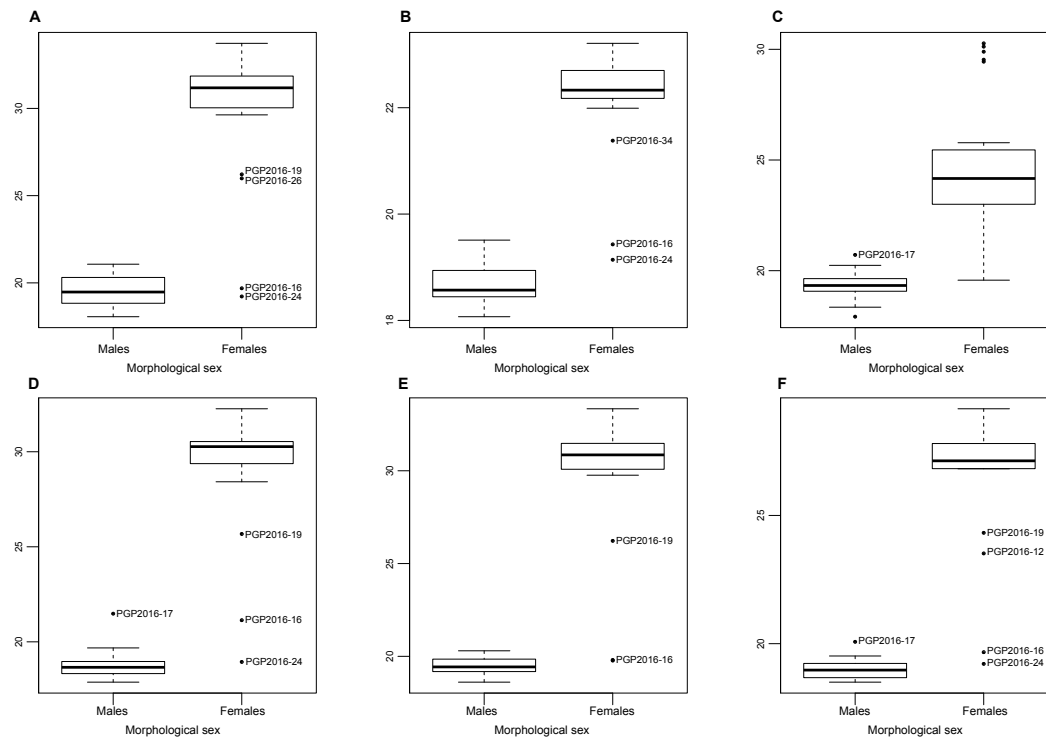


FIGURE D.4 – Result of the real-time PCR (threshold cycle, Ct) of six male-specific loci on morphologically sexed males and females from the independent validation. **A.** Locus_139082. **B.** Locus_144530. **C.** Locus_147437. **D.** Locus_139282. **E.** Locus_158229. **F.** Locus_154834.

TABLE D.1 – Characteristics of red coral sampling sites.

Population	Geographic region	Depth (m)	GPS	GPS
FIG8	Marseille	8	43 ° 12.330'N	5 ° 26.790'E
MOR40	Marseille	40	43 ° 12.060'N	5 ° 27.100'E
ELV12	Marseille	12	43 ° 19.780'N	5 ° 14.210'E
MEJ40	Marseille	40	43 ° 19.700'N	5 ° 13.480'E
BANN20	Banyuls	25	42 ° 26.890'N	3 ° 10.330'E
BANN40	Banyuls	35	42 ° 26.890'N	3 ° 10.330'E
BANS20	Banyuls	26	42 ° 26.390'N	3 ° 10.790'E
BANS40	Banyuls	36	42 ° 26.390'N	3 ° 10.790'E
POR20	Corsica	21	42 ° 16.292'N	8 ° 41.255'E
POR40	Corsica	33	42 ° 16.292'N	8 ° 41.255'E
GAL20	Corsica	26	42 ° 28.210'N	8 ° 38.950'E
GAL40	Corsica	36	42 ° 28.210'N	8 ° 38.950'E

TABLE D.2 – p-values of the Wilcoxon–Mann–Whitney test of amplification differences between males and females (real-time PCR experiment).

Locus	P-values, Corsica individuals	P-values, Marseille individuals
Locus_139082	$3.48.10^{-8}$	$4.81.10^{-6}$
Locus_144530	$8.87.10^{-9}$	$6.11.10^{-7}$
Locus_147437	$1.98.10^{-7}$	$6.85.10^{-7}$
Locus_139282	$2.86.10^{-8}$	$6.11.10^{-7}$
Locus_158229	$2.41.10^{-11}$	$1.11.10^{-8}$
Locus_154834	$2.96.10^{-8}$	$5.46.10^{-7}$

TABLE D.3 – Counts of SNP loci after each step of filtering.

Step	Number of SNPs	Software
After assembly raw data	138 810	Stacks [57, 58]
Excluding loci not in within population HWE	86 520	VCFtools [87]
MAF 1 %	56 844	VCFtools [87]
One SNPs per RAD-tag	27461	

TABLE D.4 – Primers sequences of male-specific markers tested by real-time PCR.

Marker name	Primers sequences
Locus_139082	F TTCCAGGTATATGCCTTTCCATAA R GTGAGCCCAAATCCCTCAA
Locus_147437	F AAGATCAACTCGACACCACG R TGTACAGGCCAACAATAATTCCA
Locus_144530	F GCAGAGCCGTTTACCTTGAA R CCTCTCCGACTAACCATCC
Locus_139282	F GATTCTCCTGCTAGAGCAATTGTA R CCCTTCAAACCTCACACCAGTC
Locus_158229	F CTASTTTGTGCAATAAATGTGCCT R GTACAAGTCCTAGAGCTCAGAT
Locus_154834	F TGCAGATGCCATGCCTTAAAAAAT R TCCGTTTTTCGTTTTTTCAGGTGCAT

Références

- [1] F. Abascal, R. Zardoya, and D. Posada. ProtTest: selection of best-fit models of protein evolution. *Bioinformatics*, 21(9):2104–2105, 2005.
- [2] M. Abbiati, G. Santangelo, and S. Novelli. Genetic variation within and between two Tyrrhenian populations of the Mediterranean alcyonarian *Coralium rubrum*. *Marine Ecology Progress Series*, 95:245–250, 1993.
- [3] M. Alleaume-Benharira, I. R. Pen, and O. Ronce. Geographical patterns of adaptation within a species’ range: interactions between drift and gene flow. *Journal of evolutionary biology*, 19(1):203–215, 2006.
- [4] S. F. Altschul, W. Gish, W. Miller, E. W. Myers, and D. J. Lipman. Basic local alignment search tool. *Journal of Molecular Biology*, 215(3):403–410, 1990.
- [5] S. L. Ament-Velásquez, O. Breedy, J. Cortés, H. M. Guzman, G. Wörheide, and S. Vargas. Homoplasious colony morphology and mito-nuclear phylogenetic discordance among Eastern Pacific octocorals. *Molecular Phylogenetics and Evolution*, 98:373–381, 2016.
- [6] S. Anders and W. Huber. Differential expression analysis for sequence count data. *Genome Biology*, 11(10):R106, 2010.
- [7] C. Araneda, M. A. Larraín, B. Hecht, and S. Narum. Adaptive genetic variation distinguishes Chilean blue mussels (*Mytilus chilensis*) from different marine environments. *Ecology and Evolution*, 6(11):3632–3644, 2016.
- [8] R. H. Arch, R. W. Gedrich, and C. B. Thompson. Tumor necrosis factor receptor-associated factors (TRAFs)—a family of adapter proteins that regulates life and death. *Genes & Development*, 12(18):2821–2830, 1998.
- [9] R. Arizmendi-Mejía, J.-B. Ledoux, S. Civit, A. Antunes, Z. Thanopoulou, J. Garrabou, and C. Linares. Demographic responses to warming: reproduc-

- tive maturity and sex influence vulnerability in an octocoral. *Coral Reefs*, 34(4):1207–1216, 2015.
- [10] S. Arnaud-Haond, S. Teixeira, S. I. Massa, C. Billot, P. Saenger, G. Coupland, C. M. Duarte, and E. A. Serrão. Genetic structure at range edge: low diversity and high inbreeding in Southeast Asian mangrove (*Avicennia marina*) populations: genetic and inbreeding at range edge. *Molecular Ecology*, 15(12):3515–3525, 2006.
 - [11] M. Ashburner, C. A. Ball, J. A. Blake, D. Botstein, H. Butler, J. M. Cherry, A. P. Davis, K. Dolinski, S. S. Dwight, and J. T. Eppig. Gene Ontology: tool for the unification of biology. *Nature genetics*, 25(1):25–29, 2000.
 - [12] D. Aurelle, J.-B. Ledoux, C. Rocher, P. Borsa, A. Chenuil, and J.-P. Féral. Phylogeography of the red coral (*Corallium rubrum*): inferences on the evolutionary history of a temperate gorgonian. *Genetica*, 139(7):855–869, 2011.
 - [13] D. J. Ayre. Localized adaptation of sea anemone clones: evidence from transplantation over two spatial scales. *Journal of animal ecology*, (64):186–196, 1995.
 - [14] D. Bachtrog, J. E. Mank, C. L. Peichel, M. Kirkpatrick, S. P. Otto, T.-L. Ashman, M. W. Hahn, J. Kitano, I. Mayrose, R. Ming, N. Perrin, L. Ross, N. Valenzuela, J. C. Vamosi, and T. T. o. S. Consortium. Sex Determination: why so many ways of doing it? *PLOS Biol*, 12(7):e1001899, 2014.
 - [15] M. Bally and J. Garrahou. Thermodependent bacterial pathogens and mass mortalities in temperate benthic communities: a new case of emerging disease linked to climate change. *Global Change Biology*, 13(10):2078–2088, 2007.
 - [16] D. J. Barshis. Genomic potential for coral survival of climate change. In *Coral Reefs in the Anthropocene*, pages 133–146. Springer, 2015.
 - [17] D. J. Barshis, J. T. Ladner, T. A. Oliver, F. O. Seneca, N. Traylor-Knowles, and S. R. Palumbi. Genomic basis for coral resilience to climate change. *Proceedings of the National Academy of Sciences*, 110(4):1387–1392, 2013.
 - [18] D. J. Barshis, J. H. Stillman, R. D. Gates, R. J. Toonen, L. W. Smith, and C. Birkeland. Protein expression and genetic structure of the coral *Porites*

- lobata* in an environmentally extreme Samoan back reef: does host genotype limit phenotypic plasticity? *Molecular Ecology*, 19(8):1705–1720, 2010.
- [19] L. K. Bay, A. Guérécheau, N. Andreakis, K. E. Ulstrup, and M. V. Matz. Gene expression signatures of energetic acclimatisation in the reef building coral *Acropora millepora*. *PLoS ONE*, 8(5):e61736, 2013.
- [20] R. A. Bay and S. R. Palumbi. Multilocus adaptation associated with heat resistance in reef-building corals. *Current Biology*, 24(24):2952–2956, 2014.
- [21] M. A. Beaumont and D. J. Balding. Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, 13(4):969–980, 2004.
- [22] L. Benestan, B. K. Quinn, H. Maaroufi, M. Laporte, R. Rochette, and L. Bernatchez. Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in American lobster (*Homarus americanus*). *Molecular Ecology*, 2016.
- [23] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300, 1995.
- [24] P. R. Berg, S. Jentoft, B. Star, K. H. Ring, H. Knutsen, S. Lien, K. S. Jakobsen, and C. Andre. Adaptation to low salinity promotes genomic divergence in Atlantic cod (*Gadus morhua* L.). *Genome Biology and evolution*, 7(6):1644–1663, 2015.
- [25] R. Bergero and D. Charlesworth. The evolution of restricted recombination in sex chromosomes. *Trends in Ecology & Evolution*, 24(2):94–102, 2009.
- [26] R. Bergero, S. Qiu, A. Forrest, H. Borthwick, and D. Charlesworth. Expansion of the pseudo-autosomal region and ongoing recombination suppression in the *Silene latifolia* sex chromosomes. *Genetics*, 194(3):673–686, 2013.
- [27] G. Bernardi, E. Azzurro, D. Golani, and M. R. Miller. Genomic signatures of rapid adaptive evolution in the bluespotted cornetfish, a Mediterranean Lessepsian invader. *Molecular Ecology*, 2016.
- [28] E. A. Berntson, S. C. France, and L. S. Mullineaux. Phylogenetic relationships within the class Anthozoa (phylum Cnidaria) based on nuclear 18S

- rDNA sequences. *Molecular Phylogenetics and Evolution*, 13(2):417–433, 1999.
- [29] L. W. Beukeboom and N. Perrin. *The evolution of sex determination*. Oxford University Press, USA, 2014.
- [30] N. O. Bianchi. Akodon sex reversed females: the never ending story. *Cytogenetic and genome research*, 96(1-4):60–65, 2002.
- [31] N. Bierne, D. Roze, and J. J. Welch. Pervasive selection or is it...? why are F_{ST} outliers sometimes so frequent? *Molecular Ecology*, 22(8):2061–2064, 2013.
- [32] N. Bierne, J. Welch, E. Loire, F. Bonhomme, and P. David. The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Molecular Ecology*, 20(10):2044–2072, 2011.
- [33] W. C. Black IV, C. F. Baer, M. F. Antolin, and N. M. DuTeau. Population genomics: genome-wide sampling of insect populations. *Annual review of entomology*, 46(1):441–469, 2001.
- [34] J. Boavida, D. Paulo, D. Aurelle, S. Arnaud-Haond, C. Marschal, J. Reed, J. M. S. Goncalves, and E. A. Serrao. A well-kept treasure at depth: precious red coral rediscovered in atlantic deep coral gardens (SW Portugal) after 300 years. *Plos One*, 11(1):e0147228, 2016. WOS:000368655300077.
- [35] D. I. Bolnick and P. Nosil. Natural selection in populations subject to a migration load. *Evolution*, 61(9):2229–2243, 2007.
- [36] P. Bongaerts, C. Riginos, K. B. Hay, M. J. van Oppen, O. Hoegh-Guldberg, and S. Dove. Adaptive divergence in a scleractinian coral: physiological adaptation of *Seriatopora hystrix* to shallow and deep reef habitats. *BMC evolutionary biology*, 11(1):303, 2011.
- [37] A. Bonin, P. Taberlet, C. Miaud, and F. Pompanon. Explorative genome scan to detect candidate loci for adaptation along a gradient of altitude in the common frog (*Rana temporaria*). *Molecular Biology and Evolution*, 23(4):773–783, 2006.
- [38] V. Bourret, M. Dionne, and L. Bernatchez. Detecting genotypic changes associated with selective mortality at sea in Atlantic salmon: polygenic multi-

- locus analysis surpasses genome scan. *Molecular Ecology*, 23(18):4444–4457, 2014.
- [39] I. R. Bradbury, L. C. Hamilton, B. Dempson, M. J. Robertson, V. Bourret, L. Bernatchez, and E. Verspoor. Transatlantic secondary contact in Atlantic Salmon, comparing microsatellites, a single nucleotide polymorphism array and restriction-site associated DNA sequencing for the resolution of complex spatial structure. *Molecular Ecology*, 24(20):5130–5144, 2015.
- [40] I. R. Bradbury, S. Hubert, B. Higgins, T. Borza, S. Bowman, I. G. Paterson, P. V. Snelgrove, C. J. Morris, R. S. Gregory, D. C. Hardie, and others. Parallel adaptive evolution of Atlantic cod on both sides of the Atlantic Ocean in response to temperature. *Proceedings of the Royal Society of London B: Biological Sciences*, 277(1701):3725–3734, 2010.
- [41] J. R. Bradley and J. S. Pober. Tumor necrosis factor receptor-associated factors (TRAFs). *Oncogene*, 20(44), 2001.
- [42] L. Bramanti, J. Movilla, M. Guron, E. Calvo, A. Gori, C. Dominguez-Carrió, J. Grinyó, A. Lopez-Sanz, A. Martinez-Quintana, C. Pelejero, and others. Detrimental effects of ocean acidification on the economically important Mediterranean red coral (*Corallium rubrum*). *Global Change Biology*, 19(6):1897–1908, 2013.
- [43] D. Bridge, C. W. Cunningham, R. DeSalle, and L. W. Buss. Class-level relationships in the phylum Cnidaria: molecular and morphological evidence. *Molecular Biology and Evolution*, 12(4):679–689, 1995.
- [44] L. Brocchieri. Phylogenetic inferences from molecular sequences: review and critique. *Theoretical Population Biology*, 59(1):27–40, 2001.
- [45] J. E. Brommer. Whither P_{ST} ? The approximation of Q_{ST} by P_{ST} in evolutionary and conservation biology. *Journal of evolutionary biology*, 24(6):1160–1168, 2011.
- [46] B. Brown, R. Dunne, M. Goodson, and A. Douglas. Experience shapes the susceptibility of a reef coral to bleaching. *Coral Reefs*, 21(2):119–126, 2002.
- [47] J. F. Bruno and P. J. Edmunds. Clonal variation for phenotypic plasticity in the coral *Madracis mirabilis*. *Ecology*, 78(7):2177–2190, 1997.

- [48] P. F. Brussard. Geographic patterns and environmental gradients: the central-marginal model in *Drosophila* revisited. *Annual Review of Ecology and Systematics*, (15):25–64, 1984.
- [49] C. A. Burge, M. E. Mouchka, C. D. Harvell, and S. Roberts. Immune response of the Caribbean sea fan, *Gorgonia ventalina*, exposed to an *Aplanochytrium* parasite as revealed by transcriptome sequencing. *Frontiers in Physiology*, 4, 2013.
- [50] R. S. Burton. Evolutionary consequences of restricted gene flow among natural populations of the copepod, *Trigriopus californicus*. *Bulletin of Marine Science*, 39(2):526–535, 1986.
- [51] R. S. Burton and M. W. Feldman. Physiological effects of an allozyme polymorphism: glutamate-pyruvate transaminase and response to hyperosmotic stress in the copepod *Tigriopus californicus*. *Biochemical genetics*, 21(3-4):239–251, 1983.
- [52] I. Calderon, J. Garrabou, and D. Aurelle. Evaluation of the utility of COI and ITS markers as tools for population genetic studies of temperate gorgonians. *Journal of Experimental Marine Biology and Ecology*, 336(2):184–197, 2006.
- [53] R. Cannas, F. Sacco, A. Cau, E. Coluccia, M. C. Follesa, and A. Cau. New insights into connectivity patterns of mesophotic red coral (*Corallium rubrum*) populations. *Hydrobiologia*, 759(1):63–73, 2015.
- [54] R. Cannas, F. Sacco, A. Cau, D. Cuccu, M. C. Follesa, and A. Cau. Genetic monitoring of deep-water exploited banks of the precious Sardinia coral *Corallium rubrum* (L., 1758): useful data for a sustainable management: genetic monitoring of exploited banks of the precious Sardinia coral. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 26(2):236–250, 2016.
- [55] S. N. Carmichael, M. Bekaert, J. B. Taggart, H. R. Christie, D. I. Bassett, J. E. Bron, P. J. Skuce, K. Gharbi, R. Skern-Mauritzen, and A. Sturm. Identification of a sex-linked SNP marker in the salmon louse (*Lepeophtheirus salmonis*) using RAD sequencing. *PloS one*, 8(10):e77832, 2013.
- [56] J. Castresana. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. *Molecular Biology and Evolution*, 17(4):540–552, 2000.

- [57] J. Catchen, P. A. Hohenlohe, S. Bassham, A. Amores, and W. A. Cresko. Stacks: an analysis tool set for population genomics. *Molecular Ecology*, 22(11):3124–3140, 2013.
- [58] J. M. Catchen, A. Amores, P. Hohenlohe, W. Cresko, and J. H. Postlethwait. Stacks: building and genotyping loci de novo from short-read sequences. *G3: Genes, Genomes, Genetics*, 1(3):171–182, 2011.
- [59] T. Cavalier-Smith, M. T. E. P. Allsopp, E. E. Chao, N. Boury-Esnault, and J. Vacelet. Sponge phylogeny, animal monophyly, and the origin of the nervous system: 18s rRNA evidence. *Canadian Journal of Zoology*, 74(11):2031–2045, 1996.
- [60] J. Chaline. *Quoi de neuf depuis Darwin?: la théorie de l'évolution des espèces dans tous ses états*. Ellipses, 2006.
- [61] E. S. Chang, M. Neuhof, N. D. Rubinstein, A. Diamant, H. Philippe, D. Huchon, and P. Cartwright. Genomic insights into the evolutionary origin of Myxozoa within Cnidaria. *Proceedings of the National Academy of Sciences*, 112(48):14912–14917, 2015.
- [62] C. A. Chen, D. M. Odorico, M. Tenlohuis, J. E. N. Veron, and D. J. Miller. Systematic relationships within the Anthozoa (Cnidaria: Anthozoa) using the 5'-end of the 28s rDNA. *Molecular Phylogenetics and Evolution*, 4(2):175–183, 1995.
- [63] C.-J. Chen, S. Shikina, W.-J. Chen, Y.-J. Chung, Y.-L. Chiu, J. A. Bertrand, Y.-H. Lee, and C.-F. Chang. A novel female-specific and sexual reproduction-associated Dmrt gene discovered in the stony coral, *Euphyllia ancora*. *Biology of Reproduction*, 94(2):1–13, 2016.
- [64] I.-P. Chen, C.-Y. Tang, C.-Y. Chiou, J.-H. Hsu, N. V. Wei, C. C. Wallace, P. Muir, H. Wu, and C. A. Chen. Comparative analyses of coding and noncoding DNA regions indicate that *Acropora* (Anthozoa: Scleractina) possesses a similar evolutionary tempo of nuclear vs. mitochondrial genomes as in plants. *Marine Biotechnology*, 11(1):141–152, 2009.
- [65] E. Cherif, S. Zehdi, K. Castillo, N. Chabrillange, S. Abdoukader, J.-C. Pintaud, S. Santoni, A. Salhi-Hannachi, S. Glémin, and F. Aberlenc-Bertossi. Male-specific DNA markers provide genetic evidence of an XY chromosome

- system, a recombination arrest and allow the tracing of paternal lineages in date palm. *New Phytologist*, 197(2):409–415, 2013.
- [66] N. D. Chu, S. T. Kaluziak, G. C. Trussell, and S. V. Vollmer. Phylogenomic analyses reveal latitudinal population structure and polymorphisms in heat stress genes in the North Atlantic snail *Nucella lapillus*. *Molecular Ecology*, 23(7):1863–1873, 2014.
 - [67] A. G. Collins. Evaluating multiple alternative hypotheses for the origin of Bilateria: an analysis of 18s rRNA molecular evidence. *Proceedings of the National Academy of Sciences*, 95(26):15458–15463, 1998.
 - [68] A. G. Collins. Phylogeny of Medusozoa and the evolution of cnidarian life cycles. *Journal of Evolutionary Biology*, 15(3):418–432, 2002.
 - [69] A. G. Collins. Recent insights into cnidarian phylogeny. *Smithsonian Contributions to Marine Sciences*, 38:139–149, 2009.
 - [70] A. G. Collins, P. Schuchert, A. C. Marques, T. Jankowski, M. Medina, and B. Schierwater. Medusozoan phylogeny and character evolution clarified by new large and small subunit rDNA data and an assessment of the utility of phylogenetic mixture models. *Systematic Biology*, 55(1):97–115, 2006.
 - [71] A. Conesa, S. Götz, J. M. García-Gómez, J. Terol, M. Talón, and M. Robles. Blast2go: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, 21(18):3674–3676, 2005.
 - [72] D. O. Conover, L. M. Clarke, S. B. Munch, and G. N. Wagner. Spatial and temporal scales of adaptive divergence in marine fishes and the implications for conservation. *Journal of Fish Biology*, 69(sc):21–47, 2006.
 - [73] G. Coop, D. Witonsky, A. D. Rienzo, and J. K. Pritchard. Using environmental correlations to identify loci underlying local adaptation. *Genetics*, 185(4):1411–1423, 2010.
 - [74] F. Costantini and M. Abbiati. Into the depth of population genetics: pattern of structuring in mesophotic red coral populations. *Coral Reefs*, 35(1):39–52, 2016.
 - [75] F. Costantini, L. Carlesi, and M. Abbiati. Quantifying spatial genetic structuring in mesophotic populations of the precious coral *Corallium rubrum*. *PloS one*, 8(4):e61546, 2013.

- [76] F. Costantini, C. Fauvelot, and M. Abbiati. Fine-scale genetic structuring in *Corallium rubrum*: evidence of inbreeding and limited effective larval dispersal. *Marine Ecology Progress Series*, 340:109–119, 2007.
- [77] F. Costantini, C. Fauvelot, and M. Abbiati. Genetic structuring of the temperate gorgonian coral (*Corallium rubrum*) across the western Mediterranean Sea revealed by microsatellites and nuclear sequences. *Molecular Ecology*, 16(24):5168–5182, 2007.
- [78] F. Costantini, A. Gori, P. Lopez-González, L. Bramanti, S. Rossi, J.-M. Gili, and M. Abbiati. Limited genetic connectivity between gorgonian morphotypes along a depth gradient. *PloS one*, 11(8):e0160678, 2016.
- [79] F. Costantini, S. Rossi, E. Pintus, C. Cerrano, J.-M. Gili, and M. Abbiati. Low connectivity and declining genetic variability along a depth gradient in *Corallium rubrum* populations. *Coral reefs*, 30(4):991–1003, 2011.
- [80] F. Costantini, M. Taviani, A. Remia, E. Pintus, P. J. Schembri, and M. Abbiati. Deep-water *Corallium rubrum* (L., 1758) from the Mediterranean Sea: preliminary genetic characterisation. *Marine Ecology*, 31(2):261–269, 2010.
- [81] F. Crocetta and M. Spanu. Molluscs associated with a Sardinian deep water population of *Corallium rubrum* (Linné, 1758). *Mediterranean Marine Science*, 9(2):63–86, 2008.
- [82] J. F. Crow, M. Kimura, and others. An introduction to population genetics theory. *An introduction to population genetics theory.*, 1970.
- [83] F. Cunningham, M. R. Amode, D. Barrell, K. Beal, K. Billis, S. Brent, D. Carvalho-Silva, P. Clapham, G. Coates, S. Fitzgerald, L. Gil, C. G. Girón, L. Gordon, T. Hourlier, S. E. Hunt, S. H. Janacek, N. Johnson, T. Juettemann, A. K. Kähäri, S. Keenan, F. J. Martin, T. Maurel, W. McLaren, D. N. Murphy, R. Nag, B. Overduin, A. Parker, M. Patricio, E. Perry, M. Pignatelli, H. S. Riat, D. Sheppard, K. Taylor, A. Thormann, A. Vullo, S. P. Wilder, A. Zadissa, B. L. Aken, E. Birney, J. Harrow, R. Kinsella, M. Muffato, M. Ruffier, S. M. J. Searle, G. Spudich, S. J. Trevanion, A. Yates, D. R. Zerbino, and P. Flicek. Ensembl 2015. *Nucleic Acids Research*, 43(D1):D662–D669, 2015.

- [84] M. Daly, M. R. Brugler, P. Cartwright, A. G. Collins, M. N. Dawson, D. G. Fautin, S. C. France, C. McFadden, D. M. Opresko, E. Rodriguez, S. L. Romano, and J. L. Stake. The phylum Cnidaria: a review of phylogenetic patterns and diversity 300 years after Linnaeus. *Zootaxa*, (1668):127–182, 2007.
- [85] M. Daly, D. G. Fautin, and V. A. Cappola. Systematics of the hexacorallia (Cnidaria: Anthozoa). *Zoological Journal of the Linnean Society*, 139(3):419–437, 2003.
- [86] E. Danchin, A. Charmantier, F. A. Champagne, A. Mesoudi, B. Pujol, and S. Blanchet. Beyond DNA: integrating inclusive inheritance into an extended theory of evolution. *Nature Reviews Genetics*, 12(7):475–486, 2011.
- [87] P. Danecek, A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, and R. Durbin. The variant call format and VCFtools. *Bioinformatics*, 27(15):2156–2158, 2011.
- [88] C. Darwin. *The origin of species by means of natural selection: or, the preservation of favoured races in the struggle for life and the descent of man and selection in relation to sex*. Modern library, 1872.
- [89] S. K. Davy, D. Allemand, and V. M. Weis. Cell Biology of Cnidarian-Dinoflagellate Symbiosis. *Microbiology and Molecular Biology Reviews*, 76(2):229–261, 2012.
- [90] H. de Lacaze-Duthiers. *Histoire naturelle du corail: organisation, reproduction, pêche en Algérie, industrie et commerce*. J.-B. Baillière et fils, 1864.
- [91] P. de Villemereuil, E. Frichot, E. Bazin, O. François, and O. E. Gaggiotti. Genome scan methods against more complex models: when and how much should we trust them? *Molecular Ecology*, 23(8):2006–2019, 2014.
- [92] P. de Villemereuil and O. E. Gaggiotti. A new F_{ST} -based method to uncover local adaptation using environmental variables. *Methods in Ecology and Evolution*, 6(11):1248–1258, 2015.
- [93] P. De Wit and S. R. Palumbi. Transcriptome-wide polymorphisms of red abalone (*Haliotis rufescens*) reveal patterns of gene flow and local adaptation. *Molecular Ecology*, 22(11):2884–2897, 2013.

- [94] D. Del Gaudio, G. Fortunato, M. Borriello, J. M. Gili, P. Buono, G. Calcagno, F. Salvatore, and L. Sacchetti. Genetic typing of *Corallium rubrum*. *Marine Biotechnology*, 6(6):511–515, 2004.
- [95] M. A. DePristo, E. Banks, R. Poplin, K. V. Garimella, J. R. Maguire, C. Hartl, A. A. Philippakis, G. del Angel, M. A. Rivas, M. Hanna, A. McKenna, T. J. Fennell, A. M. Kernysky, A. Y. Sivachenko, K. Cibulskis, S. B. Gabriel, D. Altshuler, and M. J. Daly. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, 43(5):491–498, 2011.
- [96] M. DeSalvo, S. Sunagawa, C. Voolstra, and M. Medina. Transcriptomic responses to heat stress and bleaching in the elkhorn coral *Acropora palmata*. *Marine Ecology Progress Series*, 402:97–113, 2010.
- [97] T. J. DeWitt, A. Sih, and D. S. Wilson. Costs and limits of phenotypic plasticity. *Trends in ecology & evolution*, 13(2):77–81, 1998.
- [98] G. B. Dixon, L. K. Bay, and M. V. Matz. Bimodal signatures of germline methylation are linked with gene expression plasticity in the coral *Acropora millepora*. *BMC Genomics*, 15:1109, 2014.
- [99] S. C. Doney, M. Ruckelshaus, J. E. Duffy, J. P. Barry, F. Chan, C. A. English, H. M. Galindo, J. M. Grebmeier, A. B. Hollowed, N. Knowlton, and others. Climate change impacts on marine ecosystems. *Marine Science*, 4, 2012.
- [100] F. D’Ortenzio, H. Lavigne, F. Besson, H. Claustre, L. Coppola, N. Garcia, A. Laës-Huon, S. Le Reste, D. Malardé, C. Migon, P. Morin, L. Mortier, A. Poteau, L. Prieur, P. Raimbault, and P. Testor. Observing mixed layer depth, nitrate and chlorophyll concentrations in the northwestern Mediterranean: a combined satellite and NO₃ profiling floats experiment. *Geophysical Research Letters*, 41(18):6443–6451, 2014.
- [101] S. Dupont, N. Dorey, M. Stumpp, F. Melzner, and M. Thorndyke. Long-term and trans-life-cycle effects of exposure to ocean acidification in the green sea urchin *Strongylocentrotus droebachiensis*. *Marine Biology*, 160(8):1835–1843, 2013.
- [102] L. D’Croz and J. L. Maté. Experimental responses to elevated water temperature in genotypes of the reef coral *Pocillopora damicornis* from upwelling

- and non-upwelling environments in Panama. *Coral Reefs*, 23(4):473–483, 2004.
- [103] P. Edelaar and D. I. Bolnick. Non-random gene flow: an underappreciated force in evolution and ecology. *Trends in ecology & evolution*, 27(12):659–665, 2012.
 - [104] P. Edelaar, A. M. Siepielski, and J. Clobert. Matching habitat choice causes directed gene flow: a neglected dimension in evolution and ecology. *Evolution*, 62(10):2462–2472, 2008.
 - [105] R. C. Edgar. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, 32(5):1792–1797, 2004.
 - [106] P. J. Edmunds. Evidence that reef-wide patterns of coral bleaching may be the result of the distribution of bleaching-susceptible clones. *Marine Biology*, 121(1):137–142, 1994.
 - [107] P. D. Etter, S. Bassham, P. A. Hohenlohe, E. A. Johnson, and W. A. Cresko. SNP discovery and genotyping for evolutionary genetics using RAD sequencing. *Molecular methods for evolutionary genetics*, pages 157–178, 2011.
 - [108] G. Evanno, S. Regnaut, and J. Goudet. Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, 14(8):2611–2620, 2005.
 - [109] L. Excoffier, T. Hofer, and M. Foll. Detecting loci under selection in a hierarchically structured population. *Heredity*, 103(4):285–298, 2009.
 - [110] L. Excoffier and H. E. L. Lischer. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, 10(3):564–567, 2010.
 - [111] D. Falush, M. Stephens, and J. K. Pritchard. Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, 164(4):1567–1587, 2003.
 - [112] D. Falush, M. Stephens, and J. K. Pritchard. Inference of population structure using multilocus genotype data: dominant markers and null alleles. *Molecular Ecology notes*, 7(4):574–578, 2007.
 - [113] J. Felsenstein. Genetic drift in clines which are maintained by migration and natural-selection. *Genetics*, 81(1):191–207, 1975. WOS:A1975AV05400012.

- [114] C. Ferrier-Pagès, E. Tambutté, T. Zamoum, N. Segonds, P.-L. Merle, N. Bensoussan, D. Allemand, J. Garrabou, and S. Tambutté. Physiological response of the symbiotic gorgonian *Eunicella singularis* to a long-term temperature increase. *Journal of Experimental Biology*, 212(18):3007–3015, 2009.
- [115] D. F. Figueroa and A. R. Baco. Octocoral mitochondrial genomes provide insights into the phylogenetic history of gene order rearrangements, order reversals, and cnidarian phylogenetics. *Genome Biology and Evolution*, 7(1):391–409, 2014.
- [116] M. Foll and O. Gaggiotti. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a bayesian perspective. *Genetics*, 180(2):977–993, 2008.
- [117] M. Fourment and M. J. Gibbs. PATRISTIC: a program for calculating patristic distances and graphically comparing the components of genetic change. *BMC Evolutionary Biology*, 6(1):1, 2006.
- [118] B. L. Fowler and V. P. Buonaccorsi. Genomic characterization of sex-identification markers in *Sebastes carnatus* and *S. chrysomelas* rockfishes. *Molecular Ecology*, 2016.
- [119] S. C. France and L. L. Hoover. DNA sequences of the mitochondrial COI gene have low levels of divergence among deep-sea octocorals (Cnidaria: Anthozoa). *Hydrobiologia*, 471(1-3):149–155, 2002.
- [120] S. C. France, P. E. Rosel, J. E. Agenbroad, L. S. Mullineaux, and T. D. Kocher. DNA sequence variation of mitochondrial large-subunit rRNA provides support for a two-subclass organization of the Anthozoa (Cnidaria). *Molecular Marine Biology and Biotechnology*, 5(1):15–28, 1996.
- [121] S. Fraune, S. Forêt, and A. M. Reitzel. Using *Nematostella vectensis* to study the interactions between genome, epigenome, and bacteria in a changing environment. *Microbial Symbioses*, page 148, 2016.
- [122] E. Frichot, S. D. Schoville, G. Bouchard, and O. François. Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular biology and evolution*, 30(7):1687–1699, 2013.

- [123] E. Frichot, S. D. Schoville, P. de Villemereuil, O. E. Gaggiotti, and O. François. Detecting adaptive evolution based on association with ecological gradients: orientation matters ! *Heredity*, 115(1):22–28, 2015.
- [124] P.-A. Gagnaire and O. E. Gaggiotti. Detecting polygenic selection in marine populations by combining population genomics and quantitative genetics approaches. *Current Zoology*, page zow088, 2016.
- [125] N. Galtier and V. Daubin. Dealing with incongruence in phylogenomic analyses. *Philosophical Transactions of the Royal Society of London B: Biological Sciences*, 363(1512):4023–4029, 2008.
- [126] T. Gamble. Using RAD-seq to recognize sex-specific markers and sex chromosome systems. *Molecular Ecology*, 25(10):2114–2116, 2016.
- [127] T. Gamble, J. Coryell, T. Ezaz, J. Lynch, D. P. Scantlebury, and D. Zarkower. Restriction site-associated DNA sequencing (RAD-seq) reveals an extraordinary number of transitions among gecko sex-determining systems. *Molecular biology and evolution*, page msv023, 2015.
- [128] T. Gamble and D. Zarkower. Identification of sex-specific molecular markers using restriction site-associated DNA sequencing. *Molecular Ecology Resources*, 14(5):902–913, 2014.
- [129] D. Garant, S. E. Forde, and A. P. Hendry. The multifarious effects of dispersal and gene flow on contemporary adaptation. *Functional Ecology*, 21(3):434–443, 2007.
- [130] M. García-Rodríguez and C. Massò. Algunas bases para la determinación directa de la edad del coral rojo (*Corallium rubrum*, L). 1986.
- [131] J. Garrabou, R. Coma, N. Bensoussan, M. Bally, P. Chevaldonné, M. Cigliano, D. Diaz, J. G. Harmelin, M. C. Gambi, D. K. Kersting, J. B. Ledoux, C. Lejeusne, C. Linares, C. Marschal, T. Pérez, M. Ribes, J. C. Romano, E. Serrano, N. Teixido, O. Torrents, M. Zabala, F. Zuberer, and C. Cerrano. Mass mortality in Northwestern Mediterranean rocky benthic communities: effects of the 2003 heat wave. *Global Change Biology*, 15(5):1090–1103, 2009.
- [132] J. Garrabou and J. G. Harmelin. A 20-year study on life-history traits of a harvested long-lived temperate coral in the NW Mediterranean: insights into

- conservation and management needs. *Journal of Animal Ecology*, 71(6):966–978, 2002.
- [133] J. Garrabou, T. Perez, S. Sartoretto, and J. G. Harmelin. Mass mortality event in red coral *Corallium rubrum* populations in the Provence region (France, NW Mediterranean). *Marine Ecology Progress Series*, 217:263–272, 2001.
- [134] J. Garrabou, E. Sala, A. Arcas, and M. Zabala. The impact of diving use on marine rocky sublittoral communities: a case study of a Mediterranean bryozoan population. *Conservation Biology*, 12(2):1–12, 1998.
- [135] J. Garrabou and D. Team. Is global change a real threat for conservation of the NW Mediterranean marine biodiversity? In *EGS-AGU-EUG Joint Assembly*, volume 1, page 10522, 2003.
- [136] M. Gautier, K. Gharbi, T. Cezard, J. Foucaud, C. Kerdelhué, P. Pudlo, J.-M. Cornuet, and A. Estoup. The effect of RAD allele dropout on the estimation of genetic variation within and between populations. *Molecular Ecology*, 22(11):3165–3178, 2013.
- [137] L. U. Gleason and R. S. Burton. Genomic evidence for ecological divergence against a background of population homogeneity in the marine snail *Chlorostoma funebris*. *Molecular Ecology*, 25(15):3557–3573, 2016.
- [138] J. P. Gogarten, W. F. Doolittle, and J. G. Lawrence. Prokaryotic evolution in light of gene transfer. *Molecular biology and evolution*, 19(12):2226–2238, 2002.
- [139] R. Gomulkiewicz, R. D. Holt, and M. Barfield. The effects of density dependence and immigration on local adaptation and niche evolution in a black-hole sink environment. *Theoretical population biology*, 55(3):283–296, 1999.
- [140] G. Guillot. Detection of correlation between genotypes and environmental variables. A fast computational approach for genomewide studies. *arXiv preprint arXiv:1206.0889*, 2012.
- [141] B. Guo, Z. Li, and J. Merilä. Population genomic evidence for adaptive differentiation in the Baltic Sea herring. *Molecular Ecology*, 25(12):2833–2852, 2016.

- [142] T. Günther and G. Coop. Robust identification of local adaptation from allele frequencies. *Genetics*, 195(1):205–220, 2013.
- [143] A. Haguénauer, F. Zuberer, J.-B. Ledoux, and D. Aurelle. Adaptive abilities of the Mediterranean red coral *Corallium rubrum* in a heterogeneous and changing environment: from population to functional genetics. *Journal of Experimental Marine Biology and Ecology*, 449:349–357, 2013.
- [144] J. G. Harmelin. Environnement thermique du benthos côtier de l’île de Port-Cros (parc national, France, Méditerranée nord-occidentale) et implications biogéographiques. *Scientific reports of Port-Cros national park*, 20:97–146, 2004.
- [145] A. Hastings and F. J. Rohlf. Gene flow: effect in stochastic models of differentiation. *The American Naturalist*, 108(963):701–705, 1974.
- [146] T. Heams, P. Huneman, G. Lecointre, and M. Silberstein. *Les mondes darwiniens. L’évolution de l’évolution*. Editions Syllepse, Paris, 1re edition, 2009.
- [147] P. W. Hedrick, M. E. Ginevan, and E. P. Ewing. Genetic polymorphism in heterogeneous environments. *Annual Review of Ecology and Systematics*, pages 1–32, 1976.
- [148] T. Hilbish. Demographic and temporal structure of an allele frequency cline in the mussel *Mytilus Edulis*. *Marine Biology*, 86(2):163–171, 1985. WOS:A1985AJR4700009.
- [149] S. Hoban, J. L. Kelley, K. E. Lotterhos, M. F. Antolin, G. Bradburd, D. B. Lowry, M. L. Poss, L. K. Reed, A. Storfer, M. C. Whitlock, C. F. Baer, and A. A. Winn. Finding the genomic basis of local adaptation: pitfalls, practical solutions, and future directions. *The American Naturalist*, 2016.
- [150] A. Hobolth, O. F. Christensen, T. Mailund, and M. H. Schierup. Genomic relationships and speciation times of human, chimpanzee, and gorilla inferred from a coalescent hidden Markov model. *PLoS Genet*, 3(2):e7, 2007.
- [151] O. Hoegh-Guldberg, P. J. Mumby, A. J. Hooten, R. S. Steneck, P. Greenfield, E. Gomez, C. D. Harvell, P. F. Sale, A. J. Edwards, K. Caldeira, and others. Coral reefs under rapid climate change and ocean acidification. *Science*, 318(5857):1737–1742, 2007.

- [152] T. Hofer, N. Ray, D. Wegmann, and L. Excoffier. Large allele frequency differences between human continental groups are more likely to have occurred by drift during range expansions than by selection. *Annals of human genetics*, 73(1):95–108, 2009.
- [153] A. A. Hoffmann and M. W. Blows. Species borders: ecological and evolutionary perspectives. *Trends in Ecology & Evolution*, 9(6):223–227, 1994.
- [154] P. A. Hohenlohe, S. Bassham, P. D. Etter, N. Stiffler, E. A. Johnson, and W. A. Cresko. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet*, 6(2):e1000862, 2010.
- [155] E. J. Howells, D. Abrego, E. Meyer, N. L. Kirk, and J. A. Burt. Host adaptation and unexpected symbiont partners enable reef-building corals to tolerate extreme temperatures. *Global change biology*, 2016.
- [156] M. J. Hubisz, D. Falush, M. Stephens, and J. K. Pritchard. Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources*, 9(5):1322–1332, 2009.
- [157] T. P. Hughes, A. H. Baird, D. R. Bellwood, M. Card, S. R. Connolly, C. Folke, R. Grosberg, O. Hoegh-Guldberg, J. B. C. Jackson, J. Kleypas, and others. Climate change, human impacts, and the resilience of coral reefs. *science*, 301(5635):929–933, 2003.
- [158] B. C. C. Hume, C. R. Voolstra, C. Arif, C. D’Angelo, J. A. Burt, G. Eyal, Y. Loya, and J. Wiedenmann. Ancestral genetic diversity associated with the rapid spread of stress-tolerant coral symbionts in response to Holocene climate change. *Proceedings of the National Academy of Sciences*, 113(16):4416–4421, 2016.
- [159] D. J. Rankin and A. López-Sepulcre. Can adaptation lead to extinction? *Oikos*, 111(3):616–619, 2005.
- [160] Y. K. Jin, P. Lundgren, A. Lutz, J.-B. Raina, E. J. Howells, A. S. Paley, B. L. Willis, and M. J. H. v. Oppen. Genetic markers for antioxidant capacity in a reef-building coral. *Science Advances*, 2(5):e1500842, 2016.
- [161] T. Jombart. adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11):1403–1405, 2008.

- [162] F. C. Jones, M. G. Grabherr, Y. F. Chan, P. Russell, E. Mauceli, J. Johnson, R. Swofford, M. Pirun, M. C. Zody, S. White, E. Birney, S. Searle, J. Schmutz, J. Grimwood, M. C. Dickson, R. M. Myers, C. T. Miller, B. R. Summers, A. K. Knecht, S. D. Brady, H. Zhang, A. A. Pollen, T. Howes, C. Amemiya, Broad Institute Genome Sequencing Platform & Whole Genome Assembly Team, E. S. Lander, F. Di Palma, K. Lindblad-Toh, and D. M. Kingsley. The genomic basis of adaptive evolution in three-spine sticklebacks. *Nature*, 484(7392):55–61, 2012.
- [163] S. Kafkas, M. Khodaeiaminjan, M. Güney, and E. Kafkas. Identification of sex-linked SNP markers using RAD sequencing suggests ZW/ZZ sex determination in *Pistacia vera* L. *BMC genomics*, 16(1):1, 2015.
- [164] S. E. Kahng, Y. Benayahu, and H. R. Lasker. Sexual reproduction in octocorals. *Marine Ecology Progress Series*, 443:265–283, 2011.
- [165] S. Karako-Lampert, D. Zoccola, M. Salmon-Divon, M. Katzenellenbogen, S. Tambutté, A. Bertucci, O. Hoegh-Guldberg, E. Deleury, D. Allemand, and O. Levy. Transcriptome analysis of the scleractinian coral *Stylophora pistillata*. *PLoS ONE*, 9(2):e88615, 2014.
- [166] T. J. Kawecki and D. Ebert. Conceptual issues in local adaptation. *Ecology Letters*, 7(12):1225–1241, 2004.
- [167] E. Kayal and D. V. Lavrov. The mitochondrial genome of *Hydra oligactis* (Cnidaria, Hydrozoa) sheds new light on animal mtDNA evolution and cnidarian phylogeny. *Gene*, 410(1):177–186, 2008.
- [168] E. Kayal, B. Roure, H. Philippe, A. G. Collins, and D. V. Lavrov. Cnidarian phylogenetic relationships as revealed by mitogenomics. *BMC Evolutionary Biology*, 13(1):5, 2013.
- [169] C. Kenkel and M. V. Matz. Enhanced gene expression plasticity as a mechanism of adaptation to a variable environment in a reef-building coral. Technical Report biorxiv;059667v1, 2016.
- [170] C. Kenkel, C. Sheridan, M. C. Leal, R. Bhagooli, K. D. Castillo, N. Kurata, E. McGinty, T. Goulet, and M. Matz. Diagnostic gene expression biomarkers of coral thermal stress. *Molecular Ecology Resources*, 14(4):667–678, 2014.

- [171] A. M. Kerr, A. H. Baird, and T. P. Hughes. Correlated evolution of sex and reproductive mode in corals (Anthozoa: Scleractinia). *Proceedings of the Royal Society of London B: Biological Sciences*, 278(1702):75–81, 2011.
- [172] D. K. Kersting, N. Bensoussan, and C. Linares. Long-term responses of the endemic reef-builder *Cladocora caespitosa* to Mediterranean warming. *PLoS ONE*, 8(8):e70820, 2013.
- [173] J. Kim, W. Kim, and C. W. Cunningham. A new perspective on lower metazoan relationships from 18s rDNA sequences. *Molecular Biology and Evolution*, 16(3):423–427, 1999.
- [174] N. E. Kimes, C. J. Grim, W. R. Johnson, N. A. Hasan, B. D. Tall, M. H. Kothary, H. Kiss, A. C. Munk, R. Tapia, L. Green, and others. Temperature regulation of virulence factors in the pathogen *Vibrio coralliilyticus*. *The ISME journal*, 6(4):835–846, 2012.
- [175] E. Kisdi. Dispersal: risk spreading versus local adaptation. *The American Naturalist*, 159(6):579–596, 2002.
- [176] H. Kishino and M. Hasegawa. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. *Journal of Molecular Evolution*, 29(2):170–179, 1989.
- [177] K. M. Kocot, M. R. Citarella, L. L. Moroz, and K. M. Halanych. PhyloTreePruner: a phylogenetic tree-based approach for selection of orthologous sequences for phylogenomics. *Evolutionary Bioinformatics Online*, 9:429–435, 2013.
- [178] N. M. Kopelman, J. Mayzel, M. Jakobsson, N. A. Rosenberg, and I. Mayrose. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Molecular Ecology Resources*, 15(5):1179–1191, 2015.
- [179] E. C. E. Kvennefors, W. Leggat, O. Hoegh-Guldberg, B. M. Degnan, and A. C. Barnes. An ancient and variable mannose-binding lectin from the coral *Acropora millepora* binds both pathogens and symbionts. *Developmental & Comparative Immunology*, 32(12):1582–1592, 2008.
- [180] J.-B.-P. A. Lamarck. *Philosophie Zoologique*. 1809.

- [181] R. Lande and S. Shannon. The role of genetic variation in adaptation and population persistence in a changing environment. *Evolution*, 50(1):434–437, 1996.
- [182] B. R. Larget, S. K. Kotha, C. N. Dewey, and C. Ané. BUCKy: Gene tree/species tree reconciliation with Bayesian concordance analysis. *Bioinformatics*, 26(22):2910–2911, 2010.
- [183] J.-B. Ledoux, D. Aurelle, N. Bensoussan, C. Marschal, J.-P. Féral, and J. Garrabou. Potential for adaptive evolution at species range margins: contrasting interactions between red coral populations and their environment in a changing ocean. *Ecology and evolution*, 5(6):1178–1192, 2015.
- [184] J.-B. Ledoux, J. Garrabou, O. Bianchimani, P. Drap, J.-P. Féral, and D. Aurelle. Fine-scale genetic structure and inferences on population biology in the threatened Mediterranean red coral, *Corallium rubrum*. *Molecular Ecology*, 19(19):4204–4216, 2010.
- [185] J.-B. Ledoux, K. Mokhtar-Jamaï, C. Roby, J.-P. Féral, J. Garrabou, and D. Aurelle. Genetic survey of shallow populations of the Mediterranean red coral [*Corallium rubrum* (Linnaeus, 1758)]: new insights into evolutionary processes shaping nuclear diversity and implications for conservation. *Molecular Ecology*, 19(4):675–690, 2010.
- [186] T. Leinonen, R. S. McCairns, R. B. O’Hara, and J. Merilä. Q_{ST} – F_{ST} comparisons: evolutionary and ecological insights from genomic heterogeneity. *Nature Reviews Genetics*, 14(3):179–190, 2013.
- [187] T. Lenormand. Gene flow and the limits to natural selection. *Trends in Ecology & Evolution*, 17(4):183–189, 2002.
- [188] R. C. Lewontin and J. Krakauer. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, 74(1):175–195, 1973.
- [189] H. Li and R. Durbin. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics*, 25(14):1754–1760, 2009.
- [190] H. Li, B. Handsaker, A. Wysoker, T. Fennell, J. Ruan, N. Homer, G. Marth, G. Abecasis, R. Durbin, and 1000 Genome Project Data Processing Sub-

- group. The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)*, 25(16):2078–2079, 2009.
- [191] L. Li, C. J. Stoeckert, and D. S. Roos. OrthoMCL: Identification of Ortholog Groups for Eukaryotic Genomes. *Genome Research*, 13(9):2178–2189, 2003.
- [192] W. Li and A. Godzik. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics*, 22(13):1658–1659, 2006.
- [193] M. T. Limborg, S. J. Helyar, M. de Bruyn, M. I. Taylor, E. E. Nielsen, R. O. B. Ogden, G. R. Carvalho, and D. Bekkevold. Environmental selection on transcriptome-derived SNPs in a high gene flow marine fish, the Atlantic herring (*Clupea harengus*). *Molecular Ecology*, 21(15):3686–3703, 2012.
- [194] C. Linares, O. Bianchimani, O. Torrents, C. Marschal, P. Drap, and J. Garrabou. Marine Protected Areas and the conservation of long-lived marine invertebrates: the Mediterranean red coral. *Marine Ecology Progress Series*, 402:69–79, 2010.
- [195] K. E. Lotterhos and M. C. Whitlock. The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*, 24(5):1031–1046, 2015.
- [196] G. Luikart, P. R. England, D. Tallmon, S. Jordan, and P. Taberlet. The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*, 4(12):981–994, 2003.
- [197] P. Lundgren, J. C. Vera, L. Peplow, S. Manel, and M. J. v. Oppen. Genotype – environment correlations in corals from the Great Barrier Reef. *BMC Genetics*, 14(1):9, 2013.
- [198] H. Lõhelaïd, T. Teder, and N. Samel. Lipxygenase-allene oxide synthase pathway in octocoral thermal stress response. *Coral Reefs*, 34(1):143–154, 2015.
- [199] S. Manel and R. Holderegger. Ten years of landscape genetics. *Trends in Ecology & Evolution*, 28(10):614–621, 2013.
- [200] S. Manel, C. Perrier, M. Pratlong, L. Abi-Rached, J. Paganini, P. Pontarotti, and D. Aurelle. Genomic resources and their influence on the detection of the

- signal of positive selection in genome scans. *Molecular Ecology*, 25(1):170–184, 2016.
- [201] A. C. Marques and A. G. Collins. Cladistic analysis of Medusozoa and cnidarian evolution. *Invertebrate Biology*, 123(1):23–42, 2004.
- [202] C. Marschal, J. Garrabou, J. G. Harmelin, and M. Pichon. A new method for measuring growth and age in the precious red coral *Corallium rubrum* (L.). *Coral Reefs*, 23(3):423–432, 2004.
- [203] Y. Martin, J. L. Bonnefont, and L. Chancerelle. Gorgonians mass mortality during the 1999 late summer in French Mediterranean coastal waters: the bacterial hypothesis. *Water Research*, 36(3):779–782, 2002.
- [204] A. Martínez-Quintana, L. Bramanti, N. Viladrich, S. Rossi, and K. Guizien. Quantification of larval traits driving connectivity: the case of *Corallium rubrum* (L. 1758). *Marine Biology*, 162(2):309–318, 2015.
- [205] A. B. Mayfield, P.-H. Chan, H. M. Putnam, C.-S. Chen, and T.-Y. Fan. The effects of a variable temperature regime on the physiology of the reef-building coral *Seriatopora hystrix*: results from a laboratory-based reciprocal transplant. *The Journal of experimental biology*, 215(23):4183–4195, 2012.
- [206] J. H. McDonald, M. Kreitman, and others. Adaptive protein evolution at the Adh locus in Drosophila. *Nature*, 351(6328):652–654, 1991.
- [207] C. S. McFadden, Y. Benayahu, E. Pante, J. N. Thoma, P. A. Nevarez, and S. C. France. Limitations of mitochondrial gene barcoding in Octocorallia. *Molecular Ecology Resources*, 11(1):19–31, 2011.
- [208] A. McKenna, M. Hanna, E. Banks, A. Sivachenko, K. Cibulskis, A. Kernyt-sky, K. Garimella, D. Altshuler, S. Gabriel, M. Daly, and M. A. DePristo. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome research*, 20(9):1297–1303, 2010.
- [209] M. Medina, A. G. Collins, J. D. Silberman, and M. L. Sogin. Evaluating hypotheses of basal animal phylogeny using complete sequences of large and small subunit rRNA. *Proceedings of the National Academy of Sciences*, 98(17):9707–9712, 2001.
- [210] E. Meyer, G. V. Aglyamova, S. Wang, J. Buchanan-Carter, D. Abrego, J. K. Colbourne, B. L. Willis, and M. V. Matz. Sequencing and de novo analysis

- of a coral larval transcriptome using 454 GSFlx. *BMC Genomics*, 10(1):219, 2009.
- [211] I. Milano, M. Babbucci, A. Cariani, M. Atanassova, D. Bekkevold, G. R. Carvalho, M. Espiñeira, F. Fiorentino, G. Garofalo, A. J. Geffen, J. H. Hansen, S. J. Helyar, E. E. Nielsen, R. Ogden, T. Patarnello, M. Stagioni, FishPopTrace Consortium, F. Tinti, and L. Bargelloni. Outlier SNP markers reveal fine-scale genetic structuring across European hake populations (*Merluccius merluccius*). *Molecular Ecology*, 23(1):118–135, 2014.
- [212] S. Mita, A.-C. Thuillet, L. Gay, N. Ahmadi, S. Manel, J. Ronfort, and Y. Vigouroux. Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, 22(5):1383–1399, 2013.
- [213] K. Mokhtar-Jamaï, M. Pascual, J.-B. Ledoux, R. Coma, J.-P. Féral, J. Garrabou, and D. Aurelle. From global to local genetic structuring in the red gorgonian *Paramuricea clavata*: the interplay between oceanographic conditions and limited larval dispersal. *Molecular Ecology*, 20(16):3291–3305, 2011.
- [214] A. Mortazavi, B. A. Williams, K. McCue, L. Schaeffer, and B. Wold. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods*, 5(7):621–628, 2008.
- [215] A. Moya, L. Huisman, E. E. Ball, D. C. Hayward, L. C. Grasso, C. M. Chua, H. N. Woo, J.-P. Gattuso, S. Forêt, and D. J. Miller. Whole transcriptome analysis of the coral *Acropora millepora* reveals complex responses to CO₂-driven acidification during the initiation of calcification. *Molecular Ecology*, 21(10):2440–2454, 2012.
- [216] A. P. Muir, R. Biek, R. Thomas, and B. K. Mable. Local adaptation with high gene flow: temperature parameters drive adaptation to altitude in the common frog (*Rana temporaria*). *Molecular Ecology*, 23(3):561–574, 2014.
- [217] P. J. Mumby, I. A. Elliott, C. M. Eakin, W. Skirving, C. B. Paris, H. J. Edwards, S. Enríquez, R. Iglesias-Prieto, L. M. Cherubin, and J. R. Stevens. Reserve design for uncertain responses of coral reefs to climate change. *Ecology Letters*, 14(2):132–140, 2011.

- [218] T. Nagylaki. Random Genetic Drift in a Cline. *Proceedings of the National Academy of Sciences of the United States of America*, 75(1):423–426, 1978. WOS:A1978EM12800093.
- [219] E. E. Nielsen, J. Hemmer-Hansen, N. A. Poulsen, V. Loeschcke, T. Moen, T. Johansen, C. Mittelholzer, G.-L. Taranger, R. Ogden, and G. R. Carvalho. Genomic signatures of local directional selection in a high gene flow marine organism; the Atlantic cod (*Gadus morhua*). *BMC Evolutionary Biology*, 9(1):276, 2009.
- [220] R. Nielsen. Molecular signatures of natural selection. *Annual Review of Genetics*, 39:197–218, 2005.
- [221] T. Nosenko, F. Schreiber, M. Adamska, M. Adamski, M. Eitel, J. Hammel, M. Maldonado, W. E. G. Müller, M. Nickel, B. Schierwater, J. Vacelet, M. Wiens, and G. Wörheide. Deep metazoan phylogeny: when different genes tell different stories. *Molecular Phylogenetics and Evolution*, 67(1):223–233, 2013.
- [222] D. M. Odorico and D. J. Miller. Internal and external relationships of the Cnidaria: implications of primary and predicted secondary structure of the 5'-end of the 23s-like rDNA. *Proceedings of the Royal Society B: Biological Sciences*, 264(1378):77–82, 1997.
- [223] T. A. Oliver and S. R. Palumbi. Do fluctuating temperature environments elevate coral thermal tolerance? *Coral Reefs*, 30(2):429–440, 2011.
- [224] M. J. H. v. Oppen, B. J. McDonald, B. Willis, and D. J. Miller. 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? *Molecular Biology and Evolution*, 18(7):1315–1329, 2001.
- [225] H.-J. Osigus, M. Eitel, M. Bernt, A. Donath, and B. Schierwater. Mitogenomics at the base of Metazoa. *Molecular Phylogenetics and Evolution*, 69(2):339–351, 2013.
- [226] C. Palaioikostas, M. Bekaert, M. G. Khan, J. B. Taggart, K. Gharbi, B. J. McAndrew, and D. J. Penman. Mapping and validation of the major sex-

- determining region in Nile tilapia (*Oreochromis niloticus* L.) using RAD sequencing. *PLoS One*, 8(7):e68389, 2013.
- [227] S. R. Palumbi, D. J. Barshis, N. Traylor-Knowles, and R. A. Bay. Mechanisms of reef coral resistance to future climate change. *Science*, 2014.
- [228] E. Pante, N. Puillandre, A. Viricel, S. Arnaud-Haond, D. Aurelle, M. Castelin, A. Chenuil, C. Destombe, D. Forcioli, M. Valero, and others. Species are hypotheses: avoid connectivity assessments based on pillars of sand. *Molecular Ecology*, 24(3):525–544, 2015.
- [229] E. Park, D.-S. Hwang, J.-S. Lee, J.-I. Song, T.-K. Seo, and Y.-J. Won. Estimation of divergence times in cnidarian evolution based on mitochondrial protein-coding genes and the fossil record. *Molecular Phylogenetics and Evolution*, 62(1):329–345, 2012.
- [230] L. M. Parker, P. M. Ross, W. A. O’Connor, L. Borysko, D. A. Raftos, and H.-O. Pörtner. Adult exposure influences offspring response to ocean acidification in oysters. *Global Change Biology*, 18(1):82–92, 2012.
- [231] S. A. Pavey, J. Gaudin, E. Normandeau, M. Dionne, M. Castonguay, C. Audet, and L. Bernatchez. RAD sequencing highlights polygenic discrimination of habitat ecotypes in the panmictic American eel. *Current Biology*, 25(12):1666–1671, 2015.
- [232] P. Pavlidis, J. D. Jensen, W. Stephan, and A. Stamatakis. A critical assessment of storytelling: gene ontology categories and the importance of validating genomic scans. *Molecular biology and evolution*, 29(10):3237–3248, 2012.
- [233] T. Perez, J. Garrabou, S. Sartoretto, J.-G. Harmelin, P. Francour, and J. Vacelet. Mortalité massive d’invertébrés marins: un événement sans précédent en Méditerranée nord-occidentale. *Comptes Rendus de l’Académie des Sciences-Series III-Sciences de la Vie*, 323(10):853–865, 2000.
- [234] G. Pertea, X. Huang, F. Liang, V. Antonescu, R. Sultana, S. Karamycheva, Y. Lee, J. White, F. Cheung, B. Parvizi, J. Tsai, and J. Quackenbush. TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics*, 19(5):651–652, 2003.

- [235] M. H. Pespeni, B. T. Barney, and S. R. Palumbi. Differences in the regulation of growth and biomineralization genes revealed through long-term common-garden acclimation and experimental genomics in the purple sea urchin. *Evolution*, 67(7):1901–1914, 2013.
- [236] M. H. Pespeni and S. R. Palumbi. The purple sea urchin genome suggests local adaptation along a latitudinal gradient despite high gene flow. In *Integrative and Comparative Biology*, volume 50, pages E136–E136. OXFORD UNIV PRESS INC JOURNALS DEPT, 2001 EVANS RD, CARY, NC 27513 USA, 2010.
- [237] M. H. Pespeni and S. R. Palumbi. Signals of selection in outlier loci in a widely dispersing species across an environmental mosaic. *Molecular Ecology*, 22(13):3580–3597, 2013.
- [238] A. Pey, J. Catanéo, D. Forcioli, P.-L. Merle, and P. Furla. Thermal threshold and sensitivity of the only symbiotic mediterranean gorgonian *Eunicella singularis* by morphometric and genotypic analyses. *Comptes rendus biologies*, 336(7):331–341, 2013.
- [239] A. Pey, J. Catanéo, D. Forcioli, P.-L. Merle, and P. Furla. Thermal threshold and sensitivity of the only symbiotic Mediterranean gorgonian *Eunicella singularis* by morphometric and genotypic analyses. *Comptes rendus biologies*, 336(7):331–341, 2013.
- [240] A. Pey, T. Zamoum, D. Allemand, P. Furla, and P.-L. Merle. Depth-dependant thermotolerance of the symbiotic Mediterranean gorgonian *Eunicella singularis*: evidence from cellular stress markers. *Journal of Experimental Marine Biology and Ecology*, 404(1):73–78, 2011.
- [241] M. Pigliucci. *Phenotypic plasticity: beyond nature and nurture*. JHU Press, 2001.
- [242] I. D. Pivotto, D. Nerini, M. Masmoudi, H. Kara, L. Chaoui, and D. Aurelle. Highly contrasted responses of Mediterranean octocorals to climate change along a depth gradient. *Royal Society open science*, 2(5):140493, 2015.
- [243] S. Planes and P. Romans. Evidence of genetic selection for growth in new recruits of a marine fish. *Molecular Ecology*, 13(7):2049–2060, 2004.

- [244] M. Plummer, N. Best, K. Cowles, and K. Vines. CODA: Convergence diagnosis and output analysis for MCMC. *R news*, 6(1):7–11, 2006.
- [245] N. R. Polato, J. C. Vera, and I. B. Baums. Gene discovery in the threatened elkhorn coral: 454 sequencing of the *Acropora palmata* transcriptome. *PLoS ONE*, 6(12):e28634, 2011.
- [246] B. N. Poncet, D. Herrmann, F. Gugerli, P. Taberlet, R. Holderegger, L. Gielly, D. Rioux, W. Thuiller, S. Aubert, and S. Manel. Tracking genes of ecological relevance using a genome scan in two independent regional population samples of *Arabis alpina*. *Molecular Ecology*, 19(14):2896–2907, 2010.
- [247] D. C. Potts. Natural selection in experimental populations of reef-building corals (Scleractinia). *Evolution*, pages 1059–1078, 1984.
- [248] D. A. Powers and P. M. Schulte. Evolutionary adaptations of gene structure and expression in natural populations in relation to a changing environment: a multidisciplinary approach to address the million-year saga of a small fish. *The Journal of experimental zoology*, 282(1-2):71–94, 1998.
- [249] C. Prada, N. V. Schizas, and P. M. Yoshioka. Phenotypic plasticity or speciation? A case from a clonal marine organism. *BMC Evolutionary Biology*, 8(1):1, 2008.
- [250] M. Pratlong, A. Haguénauer, O. Chabrol, C. Klopp, P. Pontarotti, and D. Aurelle. The red coral (*Corallium rubrum*) transcriptome: a new resource for population genetics and local adaptation studies. *Molecular Ecology Resources*, 15(5):1205–1215, 2015.
- [251] M. N. Price, P. S. Dehal, and A. P. Arkin. FastTree: computing large minimum evolution trees with profiles instead of a distance matrix. *Molecular Biology and Evolution*, 26(7):1641–1650, 2009.
- [252] C. Priori, F. Erra, M. Angiolillo, and G. Santangelo. Effects of gastropod predation on the reproductive output of an overexploited deep octocoral. *Coral Reefs*, 34(1):59–63, 2015.
- [253] J. K. Pritchard and A. Di Rienzo. Adaptation—not by sweeps alone. *Nature Reviews Genetics*, 11(10):665–667, 2010.
- [254] J. K. Pritchard, M. Stephens, and P. Donnelly. Inference of population structure using multilocus genotype data. *Genetics*, 155(2):945–959, 2000.

- [255] H. M. Putnam, J. M. Davidson, and R. D. Gates. Ocean acidification influences host DNA methylation and phenotypic plasticity in environmentally susceptible corals. *Evolutionary Applications*, 2016.
- [256] H. M. Putnam and R. D. Gates. Preconditioning in the reef-building coral *Pocillopora damicornis* and the potential for trans-generational acclimatization in coral larvae under future climate change conditions. *Journal of Experimental Biology*, 218(15):2365–2372, 2015.
- [257] S. Qiu, R. Bergero, S. Guirao-Rico, J. L. Campos, T. Cezard, K. Gharbi, and D. Charlesworth. RAD mapping reveals an evolving, polymorphic and fuzzy boundary of a plant pseudoautosomal region. *Molecular Ecology*, 25(1):414–430, 2016.
- [258] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2016.
- [259] V. Ravigné, U. Dieckmann, and I. Olivieri. Live where you thrive: joint evolution of habitat choice and local adaptation facilitates specialization and promotes diversity. *The American Naturalist*, 174(4):E141–E169, 2009.
- [260] A. J. Reft and M. Daly. Morphology, distribution, and evolution of apical structure of nematocysts in hexacorallia. *Journal of Morphology*, 273(2):121–136, 2012.
- [261] C. Rellstab, F. Gugerli, A. J. Eckert, A. M. Hancock, and R. Holderegger. A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*, 24(17):4348–4370, 2015.
- [262] T. B. Reusch and T. E. Wood. Molecular ecology of global change. *Molecular Ecology*, 16(19):3973–3992, 2007.
- [263] T. B. H. Reusch. Climate change in the oceans: evolutionary versus phenotypically plastic responses of marine animals and plants. *Evolutionary Applications*, 7(1):104–122, 2014.
- [264] P. Rice, I. Longden, and A. Bleasby. EMBOSS: The European Molecular Biology Open Software Suite. *Trends in Genetics*, 16(6):276–277, 2000.
- [265] L. H. Rieseberg, J. Whitton, and C. R. Linder. Molecular marker incongruence in plant hybrid zones and phylogenetic trees. *Acta Botanica Neerlandica*, 45(3):243–262, 1996.

- [266] M. D. Robinson, D. J. McCarthy, and G. K. Smyth. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26(1):139–140, 2010.
- [267] M. Rodriguez-Lanetty, W. S. Phillips, and V. M. Weis. Transcriptome analysis of a cnidarian – dinoflagellate mutualism reveals complex modulation of host gene expression. *BMC Genomics*, 7(1):23, 2006.
- [268] J. Romiguier, P. Gayral, M. Ballenghien, A. Bernard, V. Cahais, A. Chenuil, Y. Chiari, R. Dernat, L. Duret, N. Faivre, E. Loire, J. M. Lourenco, B. Nabholz, C. Roux, G. Tsagkogeorga, A. a.-T. Weber, L. A. Weinert, K. Belkhir, N. Bierne, S. Glémin, and N. Galtier. Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature*, 515(7526):261–263, 2014.
- [269] F. Ronquist, M. Teslenko, P. v. d. Mark, D. L. Ayres, A. Darling, S. Höhna, B. Larget, L. Liu, M. A. Suchard, and J. P. Huelsenbeck. MrBayes 3.2: efficient bayesian phylogenetic inference and model choice across a large model space. *Systematic Biology*, 61(3):539–542, 2012.
- [270] S. Rossi, G. Tsounis, C. Orejas, T. Padrón, J.-M. Gili, L. Bramanti, N. Teixidó, and J. Gutt. Survey of deep-dwelling red coral (*Corallium rubrum*) populations at Cap de Creus (NW Mediterranean). *Marine Biology*, 154(3):533–545, 2008.
- [271] F. Rousset. genepop’007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources*, 8(1):103–106, 2008.
- [272] J. F. Ryan, K. Pang, C. E. Schnitzler, A.-D. Nguyen, R. T. Moreland, D. K. Simmons, B. J. Koch, W. R. Francis, P. Havlak, S. A. Smith, N. H. Putnam, S. H. D. Haddock, C. W. Dunn, T. G. Wolfsberg, J. C. Mullikin, M. Q. Martindale, and A. D. Baxevanis. The genome of the ctenophore *Mnemiopsis leidyi* and its implications for cell type evolution. *Science*, 342(6164), 2013.
- [273] S. Salinas and S. B. Munch. Thermal legacies: transgenerational effects of temperature on growth in a vertebrate. *Ecology Letters*, 15(2):159–163, 2012.

- [274] L. V. Salvini-Plawen. On the origin and evolution of the lower Metazoa. *Journal of Zoological Systematics and Evolutionary Research*, 16(1):40–87, 1978.
- [275] J. Sambrook, E. F. Fritsch, T. Maniatis, and others. *Molecular cloning*, volume 2. Cold spring harbor laboratory press New York, 1989.
- [276] E. Sanford and M. W. Kelly. Local adaptation in marine invertebrates. *Annual Review of Marine Science*, 3(1):509–535, 2011.
- [277] G. Santangelo, M. Abbiati, and G. Caforio. Struttura di età e dinamica di popolazione di *Corallium rubrum* (L.). *Il corallo rosso in Mediterraneo: Arte, Storia e Scienza (Red Coral in the Mediterranean Sea: Art, History and Science)*, pages 131–157, 1993.
- [278] G. Santangelo, L. Bongiorno, F. Giannini, M. Abbiati, and G. Buffoni. Structure analysis of two red coral populations dwelling in different habitats. *Ci-gogna F. BG, Cattaneo-Vietti R.(eds) Red coral and other Mediterranean octocorals: biology and protection*, pages 23–43, 1997.
- [279] G. Santangelo, L. Bramanti, and M. Iannelli. Population dynamics and conservation biology of the over-exploited Mediterranean red coral. *Journal of Theoretical Biology*, 244(3):416–423, 2007.
- [280] G. Santangelo, L. Bramanti, S. Rossi, G. Tsounis, I. Vielmini, C. Lott, and J. M. Gili. Patterns of variation in recruitment and post-recruitment processes of the Mediterranean precious gorgonian coral *Corallium rubrum*. *Journal of Experimental Marine Biology and Ecology*, 411:7–13, 2012.
- [281] G. Santangelo, E. Carletti, E. Maggi, and L. Bramanti. Reproduction and population sexual structure of the overexploited Mediterranean red coral *Corallium rubrum*. *Marine Ecology Progress Series*, 248:99–108, 2003.
- [282] P. A. Saunders, T. Franco, C. Sottas, T. Maurice, G. Ganem, and F. Veyrunes. Masculinised behaviour of XY females in a mammal with naturally occurring sex reversal. *Scientific Reports*, 6, 2016.
- [283] C. Sauvage, N. Bierne, S. Lapègue, and P. Boudry. Single Nucleotide polymorphisms and their relationship to codon usage bias in the Pacific oyster *Crassostrea gigas*. *Gene*, 406(1–2):13–22, 2007.

- [284] O. Savolainen, M. Lascoux, and J. Merilä. Ecological genomics of local adaptation. *Nature Reviews Genetics*, 14(11):807–820, 2013.
- [285] M. H. Schulz, D. R. Zerbino, M. Vingron, and E. Birney. Oases: robust de novo RNA-seq assembly across the dynamic range of expression levels. *Bioinformatics*, 28(8):1086–1092, 2012.
- [286] Z. Shao, S. Graf, O. Y. Chaga, and D. V. Lavrov. Mitochondrial genome of the moon jelly *Aurelia aurita* (Cnidaria, Scyphozoa): a linear DNA molecule encoding a putative DNA-dependent DNA polymerase. *Gene*, 381:92–101, 2006.
- [287] T. L. Shearer, M. J. H. van Oppen, S. L. Romano, and G. Wörheide. Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria). *Molecular Ecology*, 11(12):2475–2487, 2002.
- [288] C. D. H. Sherman and D. J. Ayre. Fine-scale adaptation in a clonal sea anemone. *Evolution; International Journal Of Organic Evolution*, 62(6):1373–1380, 2008.
- [289] H. Shimodaira. An approximately unbiased test of phylogenetic tree selection. *Systematic Biology*, 51(3):492–508, 2002.
- [290] H. Shimodaira and M. Hasegawa. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Molecular biology and evolution*, 16:1114–1116, 1999.
- [291] H. Shimodaira and M. Hasegawa. CONSEL: for assessing the confidence of phylogenetic tree selection. *Bioinformatics*, 17(12):1246–1247, 2001.
- [292] S. Siebert and C. E. Juliano. Sex, polyps, and medusae: Determination and maintenance of sex in cnidarians. *Molecular Reproduction and Development*, 2016.
- [293] M. Slatkin and T. Maruyama. Genetic drift in a cline. *Genetics*, 81(1):209–222, 1975.
- [294] L. W. Smith, D. Barshis, and C. Birkeland. Phenotypic plasticity for skeletal growth, density and calcification of *Porites lobata* in response to habitat type. *Coral Reefs*, 26(3):559–567, 2007.

- [295] G. N. Somero. The physiology of climate change: how potentials for acclimatization and genetic adaptation will determine ‘winners’ and ‘losers’. *The Journal of experimental biology*, 213(6):912–920, 2010.
- [296] G. N. Somero. The physiology of global change: linking patterns to mechanisms. *Annual Review of Marine Science*, 4:39–61, 2012.
- [297] J.-I. Song and J. H. Won. Systematic relationship of the anthozoan orders based on the partial nuclear 18s rDNA sequences. *Korean Journal of Biological Sciences*, 1(1):43–52, 1997.
- [298] V. Soria-Carrasco, Z. Gompert, A. A. Comeault, T. E. Farkas, T. L. Parchman, J. S. Johnston, C. A. Buerkle, J. L. Feder, J. Bast, T. Schwander, and others. Stick insect genomes reveal natural selection’s role in parallel speciation. *Science*, 344(6185):738–742, 2014.
- [299] A. Stamatakis. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, 30(9):1312–1313, 2014.
- [300] S. N. Stampar, M. M. Maronna, M. V. Kitahara, J. D. Reimer, and A. C. Morandini. Fast-evolving mitochondrial DNA in Ceriantharia: a reflection of Hexacorallia paraphyly? *PLoS ONE*, 9(1), 2014.
- [301] J. Stanton-Geddes, J. B. Yoder, R. Briskine, N. D. Young, and P. Tiffin. Estimating heritability using genomic data. *Methods in Ecology and Evolution*, 4(12):1151–1158, 2013.
- [302] M. Steel, P. Lockhart, and P. Penny. Confidence in evolutionary trees from biological data. *Nature*, 364:440–442, 1993.
- [303] D. J. Stefanik, T. J. Lubinski, B. R. Granger, A. L. Byrd, A. M. Reitzel, L. DeFilippo, A. Lorenc, and J. R. Finnerty. Production of a reference transcriptome and transcriptomic database (EdwardsiellaBase) for the lined sea anemone, *Edwardsiella lineata*, a parasitic cnidarian. *BMC Genomics*, 15(1):71, 2014.
- [304] J. C. Sullivan, A. M. Reitzel, and J. R. Finnerty. Upgrades to StellaBase facilitate medical and genetic studies on the starlet sea anemone, *Nematostella vectensis*. *Nucleic Acids Research*, 36(suppl 1):D607–D611, 2008.

- [305] D. L. Swofford. *Paup 4.0 for Macintosh: Phylogenetic Analysis Using Parsimony (Software and User's Book for Macintosh)*. Sinauer Associates, Incorporated, 2004.
- [306] M. Szulkin, P.-A. Gagnaire, N. Bierne, and A. Charmantier. Population genomic footprints of fine-scale differentiation between habitats in Mediterranean blue tits. *Molecular Ecology*, 25(2):542–558, 2016.
- [307] F. Tajima. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 123(3):585–595, 1989.
- [308] S. Tarazona, F. García-Alcalde, J. Dopazo, A. Ferrer, and A. Conesa. Differential expression in RNA-seq: A matter of depth. *Genome Research*, 21(12):2213–2223, 2011.
- [309] R. D. C. Team. *R 2.14. 1*. R Project for Statistical Computing Vienna, Austria, 2011.
- [310] V. Teaniniuraitemoana, A. Huvet, P. Levy, C. Klopp, E. Lhuillier, N. Gaertner-Mazouni, Y. Gueguen, and G. L. Moullac. Gonad transcriptome analysis of pearl oyster *Pinctada margaritifera*: identification of potential sex differentiation and sex determining genes. *BMC Genomics*, 15(1):491, 2014.
- [311] U. Technau, S. Rudd, P. Maxwell, P. M. K. Gordon, M. Saina, L. C. Grasso, D. C. Hayward, C. W. Sensen, R. Saint, T. W. Holstein, E. E. Ball, and D. J. Miller. Maintenance of ancestral complexity and non-metazoan genes in two basal cnidarians. *Trends in Genetics*, 21(12):633–639, 2005.
- [312] C. K. Tepolt and S. R. Palumbi. Transcriptome sequencing reveals both neutral and adaptive genome dynamics in a marine invader. *Molecular Ecology*, 24(16):4145–4158, 2015.
- [313] The UniProt Consortium. Activities at the universal protein resource (UniProt). *Nucleic Acids Research*, 42(D1):D191–D198, 2013.
- [314] O. Torrents and J. Garrabou. Fecundity of red coral *Corallium rubrum* (L.) populations inhabiting in contrasting environmental conditions in the NW Mediterranean. *Marine Biology*, 158(5):1019–1028, 2011.
- [315] O. Torrents, J. Garrabou, C. Marschal, and J. G. Harmelin. Age and size at first reproduction in the commercially exploited red coral *Corallium rubrum*

- (L.) in the Marseilles area (France, NW Mediterranean). *Biological Conservation*, 121(3):391–397, 2005.
- [316] O. Torrents, E. Tambutté, N. Caminiti, and J. Garrabou. Upper thermal thresholds of shallow vs. deep populations of the precious Mediterranean red coral *Corallium rubrum* (L.): Assessing the potential effects of warming in the NW Mediterranean. *Journal of Experimental Marine Biology and Ecology*, 357(1):7–19, 2008.
- [317] N. Traylor-Knowles and S. R. Palumbi. Translational environmental biology: cell biology informing conservation. *Trends in Cell Biology*, 24(5):265–267, 2014.
- [318] N. G. Traylor-Knowles, E. G. Kane, V. Sombatsaphay, J. R. Finnerty, and A. M. Reitzel. Sex-specific and developmental expression of Dmrt genes in the starlet sea anemone, *Nematostella vectensis*. *EvoDevo*, 6:13, 2015.
- [319] G. Tsounis, S. Rossi, M. Aranguren, J.-M. Gili, and W. Arntz. Effects of spatial variability and colony size on the reproductive output and gonadal development cycle of the Mediterranean red coral (*Corallium rubrum* L.). *Marine Biology*, 148(3):513–527, 2006.
- [320] G. Tsounis, S. Rossi, J.-M. Gili, and W. Arntz. Population structure of an exploited benthic cnidarian: the case study of red coral (*Corallium rubrum* L.). *Marine Biology*, 149(5):1059–1070, 2006.
- [321] G. Tsounis, S. Rossi, J.-M. Gili, and W. E. Arntz. Red coral fishery at the Costa Brava (NW Mediterranean): case study of an overharvested precious coral. *Ecosystems*, 10(6):975–986, 2007.
- [322] K. E. Ulstrup and M. J. H. Van Oppen. Geographic and habitat partitioning of genetically distinct zooxanthellae (*Symbiodinium*) in *Acropora* corals on the Great Barrier Reef. *Molecular Ecology*, 12(12):3477–3484, 2003.
- [323] H. Van Iten, A. C. Marques, J. d. M. Leme, M. L. A. F. Pacheco, and M. G. Simões. Origin and early diversification of the phylum Cnidaria Verrill: major developments in the analysis of the taxon’s Proterozoic–Cambrian history. *Palaeontology*, 57(4):677–690, 2014.

- [324] L. Verstrepen, I. Carpentier, K. Verhelst, and R. Beyaert. ABINs: A20 binding inhibitors of NF- κ B and apoptosis signaling. *Biochemical Pharmacology*, 78(2):105–114, 2009.
- [325] J. Vidal-Dupiol, N. M. Dheilly, R. Rondon, C. Grunau, C. Cosseau, K. M. Smith, M. Freitag, M. Adjero, and G. Mitta. Thermal stress triggers broad *Pocillopora damicornis* transcriptomic remodeling, while *Vibrio coralliilyticus* infection induces a more targeted immuno-suppression response. *PloS one*, 9(9):e107672, 2014.
- [326] J. Vidal-Dupiol, D. Zoccola, E. Tambutté, C. Grunau, C. Cosseau, K. M. Smith, M. Freitag, N. M. Dheilly, D. Allemand, and S. Tambutté. Genes related to ion-transport and energy production are upregulated in response to CO₂-driven pH decrease in corals: new insights from transcriptome analysis. *PLoS ONE*, 8(3):e58652, 2013.
- [327] A. R. Wallace. *La sélection naturelle: essais*. C. Reinwald, 1872.
- [328] A. R. Wallace. On the tendency of varieties to depart indefinitely from the original type. *Journal of the Proceedings of the Linnean Society of London. Zoology*, 3:53–62, 1958.
- [329] P. J. Walsh and G. N. Somero. Temperature adaptation in sea anemones: physiological and biochemical variability in geographically separate populations of *Metridium senile*. *Marine Biology*, 62(1):25–34, 1981.
- [330] G.-R. Walther, E. Post, P. Convey, A. Menzel, C. Parmesan, T. J. Beebee, J.-M. Fromentin, O. Hoegh-Guldberg, and F. Bairlein. Ecological responses to recent climate change. *Nature*, 416(6879):389–395, 2002.
- [331] J. Wang, D.-X. Xue, B.-D. Zhang, Y.-L. Li, B.-J. Liu, and J.-X. Liu. Genome-wide SNP discovery, genotyping and their preliminary applications for population genetic inference in spotted sea bass (*Lateolabrax maculatus*). *PloS one*, 11(6):e0157809, 2016.
- [332] L. Wang, S. Liu, Z. Zhuang, L. Guo, Z. Meng, and H. Lin. Population genetic studies revealed local adaptation in a high gene-flow marine fish, the small yellow croaker (*Larimichthys polyactis*). *PloS one*, 8(12):e83493, 2013.

- [333] S. Weinberg. The light-dependent behaviour of planula larvae of *Eunicella singularis* and *Corallium rubrum* and its implication for octocorallian ecology. *Bijdragen tot de Dierkunde*, 49(1):16–30, 1979.
- [334] B. S. Weir and C. C. Cockerham. Estimating f-statistics for the analysis of population structure. *Evolution*, pages 1358–1370, 1984.
- [335] V. M. Weis. Cellular mechanisms of cnidarian bleaching: stress causes the collapse of symbiosis. *Journal of Experimental Biology*, 211(19):3059–3066, 2008.
- [336] V. M. Weis. The susceptibility and resilience of corals to thermal stress: adaptation, acclimatization or both? *Molecular Ecology*, 19(8):1515–1517, 2010.
- [337] W. I. Weis, M. E. Taylor, and K. Drickamer. The C-type lectin superfamily in the immune system. *Immunological Reviews*, 163(1):19–34, 1998.
- [338] J. Won, B. Rho, and J. Song. A phylogenetic study of the Anthozoa (phylum Cnidaria) based on morphological and molecular characters. *Coral Reefs*, 20(1):39–50, 2001.
- [339] E. M. Wood-Charlson, L. L. Hollingsworth, D. A. Krupp, and V. M. Weis. Lectin/glycan interactions play a role in recognition in a coral/dinoflagellate symbiosis. *Cellular Microbiology*, 8(12):1985–1993, 2006.
- [340] A.-L. Wrange, C. André, T. Lundh, U. Lind, A. Blomberg, P. J. Jonsson, and J. N. Havenhand. Importance of plasticity and local adaptation for coping with changing salinity in coastal areas: a test case with barnacles in the Baltic Sea. *BMC evolutionary biology*, 14(1):1, 2014.
- [341] A. E. Wright, R. Dean, F. Zimmer, and J. E. Mank. How to make a sex chromosome. *Nature Communications*, 7, 2016.
- [342] S. Wright. Evolution in Mendelian populations. *Genetics*, 16(2):97–159, 1931.
- [343] X. Xia and P. Lemey. Assessing substitution saturation with DAMBE. In *The Phylogenetic Handbook*. Cambridge University Press, second edition edition, 2009.
- [344] X. Xia and Z. Xie. DAMBE: software package for Data Analysis in Molecular Biology and Evolution. *Journal of Heredity*, 92(4):371–373, 2001.

- [345] X. Xia, Z. Xie, M. Salemi, L. Chen, and Y. Wang. An index of substitution saturation and its application. *Molecular Phylogenetics and Evolution*, 26(1):1–7, 2003.
- [346] J. Yang, B. Benyamin, B. P. McEvoy, S. Gordon, A. K. Henders, D. R. Nyholt, P. A. Madden, A. C. Heath, N. G. Martin, G. W. Montgomery, and others. Common SNPs explain a large proportion of the heritability for human height. *Nature genetics*, 42(7):565–569, 2010.
- [347] F. Zapata, F. E. Goetz, S. A. Smith, M. Howison, S. Siebert, S. H. Church, S. M. Sanders, C. L. Ames, C. S. McFadden, S. C. France, M. Daly, A. G. Collins, S. H. D. Haddock, C. W. Dunn, and P. Cartwright. Phylogenomic analyses support traditional relationships within cnidaria. *PLoS ONE*, 10(10):e0139068, 2015.
- [348] M. Ziegler, C. M. Roder, C. Büchel, and C. R. Voolstra. Limits to physiological plasticity of the coral *Pocillopora verrucosa* from the central Red Sea. *Coral Reefs*, 33(4):1115–1129, 2014.

Abstract

The on-going climate change should have major impacts on the different components of biodiversity, from genes to ecosystems. The Mediterranean red coral (*Corallium rubrum*) has an important ecological role in Mediterranean benthic ecosystems and is submitted to major anthropic pressures because of its direct (exploitation) and indirect (attractivity for recreational scuba-diving) economical values. This species has recently been impacted by mortality events linked with thermal anomalies. Because of the extremely contrasted thermal conditions it deals with along his range the red coral is an interesting model for the study of local adaptation. During my thesis, I tried to understand the genetic basis of thermotolerance differences in the red coral.

We first confirmed that gene expression differences were maintained along time, in absence of thermal stress in individuals from different depths in Marseille. Some of these genes presented relevant functions which were already described in the stress response in cnidaria and were good candidates for the study of local adaptation and strong arguments supporting the conservation of this function in cnidaria.

In order to identify potential genetic basis of the local adaptation in the red coral, we built a sampling design of pairs of ‘shallow vs deep’ populations in three geographical regions and sequenced via RAD-Sequencing the corresponding individuals. The analysis of neutral genetic structure of the studied populations highlighted a limited connectivity of shallow populations with both shallow populations and deep populations that could counteract recolonization abilities of population the most exposed to global change. Several methodological obstacles have been met in the detection of loci under selection in such strongly structured species. By keeping in mind these potential biases, we highlighted a potential signal of local adaptation in Marseille and Corsica, without any convergence in candidates genes and functions. The analysis of the genetic structure of the red coral led us to the identification of sex genetic markers which is the first description in a non-bilaterian species. This observation confirms that the red coral is an interesting model for evolutionary and integrative biology.

In conclusion, this thesis completes ecological and biological genetic knowledge we have for the red coral. Here we provide a first study of local adaptation and genetic sex determination in this species. New approaches will be implemented to further explore the local adaptation of red coral to depth, particularly in light of a genome, or a more focused view of the phenotype.

Key words: Red coral, Selection, Local adaptation, Climate change, Genetic determinism of sex, Connectivity

Résumé

Le changement climatique actuel affecte déjà différents composants de la biodiversité, des gènes aux écosystèmes. En Méditerranée, le corail rouge *Corallium rubrum* présente un important rôle écologique dans les écosystèmes benthiques de substrat dur. Il subit d'importantes pressions anthropiques de par sa valeur économique directe (exploitation), et indirecte (attire pour la plongée récréative). Il a récemment été touché par des événements de mortalité liés à des anomalies thermiques. Par ailleurs, les conditions environnementales, et notamment thermiques extrêmement contrastées que cette espèce subit sur l'ensemble de son aire de répartition, en font un modèle intéressant pour l'étude de l'adaptation locale. J'ai donc cherché, au cours de ma thèse, à comprendre les bases génétiques des différences de thermotolérance chez le corail rouge.

Dans un premier temps, nous avons confirmé que les différences d'expressions de gènes étaient maintenues au cours du temps, en absence de stress thermique chez des individus issus de profondeurs différentes à Marseille. Certains des gènes identifiés, présentant des fonctions pertinentes et déjà décrites pour la réponse au stress chez les cnidaires, sont de bons candidats pour l'étude de l'adaptation locale et de forts arguments en faveur de la conservation de cette fonction chez les cnidaires.

Afin d'identifier d'éventuelles bases génétiques de l'adaptation locale chez le corail rouge, nous avons mis en place un protocole d'échantillonnage de paires de populations 'surface vs profondeurs' dans trois régions géographiques différentes suivi d'un séquençage via RAD-Séquençage. L'analyse de la structure génétique neutre indique une connectivité réduite entre les populations de surface et à fois populations de surface et les populations profondes qui pourrait limiter les capacités de recolonisation des populations les plus exposées aux pressions du changement global. En ce qui concerne les locus sous sélection, plusieurs obstacles méthodologiques ont été rencontrés dans leur détection chez cette espèce à forte structure génétique. En gardant à l'esprit ces biais possibles, nous avons identifié un signal probable d'adaptation locale à Marseille et en Corse, sans qu'une convergence dans les gènes ou les fonctions candidats n'ait été observée. L'analyse de la structure génétique chez le corail rouge a conduit à l'identification de marqueurs génétiques du sexe qui correspond à la première description chez une espèce non-bilatérienne. Cette observation confirme l'intérêt du corail rouge en tant que modèle en biologie évolutive et intégrative.

En conclusion, cette thèse complète les connaissances biologiques écologiques et génétiques que nous avons chez l'espèce emblématique qu'est le corail rouge. Nous apportons ici une première étude de l'adaptation locale et du déterminisme génétique du sexe chez cette espèce. De nouvelles approches pourront être mises en place pour continuer à explorer l'adaptation locale du corail rouge à la profondeur, notamment à la lumière d'un génome, ou d'un point de vue plus axé sur le phénotype.

Mots clés : Corail rouge, Sélection, Adaptation locale, Changement climatique, Déterminisme génétique du sexe, Connectivité